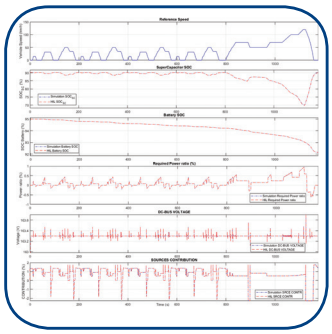
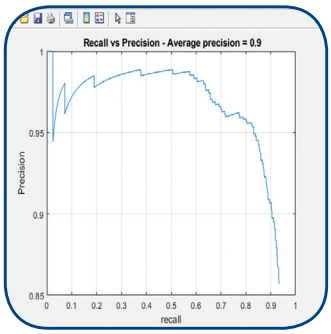
Volume 14, Number 3, 2023 ISSN 1847-6996

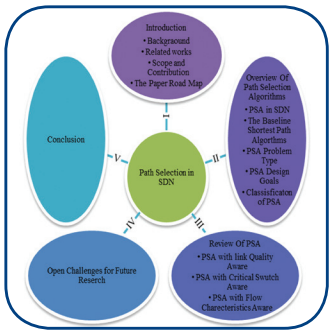
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

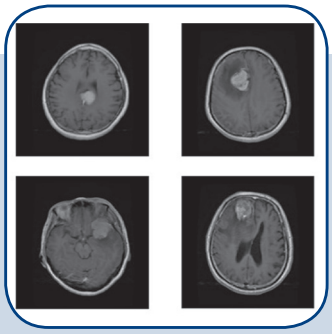
IJECES International Journal of Electrical and Computer Engineering Systems

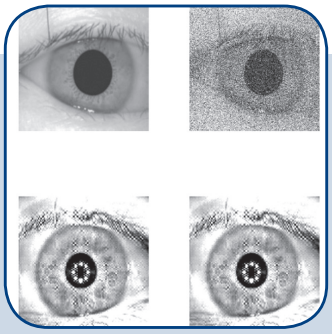
International Journal of Electrical and Computer Engineering Systems

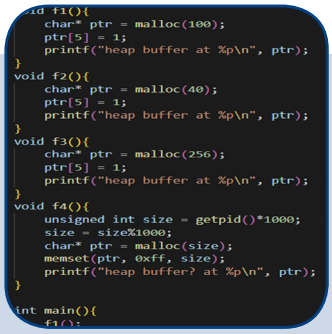












ISSN: 1847-6996

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek, Josip Juraj Strossmayer University of Osijek, Croatia

Osijek, Croatia | Volume 14, Number 3, 2023 | Pages 241 - 370

The International Journal of Electrical and Computer Engineering Systems is published with the financial support of the Ministry of Science and Education of the Republic of Croatia

CONTACT

International Journal of Electrical and Computer Engineering Systems (IJECES)

Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b, 31000 Osijek, Croatia
Phone: +38531224600, Fax: +38531224605
e-mail: ijeces@ferit.hr

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.
Giro account: 2390001 - 1100016777,
Croatian Postal Bank

Journal is referred in:

- Scopus
- Web of Science Core Collection (Emerging Sources Citation Index - ESCI)
- Google Scholar
- CiteFactor
- Genamics
- Hrčak
- Ulrichweb
- Reaxys
- Embase
- Engineering Village

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003
Published: quarterly
Circulation: 300

IJECES online

https://ijeces.ferit.hr

TABLE OF CONTENTS

Breast Cancer Classification by Gene Expression Analysis using Hybrid Feature Selection and Hyper-heuristic Adaptive Universum Support Vector Machine

Original Scientific Paper

V. Murugesan
Department of Computer Science, VLB Janakiammal College of Arts and Science, Coimbatore,
Tamil Nadu- 641042, India
murugeshvlb2020@gmail.com

P. Balamurugan
Department of Computer Science, Government Arts College, Coimbatore,
Tamil Nadu-641018, India
spbalamurugan@rediffmail.com

Abstract *– Comprehensive assessments of the molecular characteristics of breast cancer from gene expression patterns can aid in the early identification and treatment of tumor patients. The enormous scale of gene expression data obtained through microarray sequencing increases the difficulty of training the classifier due to large-scale features. Selecting pivotal gene features can minimize high dimensionality and the classifier complexity with improved breast cancer detection accuracy. However, traditional filter and wrapper-based selection methods have scalability and adaptability issues in handling complex gene features. This paper presents a hybrid feature selection method of Mutual Information Maximization - Improved Moth Flame Optimization (MIM-IMFO) for gene selection along with an advanced Hyper-heuristic Adaptive Universum Support classification model Vector Machine (HH-AUSVM) to improve cancer detection rates. The hybrid gene selection method is developed by performing filter-based selection using MIM in the first stage followed by the wrapper method in the second stage, to obtain the pivotal features and remove the inappropriate ones. This method improves standard MFO by a hybrid exploration/exploitation phase to accomplish a better trade-off between exploration and exploitation phases. The classifier HH-AUSVM is formulated by integrating the Adaptive Universum learning approach to the hyper-heuristics-based parameter optimized SVM to tackle the class samples imbalance problem. Evaluated on breast cancer gene expression datasets from Mendeley Data Repository, this proposed MIM-IMFO gene selection-based HH-AUSVM classification approach provided better breast cancer detection with high accuracies of 95.67%, 96.52%, 97.97% and 95.5% and less processing time of 4.28, 3.17, 9.45 and 6.31 seconds, respectively.*

Keywords*: Gene expression analysis, Breast cancer, hybrid gene selection, Mutual Information Maximization, Improved Moth Flame Optimization, Support Vector Machine, Adaptive Universum learning, Hyper-heuristic algorithm*

1. INTRODUCTION

Breast cancer is widespread among women with a high global death rate and has multiple causes, including genetic and hereditary factors [1]. The genetic factors can be estimated only through learning the gene expression data for molecular analysis of breast cancer pathogenesis. This data is a part of the cancer transcriptome, including the different RNA sequencing data types. The transcriptome of an organism is measurable using RNA-seq or DNA microarrays. The molecular analysis of this genetic information from DNA can provide all the information about the features and functions of all the body cells [2]. The genes also provide the vital specification of the phenotypes, which can be identified by analyzing the gene expression profiles of all diseased tissues and healthy tissues for obtaining the genetic variables for the pathological process [3]. The gene expression data can provide information about the cancer cells, the impact of drugs on the tissues and genetic variations in the diseased cells. Therefore, the gene expression data helps obtain the different features associated with breast cancer, which can be analyzed using advanced computational methods to

identify the gene targets, detect the disease's presence and develop suitable drugs [4]. Many studies have used gene expression data for obtaining deep tumor characteristics, which provide options for treating, caring and monitoring cancer patients. Detection of the genes more highly expressive of the tumor characteristics than the normal cell characteristics is often challenging when selecting the best computation method [5]. The gene expression data analysis also has challenges, such as high dimensionality of gene features, moderately smaller sample size and a higher noise ratio.

Numerous studies have utilized supervised and unsupervised learning systems for cancer identification from gene expression data. The unsupervised category includes cluster-based methods and decision tree classifiers, while the supervised category includes statistical and machine learning (ML) algorithms [6]. ML algorithms are predominantly utilized for the classification of disease data. They have often produced efficient results using the algorithms such as support vector machines (SVM), Random forests (RF), artificial neural networks (ANN), etc. The latest studies have used deep learning (DL) methods, the so-called complicated algorithms of the ML family, for cancer classification tasks. Algorithms such as Deep Neural Networks (DNN), Convolutional Neural networks (CNN), etc., have better learning rates and improved deterministic powers than ML algorithms. Still, these algorithms require more training data to learn the deep features. They are also mostly limited by the high dimensionality and the sparse sample size of the gene expression data [7]. Considering such limitations of DL methods and the extensive research still needed to integrate them for the genomics data analysis, ML algorithms are suggested for breast cancer classification from gene expression data.

SVM has provided better performance for breast cancer classification with reduced training and testing time [8]. Yet, SVM also has its share of limitations, namely the limited ability to handle high dimensions, underperformance when the target classes are overlapping or unbalanced, and, most importantly, the parameters of SVM do not adapt automatically to the given problem [9]. This paper has focused on developing an efficient breast cancer classification approach by improving the SVM classifier's performance and establishing a hybrid algorithm for improved gene selection so that the high-dimension problem is solved. The proposed approach has developed a feature selection method in which the filter-based method of MIM is combined with the wrapper method of IMFO to form the MIM-IMFO method. This proposed approach is based on two stages; the first is selecting the most important gene features using Mutual information. The second stage reduces the irrelevant features by the optimal feature subset selection using IMFO. Here, IMFO is developed by introducing a hybrid exploration/exploitation phase to the standard MFO [10] to achieve a good trade-off between the exploration and exploitation phases. The proposed classifier HH-AUSVM is an improved model of SVM in which the Adaptive Universum (AU) learning approach is applied to provide prior knowledge adaptively about the optimal classification problem and minimize the class imbalance problem. The Universum samples are data added to the imbalance classes as false data to balance the data distribution for easy computation without impacting the final output. Additionally, the SVM parameters are tuned using hyper-heuristics to form optimal SVM configuration with high accuracy and reduced model complexity. Evaluation of the MIM-IMFO and HH-AUSVM is performed using gene expression datasets from Mendeley Data Repository for breast cancer.

2. RELATED WORKS

Recent studies have presented various feature selection and classification methods for breast cancer analysis from gene expression data. Feature or Gene selection methods can be filter, wrapper or embedded methods. Statistical measures such as correlation coefficient and mutual information are used in filter methods to select genes based on relevancy. Vanitha et al. [11] computed MI between genes and class labels to select the best genes and applied SVM for classification. Recently, Rahmanian and Mansoori [12] developed an unsupervised gene selection method using multivariate normalized MI (MNMI) with higher classification accuracy. Wrapper methods have mostly utilized one or more metaheuristic algorithms. Several algorithms, such as GA [13], PSO [14], GSA [15], etc., provided a training-based selection of genes. Embedded methods are a hybrid of the wrapper and filter methods and select genes based on relevancy and training accuracy. In [16], the authors combined MI and GA to form a hybrid selection method and achieved 90% accuracy. Sun et al. [17] presented hybrid gene selection using ReliefF and ant colony optimization (RFACO) and achieved an average accuracy of 94.3%. Although the embedded methods provide a better performance, the complexity of these methods must be reduced.

Zhang et al. [18] proposed an efficient feature selection strategy using an SVM based on recursive feature elimination and parameter optimization (SVM-RFE-PO). Evaluated on GEO and TCGA datasets, this model achieved an AUC of 96% but also increased the complexity of training. Kong and Yu [19] presented Forest Deep Neural Network (fDNN) model using RF and DNN to extract features to increase the classification accuracy of gene expression data. This model used GEO repository datasets, namely GSE99095 and GSE106291, for evaluation and achieved testing AUC of 0.986 and 0.778 for the two datasets. But this model has limited performance when there are overlapping classes in the dataset. Zhang et al. [20] developed an ensemble classifier based on the principal component analysis (PCA), deep Auto-Encoders and AdaBoost algorithm (PCA-AE-Ada). This model obtained 0.714 AUC and increased the

accuracy from 75% to 85%. Yet, this model is prone to over-fitting owing to the sparse gene datasets.

Elbashir et al. [21] introduced Lightweight CNN with selected hyper-parameters for breast cancer classification with Array-Array Intensity Correlation (AAIC) outlier removal, filtering and normalization. This model obtained 98.76% accuracy on the TCGA dataset for breast cancer gene expression data. However, this method has higher computational complexity. Mondal et al. [22] developed an entropy-based supervised learning method of SVM, RF, k-nearest neighbor (KNN) and naive Bayes for cancerous breast genes. Among them, SVM achieved 91.5% classification accuracy on the GSE349 & GSE350 datasets from the GEO repository. Yet, this model suffers from a class imbalance problem. AbdElNabi et al. [23] developed a cancer classification approach using information gain (IG)-grey wolf optimization (GWO) feature selection and SVM classifier to overcome the over-fitting problems. Evaluated on skewed cancer datasets from Kent Ridge Bio-Medical Data websites, this approach obtained 94.87% accuracy for breast cancer and 95.935% for colon cancer. However, this approach has generated high false positives due to reduced feature selection.

Pham et al. [24] established a model for subtyping breast cancer from gene expression data using an SVM-RFE classifier with GS (SVM-RFE-GS). This model achieved an accuracy of 89.40% with an improvement of 5.44% on TCGA datasets but did not consider the class imbalance problem. Gupta and Gupta [25] presented an improved SVM-RFE gene selection scheme with the Least Absolute Shrinkage Selector Operator (LASSO) and Ridge regression for classifying breast cancer genes. This method reduced the RMSE values from 0.15 to 0.24. Yet, this method has limited learning capability. Hosseinpour et al. [26] developed a Hybrid High-order Type-2 Fuzzy Cognitive Map Improved RF classification (HHTFCMIRF) approach. This approach utilized the improved RF for classifying the breast cancer gene data and achieved 93.5% on the TCGA dataset. But this method also increased the false positives in the presence of the class imbalance problem. Wei et al. [27] proposed generative adversarial networks (GAN) model with data augmentation to detect breast cancers with an accuracy of 92.6%. Still, the class imbalance problem or overlapping class labels is not considered.

Few inferences are obtained from the studies in the literature. The main inference is that the ML algorithms can be more suited for the sparsely sampled gene expression data irrespective of the emergence of complicated DL methods. Though DL methods have better feature learning, ML methods can avoid over-fitting and class imbalance problems more effectively. The other main inference is that feature learning methods can improve classification accuracy with reduced complexity in high-dimensional data. Considering these inferences, the proposed approach has developed the MIM-IMFO feature selection and HH-AUSVM classifier.

3. METHODS

The overview of the proposed approach for breast cancer classification is shown in Fig. 1.

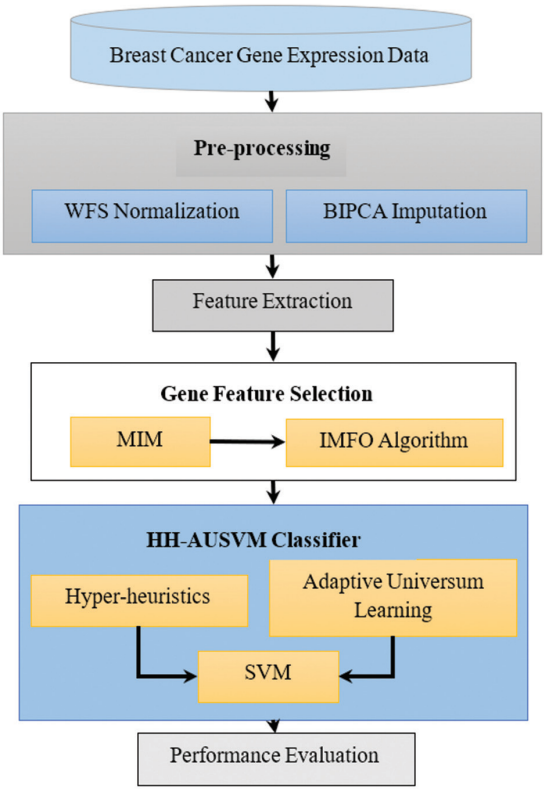


Fig. 1. Overview of the Proposed Breast Cancer Classification model

The proposed approach for breast cancer classification using gene expression analysis includes three main steps: pre-processing, feature extraction and selection, and classification. The pre-processing step intends to offer high-quality gene expression data for investigating breast cancer physiognomies. The pre-processing method has two major processes, namely Weighted Fuzzy Score (WFS) based data normalization and Bayesian Independent Principal Component Analysis (BIPCA) based missing value imputation [28]. After pre-processing, the gene feature vectors are extracted from the datasets, and then the hybrid feature selection method is utilized. First, the mutual information values are computed for the gene vector pairs, and the best gene pairs are selected based on the maximum mutual information values. Then the gene pairs are rearranged using the hybrid phase, and the best genes are ranked in descending order with respect to the fitness values using the IMFO algorithm. Finally, the selected pivotal gene features are fed to the HH-AUSVM to train the classifier for obtaining accurate breast cancer classification in the testing stage.

3.1. DATASETS

The breast cancer gene expression profiles are collected from the Mendeley Data repository (https://data.mendeley.com/datasets/v3cc2p38hb/1).

The main dataset contains four separate datasets: BC-TCGA, GSE2034, GSE25066 and Simulation Data. Table 1 shows the data distribution in these datasets.

Table 1. Distribution of Breast Cancer gene expression datasets.

Datasets	Number of genes	Number of samples		
		Total	Normal Class	Cancer Class
BC-TCGA	17,814	590	61	529
GSE2034	12,634	286	179	107
GSE25066	12,634	492	100	392
Simulation Data	10,000	200	100	100

Table 2 shows examples of normal and tumor classes based on the selected five gene values from the input datasets.

Table 2. Example of Normal and Tumor data.

Tissue	Gene Expressions (Gene Names)					Class
	ELMO2	PNMA1	MMP2	ZHX3	CHD8	
1	0.2043	0.5385	0.7076	-0.117	-0.160	Normal
2	0.0645	0.2335	0.99	-0.468	-0.146	Normal
3	0.2887	0.2327	1.3211	-0.376	-0.162	Normal
4	1.2194	-0.2187	-0.148	0.2108	0.8762	Tumor
5	1.2426	-0.026	-1.073	0.3796	0.3047	Tumor
6	1.1717	-0921	-0.435	0.6231	0.2454	Tumor

3.2. PRE-PROCESSING

The gene expression datasets for breast cancer are raw data with limitations in terms of missing values and outliers. The high-quality data will ensure efficient disease classification. Therefore, data normalization and missing value imputation techniques are applied to the input datasets. WFS-based normalization was developed by integrating the Minkowski Weighted Score Functions into the gene fuzzy score computation. This WFS method used these weighted fuzzy scores to transform the gene expression data values without large variations. Similarly, the missing value problem is solved using the BIPCA imputation method, which utilizes Bayesian theory applied to a fusion model of principal component analysis (PCA) and independent component analysis (ICA) to replace the missing values through likelihood values of informative genes.

3.3. GENE FEATURE SELECTION USING MIM-IMFO

Gene feature selection is choosing the most informative features while eliminating the less informative and irrelevant features. Feature selection can be performed by wrapper, filter and hybrid methods. This study uses a hybrid method by combining the MIM method and the IMFO algorithm. The gene features are selected quickly and efficiently by collaboratively using these two methods. In this hybrid model, the Mutual Information (MI) is used inside the IMFO algorithm and as the metric to estimate the importance of the features. Then the IMFO-based optimization strategy ranks the features based on the fitness values and returns the top-ranked features for the classification. In this model, the IMFO is developed to overcome the limitation of MFO, i.e., the degeneration of the global search capability and slow convergence. To improve the MFO, a hybrid phase is added between exploration and exploitation, which can improve the search process and convergence speed.

Initially, the moths' population is assembled, and the initial parameters are defined along with the maximum number of iterations. In standard MFO, the moths are updated based on the flames, and flames are generated by sorting the best moths. Yet, this process will lead to poor population diversity leading to slow convergence and reduced global search capability. The moth population is initialized, and the elite individual information of moths is protected to eliminate the loss rate. The best solution for the entire population will be stored in a matrix form

$$H = \begin{bmatrix} h_{1,1} & h_{1,2} & \dots & h_{1,d} \\ h_{2,1} & \dots & \dots & h_{2,d} \\ \dots & h_{i,j} & \dots & \dots \\ h_{n,1} & h_{n,2} & \dots & h_{n,d} \end{bmatrix} \quad (1)$$

Here $h_{i,j}$ refers to the best position of the *i-th* moth at the *j-th* dimension. This equation is an enhanced form of the standard MFO initialization step. When the iteration k=1, this equation will equal the initialization function of MFO. This ensures all the moth individuals have a set initial position. These initial positions are updated by the logarithmic spiral as

$$M_i^k = D_i^{k-1} e^{bt} \cos(2\pi t) + F_i^{k-1} \quad (2)$$

Here, M_i^k denotes the position of *i-th* moth at the iteration k, $D_i^{k-1} = |F_i^{k-1} - M_i^{k-1}|$ denotes the distance between *i-th* moth and *i-th* flame at *k-1* iteration and F_i^{k-1} denotes the position of the *i-th* flame at iteration k-1. The parameter b defines the spiral shape, and t denotes a random number between [r,1], with r being a linearly decreasing function from −1 to −2.

Then the MI is computed for these feature subsets, which are mapped with the gene features using Eq. (5), and the process will be terminated if the maximum MI value is obtained. MI is used to compute the effectiveness of the features from high dimensional data to obtain higher classification accuracy. It is estimated as the amount of information via the reduction in entropy. Entropy can measure the diversity in the attributes and helps in obtaining the impurity of information to quantify the uncertainty of the prediction results using the given variable. Hence the entropy is first formulated to compute the MI. Let y denote the discrete random variable attribute with two possible outcomes, i.e., relevant (R) and irrelevant ($\bar{R}$) to the ideal features. The binary function H can be expressed as a logarithmic value.

$$H(y) = -p(R)\log p(R) - p(\bar{R})\log p(\bar{R})$$

Here (R, $\bar{R}$) denotes the possible classes- relevant and irrelevant, $p(R)$ denote the probability of the sample being $y\in(R)$ and $p(\bar{R})$ denote the probability of the sample being $y\in(\bar{R})$. Conditional entropy defines the quantity of the uncertainties of each feature in the decision process, and it is computed between two events, X and Y, where X has the value of feature x,

$$\begin{gathered} H(Y \mid X)=\sum_{x\in X} p_x(x)\, H(Y \mid X=x)= \\ \sum_{x\in X} p_x(x) \sum_{y\in Y} p(y \mid x)\ log\, p_y(y \mid x)= \\ \sum_{x\in X}\sum_{y\in Y} p_{xy}(x,y)\, logp_y(y|x) \end{gathered} \quad (4)$$

The smaller values of the impurity will result in more skewed class distributions. The values of entropy and the misclassification errors will be the highest when the class distribution is uniform and the minimum when all the samples belong to the same class.

The MI of y can be computed using the entropy and conditional entropy from a feature x as

$$MI(y|x) = H(y) - H(y|x) \quad (5)$$

A larger MI defines the higher discriminative power for the decision process and determines the relevance of the features with respect to the classification problem. The gene pairs are rearranged in the hybrid phase to obtain improved gene features.

IMFO includes a new step compared to the MFO. As stated before, the IMFO performs three phases. The primary phase pledges an optimal exploration, the intermediary phase is a hybrid exploration/exploitation, and the final phase improves the exploitation. The iterations determine it, and hence the iterations are divided into $I1$, $I2$ and $I3$ for each phase. It is defined as

$$I_1 = [1, \delta_1] \cap N, I_2 = [\delta_1, \delta_2] \cap N, I_3 = [\delta_2, K] \cap N \quad (6)$$

Here $\delta1=\alpha K$, $\delta2=\beta K$ with $\alpha,\beta\in[0,1]$ and N denotes the set of numbers.

The hybrid phase has been introduced to avoid an abrupt transition between the exploration and exploitation phases. A weighted factor is added to the fitness function (accuracy or error rate) to improve the exploitation phase without downgrading the exploration capability. Weight factor w is given as

$$w = \left|\frac{f(M_{best})}{f(M_i^k)}\right| \quad (7)$$

Here, $f(M_{best})$ denotes the fitness value of the best solution, and $f(M_i^k)$ denotes the fitness values of the i-th moth at iteration k. The exploitation is improved by using this factor during the hybrid phase without influencing the exploration phase. The moth positions are updated as

$$M_i^k = D_i^{k-1} e^{bt}\cos(2\pi t) + w.F_i^{k-1} + (1-w).M_{best} \quad (8)$$

Thus, a good balance is obtained in the trade-off between exploration and exploitation. The fitness value is computed for the new features obtained after the hybrid phase using these steps. Then these features are ranked based on their relevance with respect to the fitness value (accuracy or error rate) computed in IMFO. Thus, the MIM-IMFO gene feature selection helps deep explore and decide the best gene subsets.

3.4. Classification using HH-AUSVM

The benefits of using SVM-based classifiers are that they have universal optimization and a great simplification facility to make the grouping precise. Moreover, it resolves over-fitting problems and reduces computational complications. However, the standard SVM cannot handle the noises and unknown class samples effectively. Hence Adaptive Universum learning is utilized with the SVM so that the classifier learns the patterns of unknown classes and the known classes effectively with prior knowledge. AUSVM constructs the data-dependent architecture of SVM based on the set of tolerable functions for ensuring adaptability. It is more appropriate to obtain the set of Universum samples for the SVM learning process instead of defining the data distributions explicitly. This Universum Learning solves the class imbalance problem better than the other sampling-based methods with good regularization and generalization for the data. The Universum samples are the additional samples generated based on current data but do not belong to the current classes. These data are added to the imbalance classes as false data to balance the data distribution for easy computation without impacting the final output. Since the Universum samples do not belong to any predefined classes, the AUSVM hyper-plane will fall inside the margin borders determined by C due to the usage of the maximal margin procedure. Therefore, the AUSVM must utilize a maximal soft-margin procedure and maximize the number of Universum samples distributed around the hyper-plane.

A training set with given Universum samples is defined as

$$S = \{(x_1, y_1), (x_2, y_2), \ldots . (x_s, y_s)\} \cup \{x_1^*, x_2^*, \ldots, x_u^*\} \quad (9)$$

Here $x_j^* \in R^n$,j=1,2,…,u denote the Universum samples in R^n search space, x_i R^n,i=1,2,…,s and $y_i\in\{1,-1\}$ for binary classification and $y_i R^n$, i=1,2,…,s for multi-class classification. As the Universum samples provide the prior knowledge of the network traffic classification by approximating the hyper-plane $g(x)$=0, the primal optimization algorithm of the AUSVM with the maximal soft-margin procedure is given as

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C_t\sum_{i=1}^{s}\xi_i + C_u\sum_{t=1}^{u}(\psi_t + \psi_t^*) \quad (10)$$

Subject to $y_i\,(w.\Phi\,(x_i) + b) \geq 1\text{-}\xi_i$, - ε - $\psi_t^* \leq w.\Phi\,(x_t^*) + b \leq \varepsilon + \psi_t$, $\xi_i \geq 0$, i=1,2,...,s and ψ_t, $\psi_t^* \geq 0$, t=1,2,...,u.

Here C_t denotes the margin parameter or penalty parameter of SVM, C_u denotes the margin parameter of AUSVM, ψ_t, ψ_t^* denotes the slack variables of AUSVM and ε represents the in-sensitive loss function for Universum samples. Eq. (13) of AUSVM maximizes the margin between the classifying hyper-planes and the amount of Universum samples to be distributed

around the hyper-plane. If C_u=0, then Eq. (13) will become equivalent to the standard SVM equation. This dual problem of USVM is formulated as

$$\begin{gathered}\min_{\alpha,\mu,v}\frac{1}{2}\sum_{i=1}^{s}\sum_{j=1}^{s}y_iy_j\alpha_i\alpha_j\ K(x_i,x_j)+\\ \frac{1}{2}\sum_{t=1}^{u}\sum_{z=1}^{u}(\mu_t-v_t)(\mu_z-v_z)K(x_t^*,x_z^*)+\\ \sum_{i=1}^{s}\sum_{t=1}^{u}y_i\alpha_i(\mu_t-v_t)K(x_i,x_t^*)-\sum_{i=1}^{s}\alpha_i+\\ \varepsilon\sum_{t=1}^{u}(\mu_t+v_t)\end{gathered} \tag{11}$$

Subject to $\sum_{i=1}^{s} y_i \alpha_i \sum_{t=1}^{u} (\mu_z - v_z) = 0;\ 0 \le \alpha_i \le C_t$, $i=1,2,...,s$; $0 \le \mu_t, v_t \le C_u t=1,2,...,u$.

Here μ_i and v_i are Lagrangian multipliers similar to α_i.

When the classifier is non-linear, the problem in the input space from Eq. (13) is defined as

$$f(x) = sign\left(\sum_{i=1}^{m}\alpha_i^* \times y_i \times K(x_i, y_i) + b^*\right) \tag{12}$$

Here $f(x)$ symbolizes the objective function, α_i^* denotes control parameter, and b^* indicates the bias. $K(x_i, y_i)$ represents the kernel function that creates the central product of the feature space. The parameters namely C_u, $K(x_i, y_i)$ and its parameters are selected optimally using HH. The aim is to choose an SVM structure that diminishes the miscalculation error and increases the accuracy without manipulating the complication. It is exhibited as a non-convex optimization issue conveyed as a tuple form

$$< AUSVM, \Theta, \mathcal{D}, \mathcal{C}, S > \tag{13}$$

Where AUSVM is the constructed system, Θ is the exploration area of the conceivable SVM structures, D is the dissemination of the set of cases, C is the fitness utility, and S is the arithmetic data. The objective is to reduce C to achieve the resolution set over a set of issue circumstances to discover

$$\theta^* \in \arg\min_{\theta\in\Theta}\frac{1}{|\mathcal{D}|}.\sum_{\pi\in\mathcal{D}}\mathcal{C}(\theta,\pi) \tag{14}$$

Each $\theta \in \Theta$ characterizes one conceivable structure of the SVM, and the result C is attained while analyzing the SVM through numerous illustrations. The multi-constraint optimization is expressed as

$$\min F(X) = |f_1(x), f_2(x)| \tag{15}$$

Where $f_1(x)$=Accuracy; $f_2(x)$ = Model Complexity

Here $f_1(x)$, $f_2(x)$ are the two objective functions of SVM, and Model complexity is expressed as the Number of Support Vectors (NSV). The HH algorithm reduces this function to acquire a lightweight SVM with high accuracy and less training time.

Hyper-heuristic optimization consists of high-level and low-level heuristics in the SVM design structure optimization. HH is the multi-level algorithm that accomplishes heuristic interpretations and generates low-level policies based on the problem obligation from prevalent solutions. The low-level heuristics follow the solution space and regulate the current solutions to create new solutions for assessment.

The high-level policy chooses the appropriate low-level heuristics as an alternative to probing the solutions so that the low-level policy can execute its functions without any disruptions from the advanced search policy.

4. RESULTS AND DISCUSSION

The proposed MIM-IMFO feature selection and HH-AUSVM classification approach for the breast cancer classification problem is evaluated over the Mendeley gene expression datasets. The evaluations are conducted using the MATLAB tool (R2016b version 9.1) on an Intel Core i7 processor, Windows 10 OS with 8GB RAM and 512GB SSD. The performance metrics, namely Accuracy, Precision, Recall, F-Measure, Pearson Correlation Coefficient (PCC) and Processing Time, are used for the evaluation. These parameters are chosen for evaluation since they can help determine the correctness of the models and also detect the linear relationship of the variables.

$$Accuracy = \frac{(True\ Positive + True\ Negative)}{(Total\ Classes)} \tag{16}$$

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \tag{17}$$

$$Recall = \frac{True\ Positive}{(True\ Positive + False\ Negative)} \tag{18}$$

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{19}$$

$$PCC = \frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum(x_i-\bar{x})^2\sum(y_i-\bar{y})^2}} \tag{20}$$

Here, x_i and y_i denote the x and y variables of the sample, $\bar{x}$ and $\bar{y}$ denote the mean of the values of the x and y variables.

The performance of the proposed gene selection method is evaluated with other existing gene selection methods based on the number of selected genes. All simulations are performed with the same amount of data for fair comparisons. Table 3 shows the performance of gene selection methods for the four datasets.

Table 3. Comparison of Gene Selection Methods

Method	BC-TCGA	GSE2034	GSE25066	Simulation Data
Total Genes	17814	12634	12634	10000
MI [11]	12540	9875	9833	8101
MNMI [12]	11345	8603	8628	6957
GA [13]	9790	7884	7752	5880
PSO [14]	9176	7542	7267	5541
GSA [15]	9042	7125	6823	5179
MI-GA [16]	7920	6043	6043	4222
RFACO [17]	7775	5779	5450	3980
Proposed MIM-IMFO	7108	5369	5122	3338

Table 4 shows the obtained results for the proposed methods over the testing sets of the datasets. The existing method SVM is used as the base classifier along with Adaptive Universum SVM. The proposed HH-AUSVM classifier is used without the feature selection method and with the MIM-IMFO method.

Table 4. Performance comparison of Breast Cancer Gene Expression Classification.

Accuracy				
Method	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	92.36	91.86	90.02	93.89
AUSVM	93.17	92.99	92.58	94.31
HH-AUSVM	94.49	94.26	95.67	94.94
MIM-IMFO + HH-AUSVM	95.67	96.52	97.97	95.5
Precision				
Methods	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	91.66	96.15	89.32	95.36
AUSVM	93.27	97.49	92.33	96.67
HH-AUSVM	95.88	98.67	95.69	98.5
MIM-IMFO + HH-AUSVM	98.75	99.24	96.38	100
Recall				
Methods	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	93.94	92.22	86.25	92.48
AUSVM	94.22	94.67	90.49	94.31
HH-AUSVM	94.89	95.55	92.18	95.96
MIM-IMFO + HH-AUSVM	95.72	96.2	95.04	97.78
F-measure				
Methods	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	92.79	94.14	87.76	93.9
AUSVM	93.74	96.06	91.4	95.48
HH-AUSVM	95.38	97.08	93.9	97.21
MIM-IMFO + HH-AUSVM	97.21	97.7	95.71	98.88
Pearson Correlation Coefficient				
Methods	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	0.9012	0.9147	0.812	0.8813
AUSVM	0.9225	0.9381	0.8892	0.9156
HH-AUSVM	0.9467	0.9448	0.9218	0.9271
MIM-IMFO + HH-AUSVM	0.9588	0.9633	0.9692	0.9524
Processing time (seconds)				
Methods	**BC-TCGA**	**GSE2034**	**GSE25066**	**Simulation Data**
SVM	11.98	10.96	13.56	10.55
AUSVM	10.26	10.18	12.40	9.89
HH-AUSVM	8.78	8.11	10.87	8.25
MIM-IMFO + HH-AUSVM	4.28	3.17	9.45	6.31

The results in Table 1 show that the proposed HH-AUSVM classifier with the MIM-IMFO feature selection has provided better breast cancer classification than the SVM, AUSVM and HH-AUSVM models. Among the compared methods, the proposed HH-AUSVM with MIM-IMFO feature selection has achieved 1.18%, 2.5% and 3.31% higher accuracy, 2.87%, 5.48% and 7.09% higher precision, 0.83%, 1.5% and 1.78% higher recall, 1.83%, 3.47% and 4.42% higher f-measure, 1.21%, 3.63% and 5.76% higher Pearson Correlation Coefficient, and 51.25%, 58.29% and 64.27% reduced processing time than the HH-AUSVM, AUSVM and SVM methods for the BC-TCGA dataset.

For GSE2034 data, it has achieved 2.26%, 3.53% and 4.66% higher accuracy, 0.57%, 1.75% and 3.09% higher precision, and 0.65%, 1.53% and 3.98% higher recall, 0.62%, 1.64% and 3.56% higher f-measure, 1.85%, 2.52% and 4.86% higher Pearson Correlation Coefficient, 60.91%, 68.86% and 71.08% reduced processing time than the HH-AUSVM, AUSVM and SVM methods. For GSE25066 data, HH-AUSVM with the MIM-IMFO has achieved 2.3%, 5.39% and 7.95% higher accuracy, 0.69%, 4.05% and 7.06% higher precision, 2.86%, 4.55% and 8.79% higher recall, 1.81%, 4.31% and 7.95% higher f-measure, 4.74%, 8% and 15.72% higher Pearson Correlation Coefficient and 13.06%, 23.79% and 30.31% reduced processing time than the HH-AUSVM, AUSVM and SVM methods.

For the simulation data, HH-AUSVM with the MIM-IMFO has achieved 0.56%, 1.19% and 1.61% higher accuracy, 1.5%, 3.33% and 4.64% higher precision, 1.82%, 3.47% and 5.3% higher recall, 1.67%, 3.4% and 4.98% higher f-measure, 2.53%, 3.68% and 7.11% higher Pearson Correlation Coefficient and 23.52%, 36.2% and 40.19% reduced processing time than the HH-AUSVM, AUSVM and SVM methods.

The confusion matrix evaluation obtained for the four datasets is shown in Table 5.

Table 5. Confusion Matrix of Proposed Method

Datasets	Total Data	True Positive	True Negative	False Positive	False Negative
BC-TCGA	590	59	506	11	14
GSE2034	286	173	103	3	7
GSE25066	492	98	384	4	6
Simulation Data	200	93	97	4	6

The proposed method's confusion matrix evaluation has shown a good trade-off ratio between the true and false values. In Simulation data, the 200 samples are classified correctly into 93 True Positives (Normal class) and 97 True Negatives (Tumor Class) and wrongly into 4 False Positives and 6 False Negatives. The confusion matrix complements the justification provided by the other evaluation metrics.

This better performance of the MIM-IMFO-based gene selection and HH-AUSVM classifier is because of the use of effective pre-processing by WFS-BIPCA, improved convergence and global search capability of gene feature selection and advanced learning-based optimized classification.

They are compared with the methods used in the literature studies to evaluate the proposed approach further. Since the literature methods have used different breast cancer gene expression datasets in different experimental conditions, comparing their results directly will not be ideal. Hence, for a fair comparison, the methods described in those studies are implemented in the same environment as the proposed approach over the GSE2034 and GSE25066 datasets with an equal amount of data. The comparisons are made in terms of accuracy and processing time. Table 6 shows the comparison of the proposed approach against the literature studies.

Table 6. Performance comparison against methods in the literature.

Ref. No.	Method	GSE2034		GSE25066	
		Accuracy (%)	Time (s)	Accuracy (%)	Time (s)
[18]	SVM-RFE-PO	91.5	7.54	91.87	15.76
[19]	fDNN	93.47	11.66	94.98	24.61
[20]	PCA-AE-Ada	90.88	12.89	91.45	22.55
[21]	CNN	95.25	15.63	95.79	21.8
[22]	Entropy-based SVM	91.31	6.55	92.67	17.17
[23]	IG-GWO + SVM	90.15	5.41	91.11	13.2
[24]	SVM-RFE-GS	88.91	7.88	89.28	15.78
[25]	SVM-RFE-LASSO	86.72	8.91	87.4	16.83
[26]	HHTFCMIRF	91.2	6.37	90.76	14.56
[27]	GAN	93.8	17.88	95.51	23.5
Proposed	**MIM-IMFO + HH-AUSVM**	**96.52**	**3.17**	**97.97**	**9.45**

The comparison of the proposed approach against the existing methods in the literature studies also shows that the proposed MIM-IMFO and HH-AUSVM-based breast cancer classification model performs better than the other methods for the GSE2034 and GSE25066 datasets. There have been accuracy improvements in the proposed approach by approximately 1 to 10%. The processing time of the proposed method is also less than the other methods. This concludes that the proposed approach of MIM-IMFO feature selection and HH-AUSVM classifier has provided a better analysis of the gene expression data for accurate breast cancer classification with less complexity.

5. CONCLUSION

This paper aimed to introduce a hybrid feature selection technique and advanced classifier for reducing the dimensionality of the breast cancer gene expression data to improve the classification performance. Considering the existing limitations, this paper presented an efficient classifier of HH-AUSVM to overcome the class imbalance problem, noisy data, sparse data and parameter tuning problems for analyzing the gene expression data. Utilized with the MIM-IMFO feature selection method, the HH-AUSVM classifier obtained breast cancer classification accuracies of 95.67%, 96.52%, 97.97% and 95.5% for BC-TCGA, GSE2034, GSE25066 and Simulation Data, respectively. It has also consumed about 1-10% less processing time for all four datasets than the existing methods. In the future, the possibility of improving the feature learning property of HH-AUSVM will be investigated. Although the evaluations have been made only on breast cancer gene expression datasets, the proposed method is also suitable for other cancer gene expression datasets. The efficiency of this method for classifying other types of cancer will be examined in the future.

6. REFERENCES

[1] N. Harbeck, F. Penault-Llorca, J. Cortes, M. Gnant, N. Houssami, P. Poortmans, F. Cardoso, "Breast cancer", Nature Reviews - Disease Primers, Vol. 5, No. 1, 2019, pp. 1-31.

[2] L. J. Van't Veer, H. Dai, M. J. Van De Vijver, Y. D. He, A. A. Hart, M. Mao, S. H. Friend, "Gene expression profiling predicts clinical outcome of breast cancer", Nature, Vol. 415, No. 6871, 2002, pp. 530-536.

[3] Z. Sun, Y. W. Asmann, K. R. Kalari, B. Bot, J. E. Eckel-Passow, T. R. Baker, E. A. Thompson, "Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing", Plos One, Vol. 6, No. 2, 2011, pp. 1-16.

[4] N. R. Latha, A. Rajan, R. Nadhan, S. Achyutuni, S. K. Sengodan, S. K. Hemalatha, P. Srinivas, "Gene expression signatures: A tool for analysis of breast cancer prognosis and therapy", Critical Reviews in Oncology/Hematology, Vol. 151, No. 1, 2020, pp. 102964-102978.

[5] M. Abd-Elnaby, M. Alfonse, M. Roushdy, "Classification of breast cancer using microarray gene expression data: A survey", Journal of Biomedical Informatics, Vol. 117, No. 1, 2021, pp. 103764-103782.

[6] A. Bashiri, M. Ghazisaeedi, R. Safdari, L. Shahmoradi, H. Ehtesham, "Improving the prediction of survival in cancer patients by using machine learning techniques: experience of gene expression data: a narrative review",

Iranian Journal of Public Health, Vol. 46, No. 2, 2017, pp. 165-179.

[7] R. Shahane, M. Ismail, C. S. R. Prabhu, "A survey on deep learning techniques for prognosis and diagnosis of cancer from microarray gene expression data", Journal of computational and theoretical Nanoscience, Vol. 16, No. 12, 2019, pp. 5078-5088.

[8] A. Statnikov, L. Wang, C. F. Aliferis, "A comprehensive comparison of random forests and support vector machines for microarray-based cancer classification", BMC Bioinformatics, Vol. 9, No. 1, 2008, pp. 1-10.

[9] S. Huang, N. Cai, P. P. Pacheco, S. Narrandes, Y. Wang, W. Xu, "Applications of support vector machine (SVM) learning in cancer genomics", Cancer Genomics & Proteomics, Vol. 15, No. 1, 2018, pp. 41-51.

[10] S. Mirjalili, "Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm", Knowledge-based Systems, Vol. 89, No. 1, 2015, pp. 228-249.

[11] C. D. A. Vanitha, D. Devaraj, M. Venkatesulu, "Gene expression data classification using support vector machine and mutual information-based gene selection", Procedia Computer Science, Vol. 47, No. 1, 2015, pp. 13-21.

[12] M. Rahmanian, E. G. Mansoori, "An unsupervised gene selection method based on multivariate normalized mutual information of genes", Chemometrics and Intelligent Laboratory Systems, Vol. 222, No. 1, 2022, pp. 104512-104522.

[13] S. Sayed, M. Nassef, A. Badr, I. Farag, "A nested genetic algorithm for feature selection in high-dimensional cancer microarray datasets", Expert Systems with Applications, Vol. 121, No. 1, 2019, pp. 233-243.

[14] Y. Prasad, K. K. Biswas, M. Hanmandlu, "A recursive PSO scheme for gene selection in microarray data", Applied Soft Computing, Vol. 71, No. 1, 2018, pp. 213-225.

[15] P. K. Ram, P. Kuila, "GSA-based approach for gene selection from microarray gene expression data", Machine Learning Algorithms and Applications, Wiley Press, 2021, pp. 159-174.

[16] M. Jansi Rani, D. Devaraj, "Two-stage hybrid gene selection using mutual information and genetic algorithm for cancer data classification", Journal of Medical Systems, Vol. 43, No. 8, 2019, pp. 1-11.

[17] L. Sun, X. Kong, J. Xu, Z. A. Xue, R. Zhai, S. Zhang, "A hybrid gene selection method based on ReliefF and ant colony optimization algorithm for tumor classification", Scientific Reports, Vol. 9, No. 1, 2019, pp. 1-14.

[18] Y. Zhang, Q. Deng, W. Liang, X. Zou, "An efficient feature selection strategy based on multiple support vector machine technology with gene expression data", BioMed Research International, Vol. 2018, No. 1, 2018, pp. 1-11.

[19] Y. Kong, T Yu, "A deep neural network model using random forest to extract feature representation for gene expression data classification", Scientific Reports, Vol. 8, No. 1, 2018, pp. 1-9.

[20] D. Zhang, L. Zou, X. Zhou, F. He, "Integrating feature selection and feature extraction methods with deep learning to predict clinical outcome of breast cancer", IEEE Access, Vol. 6, No. 1, 2018, pp. 28936-28944.

[21] M. K. Elbashir, M. Ezz, M. Mohammed, S. S. Saloum, "Lightweight convolutional neural network for breast cancer classification using RNA-seq gene expression data", IEEE Access, Vol. 7, No. 1, 2019, pp. 185338-185348.

[22] M. Mondal, R. Semwal, U. Raj, I. Aier, P. K. Varadwaj, "An entropy-based classification of breast cancerous genes using microarray data", Neural Computing and Applications, Vol. 32, No. 7, 2020, pp. 2397-2404.

[23] M. L. R. Abd El Nabi, M. Wajeeh Jasim, H. M. EL-Bakry, H. N. Taha, N. E. M. Khalifa, "Breast and colon cancer classification from gene expression profiles using data mining techniques", Symmetry, Vol. 12, No. 3, 2020, pp. 408-422.

[24] T. C. Pham et al. "A New Feature Selection and Classification Approach for Optimizing Breast Cancer Subtyping Based on Gene Expression", Advances in Intelligent Information Hiding and Multimedia Signal Processing, Springer, Singapore, 2021, pp. 298-307.

[25] M. Gupta, B. Gupta, "A novel gene expression test method of minimizing breast cancer risk in reduced cost and time by improving SVM-RFE gene selection method combined with LASSO", Journal of Integrative Bioinformatics, Vol. 18, No. 2, 2021, pp. 139-153.

[26] M. Hosseinpour, S. Ghaemi, S. Khanmohammadi, S. Daneshvar, "A hybrid high-order type-2 FCM improved random forest classification method for breast cancer risk assessment", Applied Mathematics and Computation, Vol. 424, No. 1, 2022, pp. 127038-127052.

[27] K. Wei, T. Li, F. Huang, J. Chen, Z. He, "Cancer classification with data augmentation based on generative adversarial networks", Frontiers of Computer Science, Vol. 16, No. 2, 2022, pp. 1-11.

[28] V. Murugesan, P. Balamurugan, "Weighted Fuzzy Score Normalization and Bayesian Independent Principal Component Analysis Imputation for Breast Cancer Gene Expression Analysis", International Journal of Intelligent Engineering and Systems, Vol. 15, No. 3, 2022, pp. 80-89.

Effective Prostate Cancer Detection using Enhanced Particle Swarm Optimization Algorithm with Random Forest on the Microarray Data

Original Scientific Paper

Sanjeev Prakashrao Kaulgud
Department of Computer Science, Reva University and Presidency University, Bangalore, India.
sanjeev.kaulgud@gmail.com

Siddanagouda Somanagouda Patil
Department of Applied Maths & Computer Science, University of Agricultural Sciences, Bengaluru, India
spatilsuasb@gmail.com

Vishwanath Hulipalled
Department of Computing and Information Technology, REVA University, Bengaluru, India
vishwa.gld@gmail.com

Prabhuraj Metipatil
Department of Computer Science & Engineering, REVA University, Bengaluru, India
pmetipatil@gmail.com

Abstract *– Prostate Cancer (PC) is the leading cause of mortality among males, therefore an effective system is required for identifying the sensitive bio-markers for early recognition. The objective of the research is to find the potential bio-markers for characterizing the dissimilar types of PC. In this article, the PC-related genes are acquired from the Gene Expression Omnibus (GEO) database. Then, gene selection is accomplished using enhanced Particle Swarm Optimization (PSO) to select the active genes, which are related to the PC. In the enhanced PSO algorithm, the interval-newton approach is included to keep the search space adaptive by varying the swarm diversity that helps to perform the local search significantly. The selected active genes are fed to the random forest classifier for the classification of PC (high and low-risk). As seen in the experimental investigation, the proposed model achieved an overall classification accuracy of 96.71%, which is better compared to the traditional models like naïve Bayes, support vector machine and neural network.*

Keywords*: Gene expression Omnibus, Particle Swarm Optimizer, Prostate Cancer, Random Forest*

1. INTRODUCTION

In recent decades, PC is the growing common non-cutaneous cancer type among men. According to the PC cancer agency, the death rate of PC is 2 million in 2018 [1-2]. Currently, the gene-level treatment shows great attention among the clinicians that significantly find the normal and abnormal patterns of the patients [3-4]. In Gene Expression Analysis (GEA), two major concerns in the microarray datasets are high data dimension and low samples [5]. By using the minimum number of samples, the classification of the disease may lead to incorrect decisions [6-7]. To address the above-mentioned difficulties, feature optimization is included in the GEA to find the effective subsets of features/genes [8-9]. The purpose of feature optimization is to identify the minimum number of feature subsets for attaining better classification accuracy. The conventional metaheuristics-based feature optimization algorithms like Bat optimizer, cuckoo search and artificial bee colony, include some disadvantages like it taking extra processing time, and being complex to resolve technical and scientific concerns [10]. To overcome the above-stated issues, an enhanced optimization algorithm is proposed in this manuscript to detect PC at an early stage. In this work, a new supervised system is implemented to improve the performance of PC detection using microarray data. Here, the PC-related

genes from the GEO database (GEO IDs: GSE 15484, GSE 21034, GSE 3325, and GSE 3998) were collected. The main contributions are listed below:

- After the collection of microarray data: GSE 15484, GSE 21034, GSE 3325 and GSE 3998, the gene assortment is performed utilizing an enhanced PSO algorithm on distinct GEO IDs. To enhance the searching ability of traditional PSO, many methods are included in the conventional PSO algorithm. In this article, the Interval-Newton methodology is used for keeping a reasonable search space by adjusting the swarm diversity adaptively and performing local search significantly.
- A random forest classifier is used to classify the sub-classes of PC such as low-risk (non-aggressive PC) and high-risk (aggressive PC) after selecting the optimal genes.
- The random forest classifier is the best choice for microarray data classification because it easily resolves the issue of unstructured or unbalanced data. In addition, the random forest classifier was utilized for solving both regression and classification issues, because it automatically handles the missing values in microarray data.
- The proposed model performance is related with the prior works in terms of error rate, accuracy, False Positive Rate (FPR), specificity, sensitivity and False Negative Rate (FNR).

The research paper is prepared as follows: A few research papers in PC detection using microarray data are surveyed in Section 2. The description of the proposed method is stated in Section 3. Section 4 represents the comparative and quantitative study of the work. The conclusion of this research paper is given in Section 5.

2. LITERATURE SURVEY

In recent times, microarray-based data classification gained more attention among researchers, especially for disease classification. Presently, various research works are carried out for prostate disease detection utilizing microarray data. In this section, a brief valuation of a few vital contributions to the prevailing literature is mentioned. Kim [11] developed a new inner class-clustering algorithm for classifying PC into aggressive (high-risk PC) and non-aggressive (low-risk PC). In this literature paper, the developed clustering algorithm investigates the gene pairs with a higher ranked score, which were related to the biological process of PC. In the resulting segment, the presented algorithm attained effective performance related to the existing studies by means of Area under Curve (AUC). Sharbaf [12] developed an optimization-based machine-learning scheme for prostate and Leukemia disease detection using microarray data. The developed system includes a filtering approach for reducing the dimension of the microarray data that completely lessens the system complexity and search space-time. Then, a wrapper technique was developed along with an ant colony optimizer based on cellular learning automata for extracting the feature vectors from the microarray data. Finally, the features were classified by applying three classification methods such as naïve Bayes, k-nearest neighbour, and Support Vector Machine (SVM). The experimental evaluation confirmed that the presented scheme achieved superior results by means of accuracy.

Paul and Sil, [13] presented an effective gene selection algorithm (Non-dominated Sorting Genetic Algorithm (NSGA)) for determining the biologically related genes for PC detection. The developed algorithm includes two parameters (confusion factor and risk factor). Initially, determine the risk factor of every PC-related gene to avoid data misclassification. Then, the confusion factor was calculated for each gene that helps to identify the confusion of a gene in detection, due to sample closeness in the normal and cancer classes. The experimental investigation shows that the presented gene selection algorithm attained superior performance in PC detection using specificity, accuracy, and sensitivity. Elyasigomari [14] developed a new gene selection and classification algorithm in microarray data for disease detection such as prostate, colon, leukemia, and lymphoma. In this research article, a new gene selection algorithm (cuckoo optimizer with a genetic algorithm) was developed to select the most predominant genes utilizing shuffling for better classification. After gene selection, the selected genes were classified by utilizing an artificial neural network, SVM, and Multilayer Perceptron (MP). From the experimental analysis, the SVM classifier attained better performance related to other classifiers in all the databases.

Nguyen [15] implemented a new gene selection procedure (modified analytic hierarchy procedure) for choosing the pre-dominant gene subsets for cancer DNA microarray data classification. The developed gene selection procedure chooses the relevant cancer genes based on the entropy test, Wilcoxon test, two-sample t-test, single-to-noise ratio, and receiver operating characteristics curve. Then, the selected pre-dominant gene subsets were classified using k-nearest neighbour, probabilistic neural network, Linear Discriminant Analysis (LDA), SVM, and MP. Gumaei, [16] integrated random committee ensemble learning and Correlation Feature Selection Algorithm (CFSA) for detecting PC. The experiments were performed on the public benchmark PC dataset by utilizing a tenfold cross-validation approach to analyze the developed method's performance. Alshareef, [17] developed a Chaotic Invasive Weed Optimization (CIWO) method for selecting the optimum subset of features. Further, the Deep Neural Network (DNN) model was used for detecting the existence of PC. In the present manuscript, a novel gene assortment algorithm (enhanced PSO algorithm) is proposed to improve PC recognition. Hence, the limitations and advantages of the existing papers are given in table 1.

Table 1. Limitations and the advantages of the existing papers

Author	Advantage	Limitation
Kim [11]	Developed an inner class-clustering algorithm that identifies the new unknown gene pairs, which effectively helps in differentiating the lower and higher risk cancers	The developed algorithm was only appropriate for binary classification that was considered as one of the major concerns
Sharbaf [12]	Implemented an optimization-based machine-learning scheme for prostate and Leukemia disease detection with limited computational complexity and search space time	The developed scheme was more applicable for structured data, but it showed limited results in unstructured data
Paul and Sil [13]	Developed an effective gene selection algorithm: NSGA that finds the biologically related genes for PC detection	In a few circumstances, the developed algorithm leads to class imbalance concerns
Elyasigomari [14]	Integrated cuckoo optimizer with a genetic algorithm for early disease detection such as prostate, colon, leukemia, and lymphoma	The developed gene selection algorithm needs more iterations for achieving better results which was considered as one of the key concerns in this research study
Nguyen [15]	The extensive experimental investigation shows that the LDA classifier attained good performance in all the datasets (prostate, colon, lymphoma, and leukemia cancer datasets) related to other classifiers	The undertaken classifier: LDA was mainly applicable for binary classification not for multiclass classification
Gumaei [16]	Integrated random committee ensemble learning and CFSA for early detection of PC	However, the developed method was computationally expensive

3. METHODOLOGY

In developing countries, PC is common cancer among men, which almost affects (7%) of the total population [18-19]. The major diagnostic tools applied for PC recognition are computed tomography, ultrasonic sound, magnetic resonance imaging, etc. Among these, micro-array data showed more attention among the researchers, because it represents the cell state at the molecular level [20-21]. Automatic PC detection includes a few drawbacks like limited samples and high dimensional data. In this article, a new improved gene optimization algorithm and supervised classifier are proposed for addressing the above stated drawbacks. Here, a novel supervised automated scheme is suggested for PC identification. The work proposed here contains three main stages namely data attainment, gene selection, and gene classification. The working method for the work proposed is shown in Fig. 1.

Initially, the PC-related genes are collected from the GEO dataset for automatic PC recognition. The GEO dataset is a publicly available dataset that contains original submitter-supplied records such as series, platform, and sample. Some of the series associated with PC are considered in this research article such as GSE 15484, GSE 21034, GSE 3325 and GSE 3998. A brief discussion about the undertaken GEO IDs is given as follows:

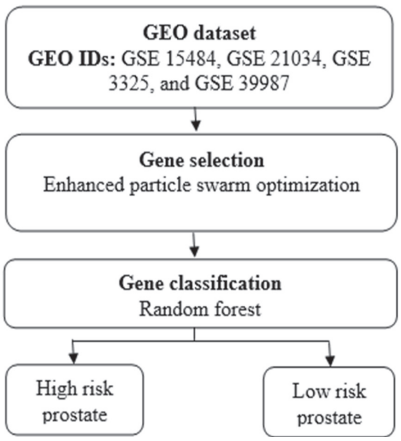


Fig. 1. Working procedure of proposed work

3.1. DATA COLLECTION

- GSE 15484 contains a total of 54675 genes. In this series, the comparison of microarray data from prostate tumors Gleason 8 (length=27), benign prostate tissues (length=13), and prostate tumors Gleason 6 (length=25).
- GSE 21034 comprises 12282 genes for 218 patients with metastatic or primary PC.
- GSE 3325 and GSE 3998 contain a total of 43419 genes and it includes thirteen individuals with metastatic and benign PC samples and six pooled samples from metastatic and benign PC tissues.
- Gene selection using enhanced PSO algorithm

Once the microarray data is collected, then the gene selection is accomplished using an enhanced PSO algorithm. Hence, PSO is a computational algorithm that mimics the behavior of a swarm rather than other evolutionary algorithms. Initially, consider a swarm size n, and then every particle i is indicated as an object with some features. Then, the particles i population are reset with position X_i and velocity V_i. Besides, estimate the objective function F_i utilizing the input value of particles position coordinates. Each particle i tracks the position coordinates, which are related to the best fitness solution achieved so far which is named Pbest. Another best value is calculated globally by the swarms, which is considered as the overall best value Gbest [22]. The particle i velocity is indicated as X_i=[X_{i1}, X_{i2},….X_{id}], and the particle i location is signified as V_i=[V_{i1}, V_{i2},….V_{id}]. In addition, every particle i has historically best position, which is specified as h_i=[h_{i1}, h_{i2},….h_{id}]. The particle's best position is identified by using the position of neighborhood particles that is specified as n_i=[n_{i1},….X_{id}]. The vectors X_0 and V_i are updated randomly by utilizing the equations (1) and (2).

$$V_{id} = wV_{id} + B_1 r_{1d}(h_{id} - X_{id}) + B_2 r_{2d}(n_{id} - X_{id}) \quad (1)$$

$$X_{id} = X_{id} + V_{id} \quad (2)$$

Where w is specified as inertia weight, B_1 and B_2 are stated as acceleration coefficients, and r_{1d} and r_{2d} are indicated as two randomly generated values within the range of [0,1] in the d dimensional space. In this article, an interval-newton method is included in the conventional PSO algorithm to keep the reasonable search space by adjusting the swarm diversity adaptively and also to perform local search effectively. The interval-newton approach is one of the effective approaches to solving a system with non-linear equations with d values. Consider F as a continuous function that has continuous partial derivatives as denoted in Eq. (3).

$$F(X^1, X^2, \dots X^d) = \begin{bmatrix} f_1(X^1, X^2, \dots X^d) \\ : \\ f_d(X^1, X^2, \dots X^d) \end{bmatrix} = 0 \qquad (3)$$

To solve the equation $F(X) = 0$, select the starting point $X_0=[X_0^1, X_0^2,\dots X_0^d]$. Then, apply an iterative formula in Eq. (3) as represented in Eq. (4).

$$X_{n+1} = X_n - J^{-1}(X_n)F(X_n),\ n = 0,1,2 \dots \qquad (4)$$

Where J is represented as the Jacobian matrix of F, and then introduce the interval–newton method $N(X)=J^{-1}(X_n)F(X_n)$ in Eq. (4), which is defined in Eq. (5) and (6).

$$X_{n+1} = X_n - N(X),\ n = 0,1,2 \dots \qquad (5)$$

$$X_{n+1} = X_n + V_{n+1} \qquad (6)$$

Where V_0 is indicated as starting velocity, X_0 is specified as starting position, V_{n+1} is represented as the present velocity of the particle i, and X_n is represented as the prior portion of the particle i. The interval-newton methodology sums the position of the particle X_n with present velocity V_{n+1}. The present velocity of particle i is identified by utilizing the acceleration constant and inertia weight, which is mathematically stated in Eq. (7).

$$V_{n+1} = wV_n + BN(X_n) \qquad (7)$$

Where, w is denoted as inertia weight, and B is indicated as acceleration constant. The gene selection after applying the enhanced PSO algorithm is denoted in Table 2.

Table 2. Gene selection after applying enhanced PSO algorithm

GEO IDs	Number of genes	Selected genes
GSE 15484	54675	300
GSE 21034	12282	300
GSE 3325 and GSE 3998	43419	300

3.3. GENE CLASSIFICATION

After gene selection, the classification of potential biomarkers is accomplished by utilizing Random Forest (RF), which solves both regression and classification issues by automatically adjusting the missing values in microarray data. RF completely lessens the issue of probability density complexity. The number of trees in an RF is stated as an individual classifier and the outcome of the classifier is chosen by all the decision trees. In this article, the tree length in the RF classification approach is fifty. The growth rule of each tree is identified to develop the RF classification approach. Then, the bootstrap samples are chosen from B for each tree in the forest, where $B(i)$ is signified as i^{th} bootstrap.

Randomly sample the micro-array data from training set S, where D is denoted as a dimension of input micro-array data, and S is stated as the number of training sets. If $d(d<D)$, select the sub-data d from the original micro-array data. From the D data, the values of d data are randomly selected and the nodes of the classifier are split by utilizing the best split on the d- dimensional data. In RF, the trees are developed until the training samples are divided without pruning. The error rate in the RF mainly depends on two dissimilar aspects:

- Power of every individual tree in the forest: The forest error rate is diminished by increasing the strength of the tree.

The connection among the trees in the forest: A higher amount of correlation leads to a higher error rate. Correspondingly, the smaller correlation leads to a smaller error rate. The pseudo-code of the proposed model is stated below:

Pseudo-code of the proposed model

Input: GSE 15484, GSE 21034, GSE 3325, and GSE 3998

Output: Classification of the potential biomarkers to predict high and low-risk prostate cancer.

Precondition: Number of relevant genes selected by enhanced PSO algorithm, number of trees in the forest F, and the training set (x_1,y_1), (x_2,y_2)..., (x_n,y_n), features (genes) D

```
Function RF (B, D)
  E ← ∅, ensemble
  For i∈1,….,F do
    B(i) ← A bootstrap section of B
    ei ← Randomized tree learn (B(i),D)
    E ← EU{ei}
  End for
  Return E
End function
Function Randomized tree learn (B, D)
  For every node:
    d ← minor subset of D
    Fragmented on top feature in d
  Return the obtained tree
End function
```

4. EXPERIMENTAL INVESTIGATION

The proposed analysis is simulated by MATLAB tool (version 2018a) in this research paper. To verify the efficacy of the proposed work, the performance of the proposed work is related with some previous works on

the GEO database. The proposed work output is validated in terms of error rate, FPR, specificity, sensitivity, accuracy and FNR. The system configuration and the parameter settings are given in table 3.

Table 3. System configuration and the parameter settings

System configuration	
Random access memory	8 GB
Hard-disk	1TB
Operating system	Windows10
Processor	Intel core i5
Enhanced PSO algorithm	
Acceleration constant	02
Initial inertia weight	0.9
Final inertia weight	0.4
Population size	300
Random forest	
Number of trees	50
Maximal depth	300

4.1. PERFORMANCE METRIC

The performance metric is described as the technique of evaluating the facts of a group or separate variable's performance by means of error rate, FPR, specificity, sensitivity, accuracy and FNR. The mathematical formula for calculating the error rate, FPR, specificity, sensitivity, accuracy and FNR are represented in Eq. (8), (9), (10), (11), (12), and (13).

$$Error\ rate = \frac{FP+FN}{P+TP+FN+TN} \times 100 \quad (8)$$

$$FPR = \frac{FP}{TN+FP} \times 100 \quad (9)$$

$$Specificity = \frac{TN}{FP+TN} \times 100 \quad (10)$$

$$Sensitivity = \frac{TP}{FN+TP} \times 100 \quad (11)$$

$$Accurcay = \frac{TP+TN}{FP+TP+FN+TN} \times 1000 \quad (12)$$

$$FNR = \frac{FN}{TP+FN} \times 100 \quad (13)$$

Where, TN, FP, TP and FN correspondingly indicate True Negative, False Positive, True Positive, and False Negative.

4.2. QUANTITATIVE INVESTIGATION

In this segment, the GEO database is applied for measuring the effectiveness of the given work. In this research, PC recognition is performed on a data mining platform for categorizing the sub-classes of PC like high and low-risk cancers. Here, the performance valuation is done with data split as 70% training and 30% testing. The performance of the suggested model is given in Table 4 and it is evaluated by considering the performance metrics like accuracy, sensitivity, and specificity for 50 iterations. Additionally, the effectiveness of the suggested model is compared with three present classification methods like SVM, Neural Network (NN), and Naive Bayes (NB). From the simulation result, the average accuracy of RF is 96.71% (combines all four GEO IDs: GSE 15484, GSE 21034, GSE 3325, and GSE 3998) and the comparative classification approaches: NN, SVM, and NB attains 72.78%, 85.05%, and 84.49% of accuracy. The average sensitivity of the RF classifier is 95.3% and the existing classification approaches achieve an average sensitivity of 81.28%, 87.22% and 85.27%. Similarly, 95.58% is the average specificity of RF and the existing classifiers (NN, SVM, and NB) attain an average specificity of 62.43%, 76.92% and 82.07%. Fig. 2 shows the Graphical representation of the suggested work in terms of accuracy, sensitivity, and specificity.

Table 4. Performance evaluation of the suggested work in terms of accuracy, sensitivity, and specificity with different classifiers

GEO IDs	Classifier	Accuracy (%)	Sensitivity (%)	Specificity (%)
GSE 15484	NN	66.67	67.17	68.12
	SVM	72.67	80	53.33
	NB	76.67	80	73.33
	RF	93.33	93.34	93.32
GSE 21034	NN	72.92	79.17	66.67
	SVM	87.50	91.67	83.33
	NB	80.56	83.33	77.78
	RF	98.11	94.89	95.33
GSE 3325 and GSE 3998	NN	78.75	97.5	52.5
	SVM	95	90	94.11
	NB	96.25	92.5	95.12
	RF	98.7	97.82	98.1

Table 5. Performance evaluation of the proposed work by means of error rate, FPR and FNR with dissimilar classifiers

GEO IDs	Classifier	Error rate	FPR	FNR
GSE 15484	NN	33.33	32.19	31.13
	SVM	27.67	20	46.67
	NB	23.21	20	26.67
	RF	13.33	17.27	23.67
GSE 21034	NN	27.08	20.83	33.33
	SVM	12.50	8.33	16.67
	NB	19.44	16.67	22.22
	RF	1.89	11.11	16.67
GSE 3325 and GSE 3998	NN	21.25	19.56	47.5
	SVM	5	12.23	34.4
	NB	13.75	19.5	36.5
	RF	1.3	11.7	26.6

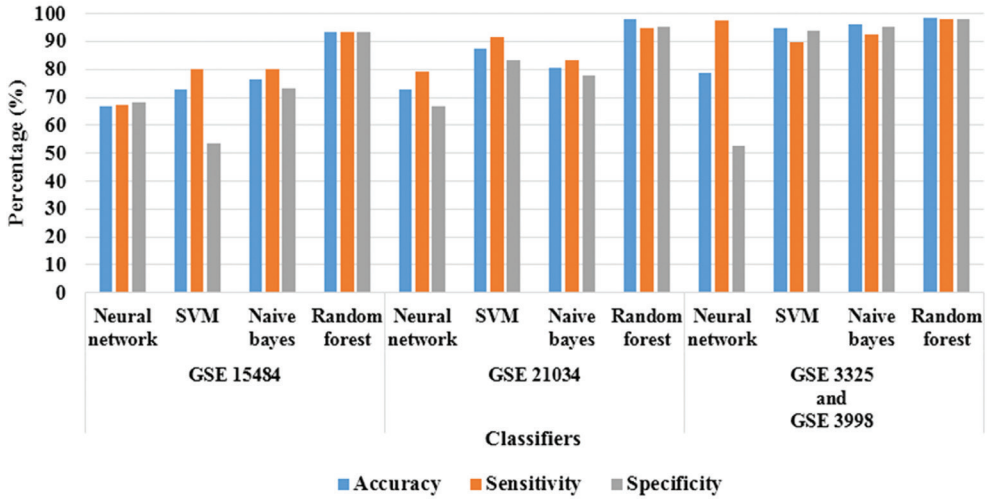


Fig. 2. Graphical representation of the suggested work in terms of accuracy, specificity, and sensitivity

In Table 5, the proposed work effectiveness is evaluated by means of error rate, FPR and FNR for 50 iterations. From the simulation consequences, the average error rate value of the RF classifier is 5.5% and the comparative classification approaches: NN, SVM, and NB classifiers achieve 27.22%, 15.05%, and 18.8% of average error rates. In addition, the FPR average value of the RF is 13.36% and existing classifiers (NN, SVM, and NB) attain 24.19%, 13.52% and 18.72% of average FPR value. At last, the FNR average value of the RF is 22.31% and the existing classifiers achieve 37.32%, 32.58% and 28.46% of the average FNR value. Tables 4 and 5 presented that the suggested work executes effectively on the GEO database (GEO IDs: GSE 15484, GSE 21034, GSE 3325, and GSE 3998) in light of error rate, FPR, specificity, sensitivity, accuracy and FNR. The graphical representation of the suggested work in terms of error rate, FPR, and FNR is shown in Fig. 3.

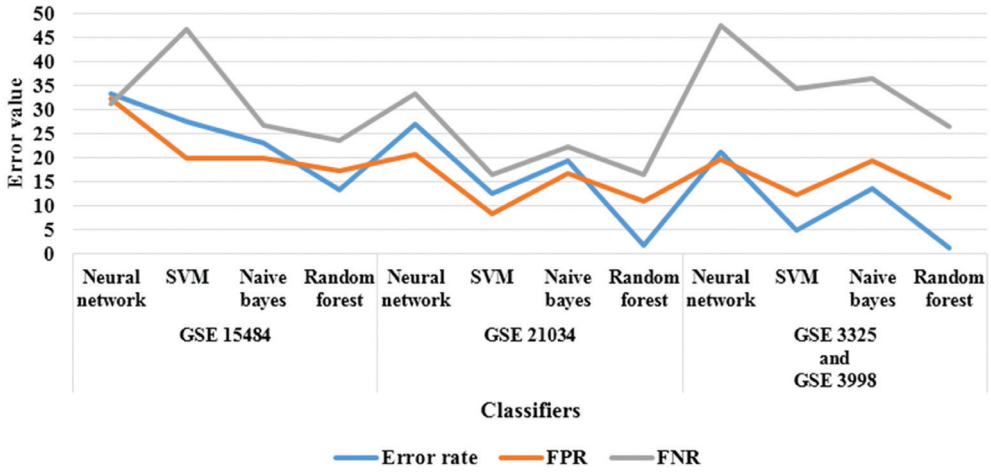


Fig. 3. Graphical representation of the suggested work in terms of error rate, FPR and FNR

Table 6 validates the performance of the suggested work with conventional PSO algorithm and enhanced PSO algorithm in light of accuracy. In the enhanced PSO algorithm, the RF averagely enhances the accuracy of PC recognition by up to 1.5-5% associated to the PSO algorithm. Here, in this work, feature optimization plays a critical role in PC recognition. Generally, the collected GEO data contains several features (genes or potential biomarkers) that might give rise to the "curse of dimensionality" problem. Therefore, reducing the dimensionality is important to optimize the genes or to select the ideal genes (associated with PC), which are appropriate for superior classification. The effectiveness of reducing dimensionality is shown in Table 6 and Fig. 4.

Table 6. Comparison of performance of the proposed work with different optimization algorithms

GEO IDs	Gene selection	Classifier	Accuracy (%)
GSE 15484	PSO	RF	90.17
	Enhanced PSO		93.33
GSE 21034	PSO		93.72
	Enhanced PSO		98.11
GSE 3325 and GSE 3998	PSO		96.29
	Enhanced PSO		98.71

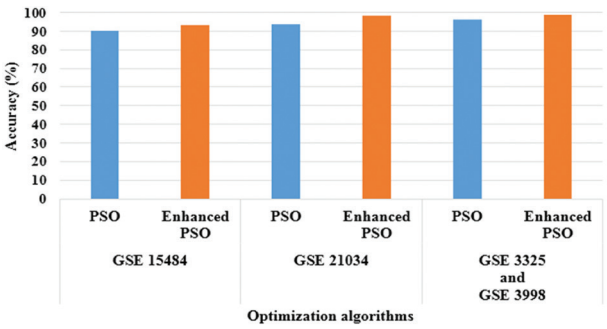


Fig. 4. Graphical illustration of the proposed work with different optimizers

4.3. COMPARATIVE ANALYSIS

In this section, the comparative evaluation of the proposed and the existing works are shown in tables 7 and 8. Kim [11] implemented a novel inner class-clustering algorithm for classifying the sub-classes of PC as aggressive or non-aggressive. In this study, the developed algorithm was tested on GEO dataset (GEO IDs: GSE 15484 and GSE 21034). From the experimental simulation, the developed algorithm attained 0.876 of AUC value on GSE 15484, and 0.989 of AUC value on GSE 21034. Compared to the published works, the proposed model achieved 0.923 of AUC value on GSE 15484 and 0.991 of AUC value on GSE 21034, which were high, related to the inner class-clustering algorithm.

Table 7. Comparative study of proposed and existing works in terms of AUC

Methods	GEO IDs	AUC
Inner class clustering algorithm [11]	GSE 15484	0.876
	GSE 21034	0.989
Enhanced PSO with RF	GSE 15484	0.923
	GSE 21034	0.991

Gumaei, [16] combined random committee ensemble learning model and CFSA for detecting PC. The developed model has averagely obtained 95.10% of recall and 95.09% of accuracy on the benchmark PC dataset. Correspondingly, Alshareef, [17] integrated CIWO and DNN for detecting the existence of PC. Hence, the developed CIWO-DNN model achieved 97.25% of recall and 97.19% of accuracy. However, the proposed enhanced PSO with RF model averagely obtained 95.35% of recall and 96.71% of accuracy in PC detection. As a future extension, a new hybrid feature optimization algorithm can be implemented to further improve PC detection.

Table 8. Comparative study of proposed and existing works by means of accuracy and recall

Methods	Accuracy (%)	Recall (%)
Correlation feature optimization with committee ensemble learning [16]	95.09	95.10
CIWO-DNN [17]	97.19	97.25
Enhanced PSO with RF	96.71	95.35

In the proposed research work, a new optimization algorithm (enhanced PSO) is developed for selecting the optimal PC-related genes. The enhanced PSO algorithm maintains better search space by adjusting swarm diversity that significantly reduces the computational complexity, which is the main concern mentioned in the literature section. Hence, the efficiency of the proposed optimization algorithm is given in tables 7 and 8.

5. CONCLUSION

This research aims to suggest new feature optimization (enhanced PSO) algorithm for choosing the most informative potential biomarkers for PC recognition. The selected potential biomarkers (PC-related genes) are classified by employing an RF classifier. The undertaken classification approach classifier effectively classifies the subclasses of PC: high-risk and low-risk PC. Related to the existing work, the developed work obtained an efficient performance in terms of accuracy. The proposed work achieved 96.71% of the overall classification accuracy in the GEO dataset from this research analysis, which is superior compared to the earlier research works. A new hybrid feature optimization algorithm is combined with a multi-class classifier to further improve the efficiency of PC identification in future work.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

The paper background work, conceptualization, methodology, dataset collection, implementation, result analysis and comparison, preparing and editing draft, as well as visualization have been done by the first and second authors. The supervision, review of work, and project administration, have been done by the third and fourth authors.

6. REFERENCES:

[1] B. Lee et al. "Long noncoding RNAs as putative biomarkers for prostate cancer detection", The Journal of Molecular Diagnostics, Vol. 16, No. 6, 2014, pp. 615-626.

[2] P. Östling et al. "Systematic analysis of microRNAs targeting the androgen receptor in prostate cancer cells", Cancer Research, Vol. 71, No. 5, 2011, pp. 1956-1967.

[3] S. M. G. Espiritu et al. "The evolutionary landscape of localized prostate cancers drives clinical aggression", Cell, Vol. 173, No. 4, 2018, pp. 1003-1013.e15.

[4] A. A. Dmitriev et al. "Identification of novel epigenetic markers of prostate cancer by NotI-microar-

ray analysis", Disease Markers, Vol. 2015, 2015, p. 241301.

[5] E. K. Markert, H. Mizuno, A. Vazquez, A. J. Levine, "Molecular classification of prostate cancer using curated expression signatures", Proceedings of the National Academy of Sciences of the United States of America, Vol. 108, No. 52, 2011, pp. 21276-21281.

[6] T. Mehmood, A. Kanwal, M. M. Butt, "Naive Bayes combined with partial least squares for classification of high dimensional microarray data", Chemometrics and Intelligent Laboratory Systems, Vol. 222, 2022, p. 104492.

[7] S. Begum, R. Sarkar, D. Chakraborty, S. Sen, U. Maulik, "Application of active learning in DNA microarray data for cancerous gene identification", Expert Systems with Applications, Vol. 177, 2021, p. 114914.

[8] S. Sucharita, B. Sahu, T. Swarnkar, "An Empirical Analysis of PCA-SVM Model for Cancer Microarray Data Classification", Advances in Intelligent Computing and Communication, Lecture Notes in Networks and Systems, Vol. 202, pp. 495-504, Springer, Singapore, 2021.

[9] J. Kim et al. "Hydrogel-based hybridization chain reaction (HCR) for detection of urinary exosomal miRNAs as a diagnostic tool of prostate cancer", Biosensors and Bioelectronics, Vol. 192, 2021, p. 113504.

[10] N. S. Mohamed, S. Zainudin, Z. A. Othman, "Metaheuristic approach for an enhanced mRMR filter method for classification using drug response microarray data", Expert Systems with Applications, Vol. 90, 2017, pp. 224-231.

[11] H. Kim, J. Ahn, C. Park, Y. Yoon, S. Park, "ICP: A novel approach to predict prognosis of prostate cancer with inner-class clustering of gene expression data", Computers in Biology and Medicine, Vol. 43, No. 10, 2013, pp. 1363-1373.

[12] F. V. Sharbaf, S. Mosafer, M. H. Moattar, "A hybrid gene selection approach for microarray data classification using cellular learning automata and ant colony optimization", Genomics, Vol. 107, No. 6, 2016, pp. 231-238.

[13] A. Paul, J. Sil, "Identification of differentially expressed genes to establish new Biomarker for cancer prediction", IEEE/ACM Transactions on Computational Biology and Bioinformatics, Vol. 16, No. 6, 2019, pp. 1970-1985.

[14] V. Elyasigomari, M. S. Mirjafari, H. R. Screen, M. H. Shaheed, "Cancer classification using a novel gene selection approach by means of shuffling based on data clustering with optimization", Applied Soft Computing, Vol. 35, 2015, pp. 43-51.

[15] T. Nguyen, A. Khosravi, D. Creighton, S. Nahavandi, "A novel aggregate gene selection method for microarray data classification", Pattern Recognition Letters, Vol. 60-61, 2015, pp. 16-23.

[16] A. Gumaei, R. Sammouda, M. Al-Rakhami, H. Al-Salman, A. El-Zaart, "Feature selection with ensemble learning for prostate cancer diagnosis from microarray gene expression", Health Informatics Journal, Vol. 27, No. 1, 2021, p. 1460458221989402.

[17] A. M. Alshareef et al. "Optimal Deep Learning Enabled Prostate Cancer Detection Using Microarray Gene Expression", Journal of Healthcare Engineering, Vol. 2022, 2022, p. 7364704.

[18] X. Wan et al. "Identification of androgen-responsive lncRNAs as diagnostic and prognostic markers for prostate cancer", Oncotarget, Vol. 7, No. 37, 2016, pp. 60503-60518.

[19] M. Lohr et al. "Identification of sample annotation errors in gene expression datasets", Archives of Toxicology, Vol. 89, No. 12, 2015, pp. 2265-2272.

[20] B. Yang et al. "Downregulation of miR-139-5p promotes prostate cancer progression through regulation of SOX5", Biomedicine & Pharmacotherapy, Vol. 109, 2019, pp. 2128-2135.

[21] D. Wen, J. Geng, W. Li, C. Guo, J. Zheng, "A computational bioinformatics analysis of gene expression identifies candidate agents for prostate cancer", Andrologia, Vol. 46, No. 6, 2014, pp. 625-632.

[22] K. L. Du, M. N. S. Swamy, "Particle Swarm Optimization", Search and Optimization by Metaheuristics, Techniques and Algorithms Inspired by Nature, Springer International Publishing, 2016, pp. 153-173.

Iris Biometric Watermarking for Authentication Using Multiband Discrete Wavelet Transform and Singular-Value Decomposition

Original Scientific Paper

S. Joyce
Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, India.
joyceimmanuel13@gmail.com

S. Veni
Department of Computer Science,
Karpagam Academy of Higher Education, Coimbatore, India
venikarthik04@gmail.com

Abstract *– The most advanced technology, watermarking enables intruders to access the database. Various techniques have been developed for information security. Watermarks and histories are linked to many biometric techniques such as fingerprints, palm positions, gait, iris and speech are recommended. Digital watermarking is the utmost successful approaches among the methods available. In this paper the multiband wavelet transforms and singular value decomposition are discussed to establish a watermarking strategy rather than biometric information. The use of biometrics instead of conservative watermarks can enhance information protection. The biometric technology being used is iris. The iris template can be viewed as a watermark, while an iris mode of communication may be used to help information security with the addition of a watermark to the image of the iris. The research involves verifying authentication against different attacks such as no attacks, Jpeg Compression, Gaussian, Median Filtering and Blurring. The Algorithm increases durability and resilience when exposed to geometric and frequency attacks. Finally, the proposed framework can be applied not only to the assessment of iris biometrics, but also to other areas where privacy is critical.*

Keywords*: Biometric Watermarking, Multiband Discrete Wavelet Transform, Information Security, Frequency Domain Watermarking*

1. INTRODUCTION

In the second half of the 20th century, the World Wide Web or WWW phenomenon, demonstrated the economic value of offering digital media free of charge. The use of digital networks to include digital media for commercial purposes is expected of multinationals (MNCs) [1]. But also, their ownership must be protected. It is thus a popular encryption method that is using encryption and other alternative methods (digital watermarking). Because of using advanced tools for copying and modifying multimedia data, safety becomes a major problem. Digital multimedia data are therefore very critical to preserve. Digital information, like digital pictures, audial and audio-visual, is widely available. Might not display the watermark. For photographs and videos, Visible watermarks are used, but they typically ruin elegance, whereby the location of the imprint is exposed to the attacker. It ran to an invisible watermark's popularity, when the watermark's position is not public. In the space or transformation field, the invisible watermarking can be performed.

The presented approach is a variant range due to the additional robustness of the same function [2].

Watermarking can be applied using a number of techniques in the transform domain, including the Fourier Transform, Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), Singular Value Decomposition (SVD), and others. A hybrid conversion domain based on DWT and SVD is used in this example. This is because DWT's multi-resolution function enhances recognizability, while SVD improves the system's robustness [3,4]. Unlike the conventional method, which uses an image or signal as a watermark, the user's iris biometric information is used as a watermark in this method. In this case user ID is used, like how a logo can be used as a watermark in different ways. The security level has improved many times since biometric technology is based on the principle of "what are you," as opposed to the conventional watermarking approach [5].

The Iris biometric watermarking enhances the security of the system by using the Fourier Transform technique. The encryption is processed along with the bio-

metric watermarking. The watermarking technology is used to protect the biometric data. Mainly the proposed method focuses on increasing the security of the iris biometric using watermarks from the host rather than the authentication from the client side. Fig1 shows

The working of the Iris Recognition System and the Feature Extraction using Fast Wavelet Transform an Iris image is segmented, normalized, Feature Extraction is done with binary signature, comparison is done and the Iris is stored in the Template.

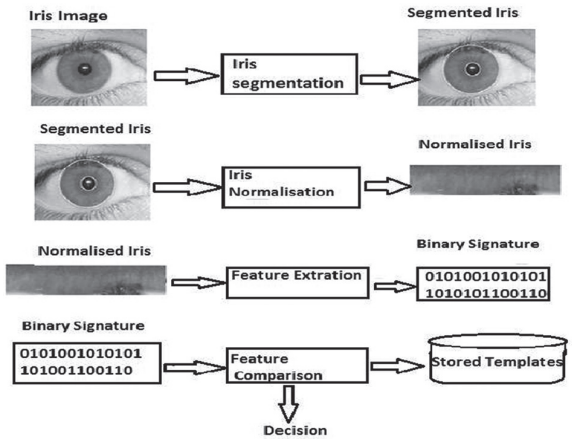


Fig. 1. Working Flow of Iris Recognition System

1.1. ADVANTAGES OF BIOMETRIC WATERMARKING

Jain et al. [19] suggested watermarking technology, as a second protective action for biometric systems. The use of watermarking will increase the safety and reliability of the biometric system efficiently [7]. Biometric watermarking benefits are listed:

The watermark is invisibly hidden in host data and can be used as an authentication token for security purposes. When biometric data are intercepted, a tracking system may be established to identify the data source.

- The host data has been connected to the watermarked information so that there is no further storage or transfer of resources needed. Furthermore, 'on-site' authentication doesn't require biometric or watermark database privileges.
- Other protection tools will not be affected by watermark. Cryptographic operations or template security methods may either be used with watermarked data or with watermarked host data.

1.2. WATERMARKING AND IRIS BIOMETRIC TECHNOLOGY

Two approaches are taken to incorporate biometric technology and watermarking technology. First, watermarking biometric data is used to protect biometric data privacy and to improve safety as a watermarking host [8]. The second is the biometric watermark used for authenticating the host image. The second form of role here is [9]. Researchers used the second watermark security system primarily with fingerprints and facets on biometric host images in the past [10,11].

2. RELATED WORKS

Watermarking technology is used to hide information in the host data image in order to protect its integrity. For embedding data into photographs, there are many watermarking techniques. They are known as spatial domain technologies [12, 13] and frequency domain technologies [14, 15, 16, 17]. While possessing the lowest complexity and maximum load power, space technology is vulnerable to occurrences such as image manipulation and low-pass filtering [18].

The spatial domain technologies include the mathematical processing of image pixel data to emphasize spatial relationships and the process again defines the homogeneous regions based on linear edges.

The frequency domain is the space defined by Fourier transform. Frequency domain analysis is used to indicate how the signal energy can be distributed in the range of frequency.

A watermark-based approach is proposed in [19, 20] in which the biometric system is safe and immune to cautious manipulation and attacks. Here it remains no research has been done into the integrity and robustness of biometric data under multiple attacks. A new biometric watermarking algorithm is developed in this white paper to reduce the vulnerability of biometric systems to geometric and frequency attacks. The mixture of wavelet and LSB-based watermarking technology, a watermark fingerprint is produced by inserting a prototype or facial image into the fingerprint image. Watermarking technology based on wavelets can withstand frequency attacks which is vulnerable to geometric outbreaks. LSB-based watermarking procedures, on the other hand, can withstand geometric attacks but are vulnerable to frequent outbreaks.

Low et al. [21] to decompose the signature into a string of binary bits, use the Discrete Random Transform (DRT) and Principal Component Analysis (PCA). The three implanting and removal approaches are compared to assess the robustness and power of JPEG compression. They are Least Significant Bit (LSB) and Code Division Multiple Access (CDMA) spread range in the spatial domain, and also in DWT domain. Humanoid visual observation, Peak Signal-to-Noise Ratio (PSNR), and alteration factor are used to determine the output of these approaches (standardized Hamming distance). The results show that, regardless of having the easiest contact to biometric watermarks, the LSB process is extremely susceptible to JPEG compression. The CDMA spread range in the DWT area is complex, but it is further hard for JPEG compression.

By combining encryption and watermarking techniques, Fouad et al. [22] identified a method for protect-

ing iris patterns. A key protects the iris file, which is inserted in the original image using LSB and DWT technology. The second key specifies the containment location. Two keys are required for iris extraction (iris and embedded).

Majumder et al. [23] apply biometric watermarking the Discrete Wavelet Transform (DWT) method, SVD (singular value decomposition) of the host copies to obtain the eigenvectors. Using Discrete Cosine Transform (DCT) technology, extract iris features and include them in the function vector. The downside of this approach is that you cannot modify the procedure for iris feature extraction.

Paunwala and Patnaik [24] inserted fingerprint and iris structures, block-divided cover image. Convert each block to a two-dimensional DCT and mark it as having or not having edges. In the low-frequency coefficient 8 x 8 DCT block, the edge block is omitted because the biometric recognition feature is built.

To improve protection, Lu et al. [25] planned a method that relies on iris recognition somewhat than digital watermarking. For error control, DCT is applied to the iris pattern, and the resulting rate is encoded using the Bose-Chaudhuri-Hocquenghem (BCH) code. The host image is split down into four similar blocks. In BCH code, the particular value is inserted in each coefficient in the host image using the key attained by the DCT procedure. The inverse cosine transforms (IDCT) are used in the image after adjusting the DCT coefficients of the host image in the watermark. Here the watermark's intensity is determined by the key used. The results show that this process is capable for extracting the watermark.

3. PROPOSED BIOMETRIC WATERMARKING ALGORITHM

The iris copy is used as the base or cover image for watermarking, a regular logo is used as the watermark. Iris characteristics are obtained through the process of iris localization and standardisation. It is classified into two parts: watermark implanting and watermark removal.

3.1 PREREQUISITES OF WATERMARK ALGORITHM

The following are the steps involved in the pre-processing steps for the cover image and watermark.

In Dual Tree Wavelet Transform, a time-frequency investigation method is used. Wavelet Transform allows for multi-resolution attribute study. In the four sub images with each point, equal sizes will be detected. LL_k will be chosen as estimated sub images, with $LH_{k'}$ $HL_{k'}$ and HH_k referring to horizontal, vertical, and diagonal direction greater-frequency aspect sub images.

The $k = 1,2,3,...(k \in N)$ will be referred to as the scale or status of the wavelet breakdown.

In the distribution of parallel spatial localization and frequency distribution relating to the watermark within the host image, the Multiband Dual Tree Discrete Wavelet Transform will be heavily used. The fundamental principle existed with the use of Multiband Dual Tree Discrete Wavelet Transform for image processing would be with many distinct halts of the image into sub images comprising diverse spatial domains and autonomous frequencies.

The data may be twisted within the calculation of the arrangement for implanting the watermarking procedure inside the element, which may be involved in debasing the intensity of the created watermark. Estimate, sub-image coefficients are chosen for imbedding watermark to ensure the provision of an invisible watermark as well as better power. Fig. 2. shows the conversion of an isolated wavelet has been completed.

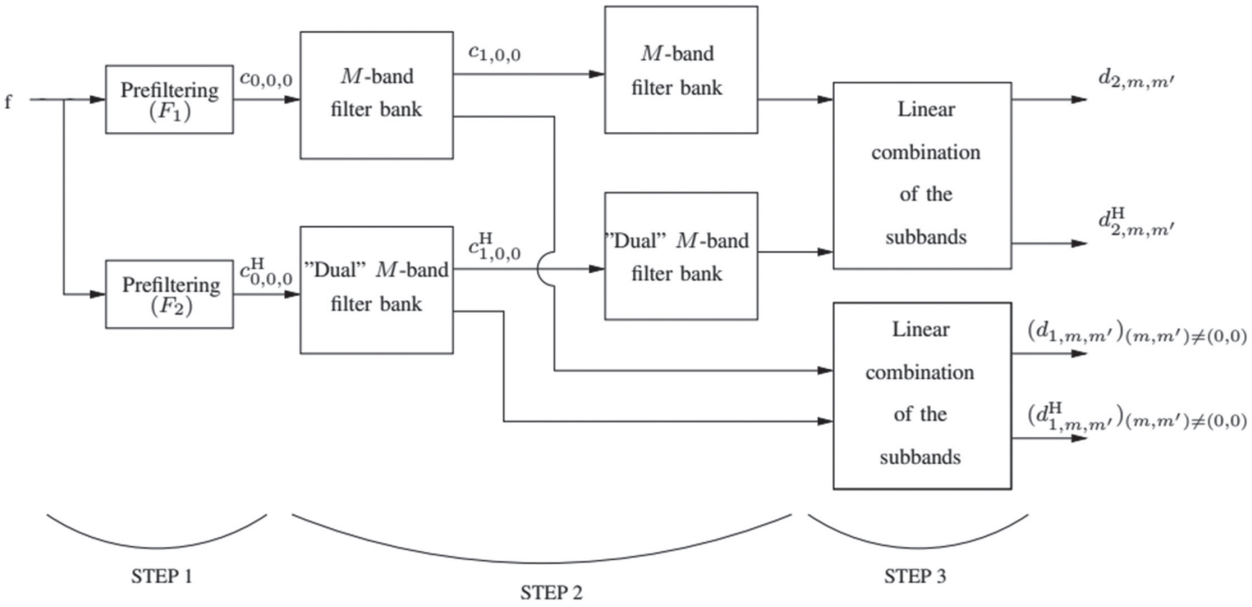


Fig. 2. M-band Dual-Tree Decomposition Scheme Over 2 Resolution Level

1) Singular Value Decomposition (SVD)

The process of Singular Value Decomposition (SVD) involves breaking down a matrix A into the form. This computation allows us to retain the important singular values that the image requires while also releasing the values that are not as necessary in retaining the quality of the image.

The original matrix A will be designated as an image with a size of $m*n$, and the breach down of A will be provided in the following expression:

$$A = U\,S\,V^{T} \tag{1}$$

Singular Value Decomposition of A will be the breakdown process, as U is a unitary matrix of dimension $m*m$, S is a matrix of positive numbers to the diagonal and Zeros found in positions other than the diagonal of dimension $*n$, and the conjugate transpose of VT, is unitary of measurement $n*n$. The luminance of the image will be represented by positive values of matrix S. Changing positive S ideals would not affect image excellence. They cannot be changed, either, even if there's an intruding impact. Watermarking techniques took advantage of these characteristics.

2) Arnold Transform

Arnold transform is widely used in image stenography, authentication, and tamper detection, self-recovery and image cryptography algorithms. In all these cases, Arnold transform is used as a scrambling step in which the number of iterations is used as a key.

Arnold Transform can perform arbitrary positioning of individual picture elements in relation to the original image. However, if the process is repeated for enough times, it is found the unchanged image reappears. In Arnold Transform, a change in position from one point to the next can be found. The Arnold transformation will be used to convert a digital image with dimension $N*N$.

$$\begin{bmatrix} i' \\ j' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} mod(n) \tag{2}$$

While $i, j \in \{0, 1,...., N\text{-}1\}$, while (i, j) are the coordinates that define the location of pixels in the actual image. (i', j') will coordinate that describe the location of pixels after the entire coordinates have been transformed, resulting in a disoriented image.

3.2. PROPOSED WATERMARKING ALGORITHM USING MULTIBAND DTDWT

The initial generation of the Watermark image will take into account for the iris image, and then the image will be transformed into one that only contains binary values. Following that, the Iris image will go through a 3-Stage Wavelet breakdown with 2D Coefficients. The watermark will be implanted after the Approximation sub-band LL3 has been selected. Prior to the implanting process, the created watermark will be converted to Arnold format. Following that, watermarked coefficients will be created by applying the SVD transform to the watermarked image along with the sub-band of choice. Following that, the watermarked image with the implanted watermark will be generated using the inverse Discrete Wavelet Transform. The procedure for extracting the watermark would be reversed from the procedure used to implant it.

Fig. 3 depicts the proposed algorithm's step-by-step movement. The iris image will be used as the concealed image, while the binary image will be used as the watermark. The proposed procedure is separated into two sections: 1. Embedding and 2. Extraction.

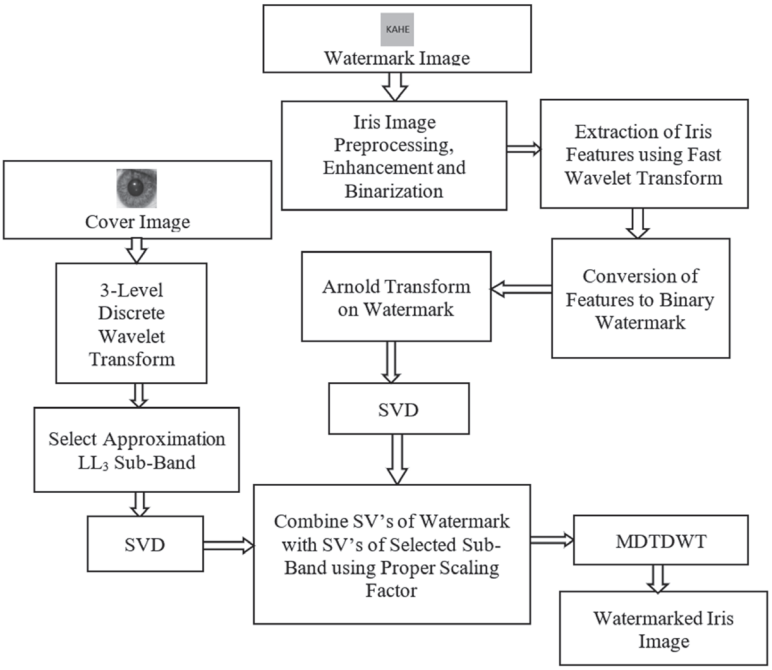


Fig. 3. Watermark Implanting Model

3.2.1. Watermark Embedding Scheme

The following are the steps involved in embedding the Watermarks into the host image:

Step 1: Multiband DTDWT is used to break down an iris image. The LL3 approximation sub-band will be chosen from all of the obtained sub-bands.

Step 2: Pre-analysis of Iris Image Breaks in the Iris Image will be observed, and fake points will be produced. The mining would subsequently be performed by standardisation and frequency assessment. The use of measured frequencies will be defined for filtering using Gabor Wavelet.

Step 3: The Multiband Wavelet Transformation is the abstraction of iris features and the replacement of features of iris images.

Step 4: Convert the Watermarked Image $_{w}$ using Arnold Transform.

Step 5: The Watermarked Coefficients A_W is achieved by accepting below three stages:

1. $A = U S V^T$

2. $S + \alpha_W = U_W S_W V_W^T$ While α is a watermark asset

3. $A_W = U S_W V^T$

Step 6: Reverse wavelets are used to modify the picture from a true double precision to an unsigned 8 bit integer. The picture of a watermark is created in relation to the implantation of a watermark.

Step 7: The procedure for embedding watermark is completed.

3.2.2. Watermark Extraction Scheme:

The watermarking can be extracted by executing the opposite order of watermark imbedding procedure.

Step 1: The low-frequency wavelet coefficient LL_3 is selected to perform a multiband dual tree wavelet transform on a watermarked image.

Step 2: Implement Singular Value Decomposition (SVD) with respect to A^*, the instruction is obtained as $A^* = U^* S_1^* V^{T*}$, along with accomplish U^*, S_1^* and V^{T*}.

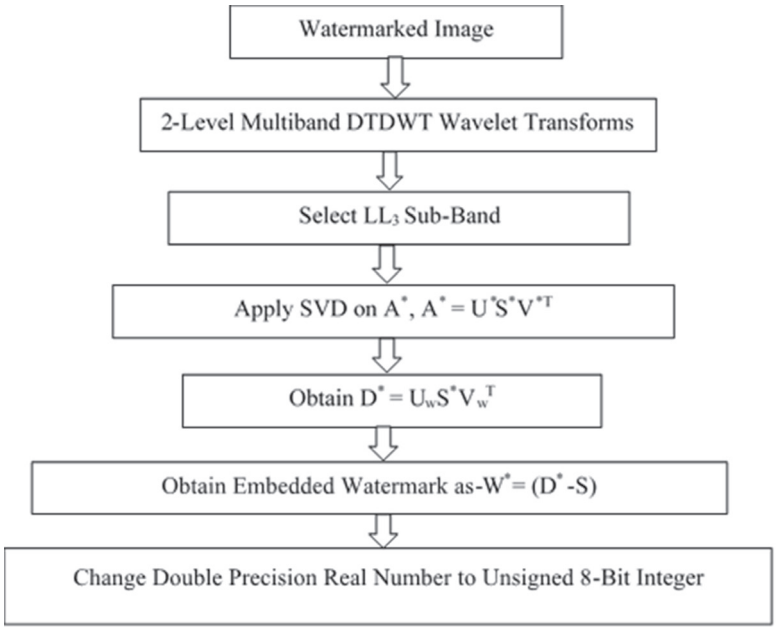


Fig. 4. Watermark Extraction Model

4. EXPERIMENTAL RESULTS

In unconditioned cases, the functionality of recognition schemes based on iris images would not produce accurate results. The availability of exactly collected iris image datasets with sufficient measurements will decide the outcome of the research involving the specified complication. Since knowledge is minimal, iris-based identification system analysis and investigation can be done by using the CASIA Iris Image Database sample is shown in Fig. 5.

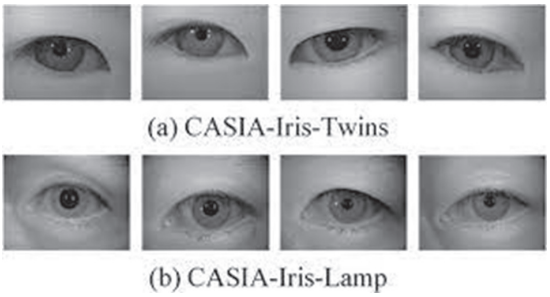


Fig 5. Sample CASIA Iris Image Database

A. Statistics and Descriptions of the Database

CASIA-IrisV3 has three following subdivisions: CASIA-Iris-Interval, CASIA-Iris-Lamp, and CASIA-Iris-Twins. CASIA-IrisV3 will contain large number of iris images: 22,035 iris images from 700 subjects. The entire set of images will be captured as an 8-bit monochromatic JPEG with near-infrared lighting.

B. Performance Analysis

CASIA iris images are used for iris recognition experiments. Each iris illustration is included with ten iris images taken at various times; the first copy will be used to implant the watermark, while the outstanding nine images will be used for identification. Table 1 shows the results of the recognition experiment.

Figure 6 shows the obtained images and extracted watermark after the attacks. The result will be accessed with the metric such as Correct Recognition Rate (CRR), False Accept Rejection Rate (FAR) and False Rejection Rate (FRR).

The results depicts that the iris biometric data is protected from the being theft by the anonymous users.

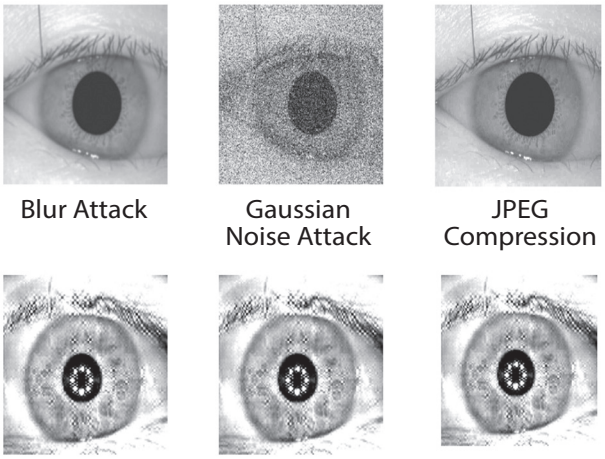


Fig 6. Obtained Result Images

Table 1. Biometric Authentication using Iris

Algorithm/Attacks	Proposed Multiband DTDWT			DWT based Watermarking		
Nature of Attacks	**CRR**	**FAR**	**FRR**	CRR	FAR	FRR
No Attack	**99.4%**	**0.0014**	**0.124**	99.4%	0.0014	0.124
Jpeg Compression Q=70%	**97.3%**	**0.0007**	**0.0041**	96.3%	0.0009	0.0049
Gaussian filtering with Var=3	**98.3%**	**0.0025**	**0.0725**	97.1%	0.0021	0.09
Median Filtering (3x3)	**97.3%**	**0.0041**	**0.034**	96.5%	0.0040	0.03
Blurring	**97.1%**	**0.0030**	**0.0281**	96.9%	0.0045	0.0881

For evaluating the functionality of suggested biometric watermarking approaches in the context of the following scenarios: Jpeg Compression kept at 70% quality, Gaussian filtering with variance kept as 3pixel, median filtering in 3x3 Window and blurring with (3*3) Mask. The proposed Multiband DTDWT rate is higher than the existing DWT based Watermarking using Iris Biometric Authentication System.

5. CONCLUSION

Multi-band DTDWT watermarks can withstand frequency attacks, while DWT-based watermarks can resist regular attacks. This article proposes a new biometric image watermarking procedure that syndicate Multi-band DTDWT and SVD-based algorithms that can increase durability and resilience when exposed to geometric and frequency attacks. The biometric recognition procedure is used as an index, comprehensive recital of the iris biometric watermark. The experimental outcomes show that the watermark procedure is more robust against geometric and frequency attacks compared to DWT-based watermarking. This algorithm can also preserve the integrity of the iris biometric template. Finally, the proposed framework can be applied not only to the assessment of iris biometrics, but also to other areas where privacy is critical. For example, user profiles on social media can provide benefits such as content or travel advice, but also recover sensitive information from seemingly anonymous data.

The proposed Multiband discrete wavelet transform and singular value decomposition technique enhances the security of the iris biometric system than the existing authentication techniques.

- We embed watermark in the principle components of the multi-band discrete wavelet coefficients. Specifically, the watermark signal is embedded into the principle components of the multi-band wavelet coefficients corresponding to the same spatial location at the same scale. With such a well-chosen embedding domain, the watermark is robustly and efficiently distributed to every detail frequency sub band. Our experimental results have shown that the watermark thus embedded has better invisibility and is more robust against JPEG compression than watermarks embedded in the DWT domain.
- Parameterized multi-band wavelet leads to a more secure watermark embedding domain, which makes the attack more difficult.
- Different from many other watermarking schemes,

in which watermark detection threshold is chosen empirically, the detection threshold of the proposed watermarking scheme can be calculated according to the targeted false positive.

6. REFERENCES

[1]. S. Katzenbeisser, F. A. P. Petitcolas, "Information hiding techniques for steganography and digital watermarking", Artech house, Computer security series, 2000.

[2]. N. F. Johnson, Z. Duric, S. Jajodia, "Information hiding, steganography", Kluwer Academic Publisher, 2003, pp. 15-29.

[3]. S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 11, 1989, pp. 674-693.

[4]. X. Zhu, J. Zhao, H. Xu, "A digital watermarking algorithm and implementation based on improved SVD watermarking-attacks and counter measures", Proceedings of the 18th IEEE Conference on Computer Vision and Pattern Recognition.

[5]. Iris recognition, https://en.wikipedia.org/wiki/Iris_recognition#Advantages (accessed: 2022)

[6]. T. Hoang, D. Tran, D. Sharma, "Bit priority based biometric watermarking", Proceedings of the Second International Conference on Communications and Electronics, Hoi An, Vietnam, 4-6 June 2008, pp. 191-195.

[7]. Technavio, https://www.techavio.com/report/iris-recognition-market-industry-analysis (accessed: 2022)

[8]. J. Dong, T. Tan, "Effects of watermarking on iris recognition performance", Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision, Hanoi, Vietnam, 17-20 December 2008.

[9]. S. Majumder, T. S. Dutta, "Watermarking of data using biometrics", Handbook of research on computational intelligence for engineering, science and business, IGI Global, 2020, p. 623.

[10]. Q. Zhao, "Advanced information security technology: watermarking and biometrics", ACM-HK Student Research and Day, 2009.

[11]. G. Varbanov, P. Blagoev, "An improving model watermarking with iris biometric code", Proceedings of the International conference on Computer systems and technologies, June 2007.

[12]. D. P. Mukherjee, S. Maitra, S. T. Acton, "Spatial domain digital watermarking of multimedia objects for buyer authentication", IEEE Transactions on Multimedia, Vol. 6, No. 1, 2004, pp. 1-15.

[13]. A. K. Singh, N. Sharma, M. Dave, A. Mohan, "A novel technique for digital image watermarking in spatial domain", Proceedings of the 2nd IEEE International Conference on Parallel, Distributed and Grid Computing, Solan, India, 6-8 December 2012, pp. 497-501.

[14]. Y. Wang, J. F. Doherty, R. E. Van Dyck, "A wavelet-based watermarking algorithm for ownership verification of digital images", IEEE Transactions on Image Processing, Vol. 11, No. 2, 2002, pp. 77-88.

[15]. W. Wang, A. Men, X. Chen, "Robust image watermarking scheme based on phase features in DFT domain and generalized radon transformations", Proceedings of the 2nd International Congress on Image and Signal Processing, Tianjin, China, 17-19 October 2009, pp. 1-5.

[16]. M. N. Sakib, S. B. Alam, A. B. M. R. Sazzad, C. Shahnaz, S. A. Fattah, "A basic digital watermarking algorithm in discrete cosine transformation domain", Proceedings of the Second International Conference on Intelligent Systems, Modelling and Simulation, Phnom Penh, Cambodia, 25-27 January 2011, pp. 419- 421.

[17]. K. Deb, M. S. Al-Seraj, M. M. Hoque, M. I. H. Sarkar, "Combined DWT-DCT based digital image watermarking technique for copyright protection", Proceedings of the 7th International Conference on Electrical and Computer Engineering, Dhaka, Bangladesh, 20-22 December 2012, pp. 458-461.

[18]. P. Dabas, K. Khanna, "A study on spatial and transform domain watermarking techniques", International Journal of Computer Applications, Vol. 71, No. 14, 2013, pp. 38-41.

[19]. A. K. Jain, U. Uludag, "Hiding biometric data", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 25, No. 11, 2003, pp. 1494-1498.

[20]. M. Vatsa, R. Singh, A. Noore, "Improving biometric recognition accuracy and robustness using DWT and SVM watermarking", IEICE Electronics Express, Vol. 2, No. 12, 2005, pp. 362-367.

[21]. C. Y. Low, A. B. J. Teoh, C. Tee, "A preliminary study on biometric watermarking for offline handwritten signature", Proceedings of the IEEE International Conference on Telecommunications and Malaysia International Conference on Communications, Penang, Malaysia, 14-17 May 2007, pp. 691-696.

[22]. M. Fouad, A. El Saddik, Z. Jiying, E. Petriu, "Combining cryptography and watermarking to secure revocable iris templates", Proceedings of the IEEE International Instrumentation and Measurement Technology Conference, Hangzhou, China, May 2011, pp. 1-4.

[23]. S. Majumder, K. J. Devi, S. K. Sarkar, "Singular value decomposition and wavelet-based iris biometric watermarking", IET Biometrics, Vol. 2, No. 1, 2013, pp. 21-27.

[24]. M. Paunwala, S. Patnaik, "Biometric template protection with DCT- based watermarking", Machine Vision and Applications, Vol. 25, No. 1, 2014, pp. 263-275.

[25]. J. Lu, T. Qu, H. R. Karimi, "Novel iris biometric watermarking based on singular value decomposition and discrete cosine transform", Mathematical Problems in Engineering, Vol. 2014, 2014, pp. 1-6.

Feature Extraction Method using HoG with LTP for Content-Based Medical Image Retrieval

Original Scientific Paper

NV Shamna
P A College of Engineering,
Department of Computer Science and Engineering,
Mangalore, India
shamnanv@gmail.com

B. Aziz Musthafa
Beary's Institute of Technology,
Department of Computer Science and Engineering,
Mangalore, India
bazizmusthafa@gmail.com

Abstract *– An accurate diagnosis is significant for the treatment of any disease in its early stage. Content-Based Medical Image Retrieval (CBMIR) is used to find similar medical images in a huge database to help radiologists in diagnosis. The main difficulty in CBMIR is semantic gaps between the lower-level visual details, captured by computer-aided tools and higher-level semantic details captured by humans. Many existing methods such as Manhattan Distance, Triplet Deep Hashing, and Transfer Learning techniques for CBMIR were developed but showed lower efficiency and the computational cost was high. To solve such issues, a new feature extraction approach is proposed using Histogram of Gradient (HoG) with Local Ternary Pattern (LTP) to automatically retrieve medical images from the Contrast-Enhanced Magnetic Resonance Imaging (CE-MRI) database. Adam optimization algorithm is utilized to select features and the Euclidean measure calculates the similarity for query images. From the experimental analysis, it is clearly showing that the proposed HoG-LTP method achieves higher accuracy of 98.8%, a sensitivity of 98.5%, and a specificity of 99.416%, which is better when compared to the existing Random Forest (RF) method which displayed an accuracy, sensitivity, and specificity of 81.1%, 81.7% and 90.5% respectively.*

Keywords*: CE-MRI dataset, Content-Based Medical Image Retrieval, Histogram of Gradient, Local Ternary Pattern*

1. INTRODUCTION

In recent days, computer-aided techniques and tools are frequently employed in medical diagnosis for better health management. Modern computer-aided medical technologies help clinicians in several areas, such as in the diagnosis and treatment of specific diseases [1]. Several computer-based approaches like Computed Tomography, Optical Projection Tomography (OPT), Magnetic Resonance Imaging (MRI), X-Ray, Angiography, Digital Mammography (DM), Ultrasonography, Optical Endoscopy (OE), Nuclear Medical Imaging (NMI), and Positron Emission Tomography (PET), give visual information that aids in the process of medical treatment and diagnosis [2]. This kind of medical photography / imagery offers details about various bodily sections needed for disease detection, diagnosis, monitoring, and therapy [3]. It takes a lot of work to compile computer-aided photographs in one place. Image retrieval has become essential for academicians, companies, and medical institutions to recover crucial information from the database [4]. Content-Based Image Retrieval (CBIR) is used in many industries, including healthcare, law enforcement, criminal justice, and heritage. CBIR is a unique tool for searching visual content amid a bigger amount of data and for retrieving images [5].

Content-Based Medical Image Retrieval (CBMIR) gives suitable results in extracting features from medical data [6]. The content-based retrieval system analyses a particular database consisting of medical images, because the database includes information related to the ranks and features of images. Those that are most comparable to the query photographs are found during retrieval, using similarity characteristics of the database [7]. The present content-based retrieval method retrieves images using similarity measures, lower-level features, and semantic gap reduction [8]. The primary challenge in CBMIR image search is identifying pertinent data amongst available photos for the given que-

ry images. Identification of similar instances based on comparable anatomy acts as a virtual peer review for diagnosis in the field of medicine [9]. Two main issues with the current CBMIR approaches are feature extraction and distance measuring. To increase the efficacy of CBMIR technique, numerous research studies are developed. However, technologies that are now in use are still unable to deliver adequate, and crucial data from image visuals, during the process of medical image retrieval [10]. The major contributions of this research are mentioned as follows,

- A feature extraction method is suggested by combining Histogram of Gradient (HoG) with Local Ternary Pattern (LTP) to automatically obtain medical images from the CE-MRI database in order to boost efficiency and reduce computational cost.
- After that, the Euclidean measure is used to calculate the similarity of the query images, and Adam optimizer is used to choose the features.
- The use of HoG and LTP to automatically retrieve medical images from CE-MRI database is offered as a solution of the Adam Optimizer method for the existing issues.

The remaining paper is given as follows: a review of existing approaches is explained in section 2, and the proposed HoG-LTP used to automatically retrieve medical images from a larger database, is explained in section 3. The results and discussions of the research work are given in section 4. The conclusion along with a brief mention of future work is presented in section 5.

2. LITERATURE REVIEW

Techniques based on content-based retrieval of images, that are currently in existence are reviewed in this section. The advantages and limitations of these methods are also described here.

Swati et al. [10] developed an automatic CBIR method based on fine-tuning and transfer learning to retrieve brain tumor-related image content. The features were extracted to identify regions with a brain tumor and were calculated using the Closed Form Metric Learning (CFML) and deep CNN VGG1. Transfer learning approach, along with fine-tuning based block-wise approach was used for the detection of tumor. The performance was carried out on CE-MRI datasets which displayed an improved value of the CFML method in comparison to existing methods. Due to time and space constraints, the method focused solely on tumor retrieval from brain MRI images.

Veerashetty et al. [11] developed an efficient Manhattan distance approach for CBMIR, wherein the histogram of gradients approach was executed. The CBMIR method was divided into two stages: extracting features and matching the features. Initially, the filter of the Gaussian model was used to enhance medical images. The established CBMIR method was used to extract preprocessed features of medical images. Euclidian distance measure was used to calculate and match the similarity between training and testing medical images. The developed method had the advantage of performing better in the retrieval of medical images. However, multiple features that were required to achieve accurate results were not extracted.

A CBMIR method known as Manhattan-distance-based Histogram of Oriented Gradients (M-HOG) was given by Ahmed et al. [12], in which the RGB-based images that were collected were first converted into Hue Saturation Value (HSV) colour space. A group of 18 features was extracted using feature colour and texture function of Grey Level Co-occurrence Matrix (GLCM), and comparable scores were computed. According to the ranking system, top photographs were selected from the query search feedback method. While employing similar coefficients, the performance was increased by means of precision and recall parameters. In the case of extracting features, the application of comparable coefficients to each pixel in the image was a time-consuming operation.

Bressan et al. [13] developed an effective diagnosis method for breast cancer detection that depended on medical active learning for a content-based image retrieval process. The proposed effective diagnosis resulted in high performance in inherent constraints and medical contexts. The group of highly informative images included a selection of conditions like uncertainty and diversity. The similarity developed using an appropriate active learning strategy. Results were evaluated in medical-based queries with an improved precision value given by the developed diagnosis method. The process of learning took more computational time in the presented content-based image retrieval process.

Zhang et al. [14] presented a triplet deep hashing approach based on the image retrieval framework's privacy preservation. Triplet deep Convolutional Neural Network (CNN) was introduced to evaluate the hash codes and visual results. The searching process in the image was improved using a suitable hash code based on S-Tree of Hierarchical Hashing S-Tree (H2S-Tree). The process of retrieving images in the cloud became easier by using an efficient structure of index in triplet deep hashing approach, which also provided an effective learning quality of hash code. Efficiency and accuracy were improved using the proposed triplet deep hashing approach, but the 1024 bits longer hash code resulted in a higher cost of storage.

Khatami et al. [15] developed a parallel deep solution method based on CNN to reduce imbalanced datasets. The majority vote with the HOG, Radon features, and Local Binary Pattern (LBP) was used in finding the best match query images. The capability of the parallel deep solution approach was improved in terms of error rate and accuracy using the IRMA dataset. The parallel deep learning approaches gave very satisfactory results in the process of feature extraction in medical fields, as compared to the existing techniques.

But, the unlabeled datasets and noisy data in the proposed deep solution approach led to unsatisfied results.

The feature extraction technique for multiple object recognition utilizing Histogram of Oriented Gradient (HOG) with Local Ternary Pattern (LTP) was shown by Akanksha and Rao [16]. The suggested technique used the Caltech 101 dataset. The proposed HoG and LTP were used to identify important regions from the image during the feature extraction procedure. Additionally, in order to detect many objects, the Deep Convolutional Neural Network (D-CNN) was utilized to fuse the collected features before sending them to the Region-based Convolutional Neural Network (R-CNN). By successfully extracting the orientation features and texture features, the suggested HoG and LTP feature extraction methods had the advantage of increasing the classification accuracy.

3. PROPOSED METHODOLOGY

The architecture and description of the proposed HoG with the LTP method is presented in this section. The diagram of the proposed feature extraction method using HoG with LTP, for CBMIR, is given in Fig. 1.

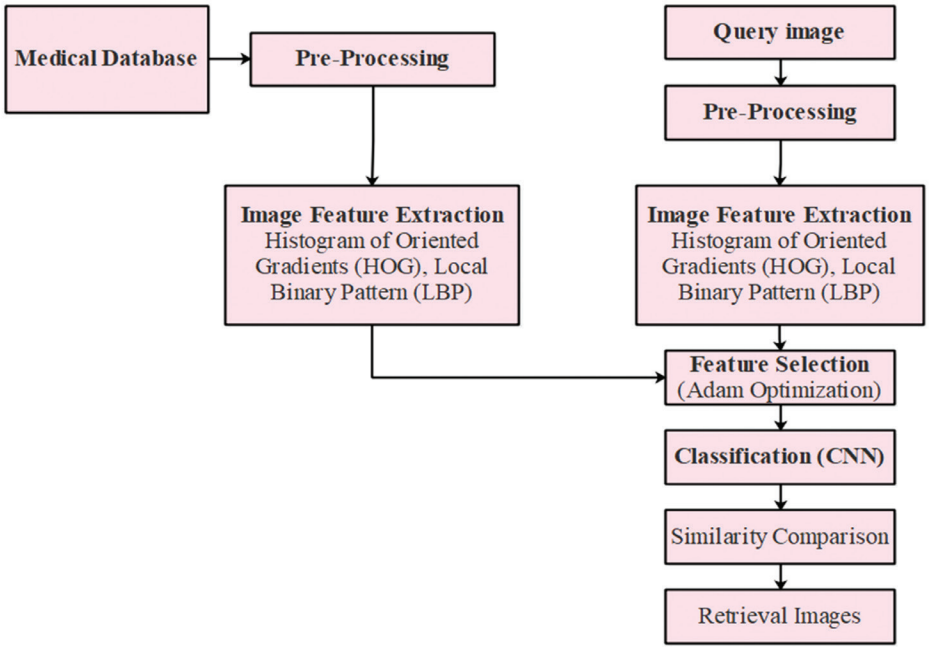


Fig. 1. The proposed block diagram of extracting features using HoG with LTP for content-based image retrieval

3.1. DATASET AND PREPROCESSING

The publicly accessible Contrast-Enhanced Magnetic Resonance Imaging (CE-MRI) dataset is utilized in the proposed HoG-LTP method. This brain tumor dataset comprises 3064 T1-weighted contrast-enhanced images containing three kinds of brain tumors. It contains 708 slices of images with Meningioma tumor, 1426 slices of tumor images with Glioma, and 930 slices of the tumor with Pituitary, consisting of pixel size

512×512 and the matrix size being 49mm×49mm. The dataset is collected from 233 patients that form a total of 3064 images containing axial, sagittal, and coronal visions. The CE-MRI datasets were gathered from the hospitals of Nan Fang, Guangzhou, China, and the Medical University of Tianjin located in China between 2005 and 2010. The 2D CE-MRI images have high semantic gaps and have tumors of Glioma, Pituitary, and Meningioma. The data is divided into a training set and a test set in the following stages: (i) The proposed system is trained using the training set, which contains 80% of the images, and (ii) The remaining 20% goes to the test set which is classified using the trained network to determine the accuracy.

3.2. PREPROCESSING

The collected greyscale images with a pixel size of 512×512 are preprocessed by reframing the size of the images. The size of the image is reduced by resizing it, decreasing the number of pixels by extracting relevant pixels and deleting the unwanted background. The image was framed with new width and height by removing the details of pixels. The same values are proportional to the scale of the image values used for both width and height.

3.3. FEATURE EXTRACTION

Grayscale image CE-MRI datasets employ the descriptor without any colour. HOG is used to produce the CE-MRI features, and LBP is used to assess the object's shape and appearance.

3.3.1 Histogram of Gradient

Histogram of Oriented Gradient (HOG) is used to detect the silent parts in the image during feature extraction after the CE-MRI images are preprocessed. The presence of gradient orientation is calculated in the spatial regions of the images, referred to as cells. Gradients of the image are calculated to extract features of HOG and an orientation histogram is created for every cell. Histograms are normalized for every cell, which results in specific blocks of the HOG descriptor. The HOG features are extracted as per the following steps:

The images are converted to greyscale after preprocessing and then are convoluted into vertical and horizontal masks to compute gradients [-1 0 1] and [-1 0 1]^T. Horizontal and vertical mask gradients are represented as shown in equations (1) and (2).

$$G_x(x,y) = [-1\ 0\ 1] * I(x,y) \tag{1}$$

$$G_y(x,y) = [-1\ 0\ 1]^T * I(x,y) \tag{2}$$

Where * represents the convolution; I is the pre-processed image; the orientation at each pixel (θ) that is computed using the gradient ratio in vertical and horizontal directions, is stated as $G_x(x, y)$, $G_y(x, y)$ which is given in equation (3).

$$\theta(x,y) = \arctan\frac{G_x(x,y)}{G_y(x,y)} \tag{3}$$

Further, every block is separated into $M \times N$ cells, where $N \geq M$. Pixels in every image are counted using a weighted vote, which shows the magnitude gradient at every pixel, and the votes are accumulated in orientation bins. The L bins are placed between $0°$ to $180°$ for un-signed gradients or $0°$ to $360°$ for signed gradients, and the l^{th} value of bins is shown in equation (4).

$$\Omega_l(x,y) = \begin{cases} G(x,y) & if\ \theta(x,y) \in bin_l \\ 0 & otherwise \end{cases} \tag{4}$$

Where pixels (x, y) are the magnitude of gradients represented as

$$G(x,y) = \sqrt{G_y(x,y)^2 + G_x(x,y)^2} \tag{5}$$

The features of HOG taken from every cell are stated as f_{cell}, and the generated features are the same as the number of bins, wherein the features are represented in equation (6).

$$f_{cell} = \frac{\sum_{(x,y)\in Block} G(x,y)+\epsilon)}{(\sum_{(x,y)\in Block} G(x,y)+\epsilon)} \tag{6}$$

Where ϵ is an immeasurable smaller quantity, and every cell's normalized features in the block, extracted from dimension vector d, are the same as the product of overall cells in the block and oriented bins, which is the final descriptor of the block.

3.2.2. Local Ternary Pattern (LTP)

A three-valued texture operator LTP provides a straightforward and efficient descriptor to describe the features. The new label is extracted from LTP label pixels of the image where the center value of every pixel based on neighborhood threshold is multiplied and added with a power of two. LTP is an extension of LBP where it mostly depends on center pixel values and pixel values existing between threshold (*t*) of *-t* to *+t*, with pixel value as zero. In case the pixel value is greater than the threshold, it results as 1 or -1 of the LTP operator.

The evaluation of the LTP operator is explained in equation (7).

$$LTP(i) = \begin{cases} 1 & if\ t \leq p_i - p_c \\ 0 & if\ |p_i - p_c| < t \\ -1 & if\ -t \geq p_i - p_c \end{cases} \tag{7}$$

Value t represents the user-specified threshold, the pixel value of the neighbor is p_i, and p_c is the central value of the pixel. Texture operators are obtained by LTP which are noise sensitive due to not being dependent on center pixel value and grey-level transformations are not followed. The LTP value is classified as upper LTP and lower LTP wherein the upper LTP is developed by applying zeroes for negative values, and lower LTP is obtained by replacing the values of 1 with zeroes and changing the negative values to one in the original LTP. The query images and datasets extract the feature data from HoG and LTP in the feature selection process.

3.4. FEATURE SELECTION

The features are selected from the CE-MRI dataset and query images using the Adam optimization algorithm.

3.4.1. Adam Optimization

The traditional stochastic gradient descent is updated as Adam Optimization which is an extension of stochastic gradient descent and productively results in the weights of the updated network. The 1st and 2nd order element is taken in the Adam Optimizer to evaluate the gradient's first-order element (mean of the gradient), and element of second-order (squared gradient based on the element), and then, bias is corrected. The learning rate is divided by the square root of the second-order moment and multiplied by the first-order element to get the update of the final weight.

The three hyper-parameters of LTP are: the learning rate, the decay rate of the first-order moment, and the decay rate of the second-order moment. This work shows that, if the optimal solution is predicted in advance, the learning rate generally takes the final resulting weight, making it simple to estimate the learning rate's scale. Firstly, the exponential moving average is biased towards zero, so it is divided by a number based on the decay rate to get an unbiased value. The Adam optimization algorithm consists of the below steps:

1. The element-wise square and gradient are computed by present parameters.

2. The 1st order and 2nd moment's exponential averages are updated and calculated as unbiased averages.
3. In the Adam optimizer, weight updation is calculated as the average of unbiased 1st order, which is divided by average of the square root of unbiased 2nd order (scaled by a learning rate of 0.001).

According to the Adam optimization algorithm, the weights are computed initially using equation (8).

$$w_t = w_{t-1} - \eta \frac{m_t^1}{\sqrt{v_t}+\epsilon} \quad (8)$$

$$m_t = (1-\beta_1)\sum_{i=0}^{t} \beta_1^{t-i} g_i \quad (9)$$

The second moving momentum is calculated by using equation (10).

$$v_t = (1-\beta_2)\sum_{i=0}^{t} \beta_2^{t-i} g_i^2 \quad (10)$$

The bias correction of momentum is calculated by using equations (11-12).

$$m_t^1 = \frac{m_t}{1-\beta_1^t} \quad (11)$$

$$v_t^1 = \frac{v_t}{1-\beta_2^t} \quad (12)$$

Where, $m_t^{\,1}$ is the moving momentum and $v_t^{\,1}$ is the second moving momentum in Adam optimization.

3.1. CLASSIFICATION

The related query images from the end user are extracted by using CNN classifier.

3.5.1 Convolutional Neural Network

The CNN model includes four various layers, presented as follows. The extracted features from input images are fed into convolutional layers, which compute the neurons that are connected to the input of local regions. Each neuron is computed using a dot product of smaller weights and the input image volume is connected. Activation function is then used to determine whether the neurons are right or not. The ReLU layer does not affect the dimensions of the images taken as input. Additionally, the pooling layer reduces the noise effect among the extracted features. Finally, the Fully Connected (FC) layer evaluates the higher-level features. The proposed method uses a pre-trained deep CNN model such as VGG19 to determine the position of the first four lumbar vertebral bodies. The VGG19 network has 19 learnable weight layers, 16 convolutional layers, and 3 FC layers with soft-max. Dropout regularisation is used in the FC layer of VGG19, and ReLU activation function is used in convolutional layers. Convolutional, pooling, ReLU, normalization, and FC layer are all parts of the VGG19 model. The convolutional layer detects the vertebral body by the images as shown in equation (13).

$$g_i^L = b_i^L + \sum_{j=1}^{m_1(L-1)} \psi_{i,j}^L \times h_j^{L-1} \quad (13)$$

Where $g_i^{\,L}$ denotes L output layer, $b_i^{\,L}$ is the base value, $\psi_{i,j}^{\,L}$ is the connection of the filter with i^{th} and j^{th} feature map, and output layer of L-1 is h_j. The pooling layer detects vertebral foramen in transverse slices and obtains maximum responses from the less convolutional layer to reduce unwanted features and solve the overfitting problem, which is explained using equations (14-16).

$$m_1^L = m_1^{L-1} \quad (14)$$

$$m_2^L = \frac{m_2^{L-1} - F(L)}{S^L} + 1 \quad (15)$$

$$m_3^L = \frac{m_3^{L-1} - F(L)}{S^L} + 1 \quad (16)$$

Where, S^L represents the strides wherein the neural network parameters are changed in the movement of images, $m_1^{\,L}$, $m_2^{\,L}$, and $m_3^{\,L}$ are the feature map filters and the layers; FC and ReLU, are explained in equations (17-18).

$$Re_i^l = max\left(h, h_i^{l-1}\right) \quad (17)$$

$$FC_i^L = f\left(z_i^l\right) with\ z_i^l = \sum_{j=1}^{m_1(l-1)} \sum_{r=1}^{m_2^{l-1}} \sum_{s=1}^{m_3^{l-1}} w_{i,j,r,s}^l \left(FC_i^{l-1}\right)_{r,s} \quad (18)$$

Where the layer of ReLU is represented as Re_i^l, h is the output layer, $FC_i^{\,L}$ is the layer of FC that follows the convolutional layer as well as pooling layer, and performs the FC layer activation function for the identification of deeper features.

4. RESULTS AND DISCUSSION

4.1. SIMILARITY CALCULATION

The Euclidean distance is used to calculate the similarity between CE-MRI images and classified query images. The resemblance between feature dataset and feature vectors of the input query image is calculated. During the image retrieval phase, the global feature vector of the query image is extracted using identified category bank. Finally, similar images from identified semantic category are retrieved and ranked using Euclidean distance between feature vector of the query image and similar candidate images. Effective sequential algorithms generate Euclidean distance maps. The map represents shortest distance from the nearby pixel in the background for each pixel present in the original binary picture's objects. The two picture scans with backward and forward movement of each line, are used to create a map including minor errors. As a result, successful computation is done with parallel propagation, which is an iterative process for expanding/shrinking purpose. In distance mapping, image processing is used frequently. Typically, it is based on one of the metrics calculated using equation (19).

$$d_4\big((i,j),(h,k)\big) = |i-h| + |j-k| \quad (19)$$

which is known as "distance of chessboard," or a combination of it, such as "octagonal distance" in the space of a two-dimensional rectangle with integers *i, j, h,* and *k*.

The binary image includes two sets of pixels labeled 0 and 1, the $L(S)$ which represents the distance map, is an image in which each pixel $(i, j) \in S$ corresponds to a pixel in $L(S)$, wherein the distance map is given in equation (20).

$$L(i,j) = \min\left(d[(i,j),S]\right) \tag{20}$$

Every pixel in S is having a particular $L(S)$ label which is equal to the neighboring distance, hence the same map $L(S)$ is defined for the background.

Content-based medical images are retrieved efficiently using the proposed feature detection; HoG with LTP method, for high-quality data. In contrast to the approaches that existed, this proposed feature detection method gives the desired results in recognizing features using CE-MRI dataset. The retrieval method is applied on a computer with 2.2 GHz, Python 3.7.3 and 8GB RAM. This proposed method demonstrates considerably good performance in extracting features from the image, in comparison to existing algorithms.

4.2. PERFORMANCE METRICS

The image retrieval performance of the proposed HoG with LTP uses 70% of training images and 30% of testing images. The proposed HoG with LTP is compared with existing techniques to check the nature of the system parameters. performance metrics considered in the assessment of HoG with LTP are explained as follows:

4.3.1. Accuracy

Accuracy is used to evaluate the model's classification by measuring ratio of the number of right predictions to total number of predictions, which is represented in equation (21).

$$Accuracy = \frac{Number\ of\ correct\ predictions}{overall\ predictions} \tag{21}$$

4.2.2. Sensitivity

Sensitivity measures the ratio of positive values calculated using equation (22).

$$Sensitivity = \frac{TP}{TP+FN} \tag{22}$$

4.2.3. Specificity

Specificity calculates the negative ratio that is predicted correctly and it is calculated using equation (23).

$$Specificity = \frac{TN}{TN+FP} \tag{23}$$

4.2.4. Precision

Precision is the ratio of overall truly classified positives to predicted positives. It is defined in equation (24).

$$Precision = \frac{TP}{TP+FP} \times 100 \tag{24}$$

4.2.5. F-score

F-Score is calculated from the precision and recall of the test model which is expressed in equation (25).

$$F - score = \frac{TP}{TP+1/2(FP+FN)} \tag{25}$$

4.2.6. Error Rate

The measure of error units present in the model is defined as the Error rate and measured by equation (26).

$$Error\ Rate=100\text{-}Accuracy \tag{26}$$

4.3. Performance Analysis

The proposed feature detection HoG and LTP methods are used to analyze CE-MRI datasets and provide much more optimal results as compared to existing SVM, RF and RCNN [16]. Results of the proposed HoG and LTP methods are compared to existing models in tables 1 and 2.

Table1. Comparison of HoG-LTP method with existing techniques related to sensitivity, specificity, and accuracy for image retrieving

Methods	Accuracy (%)	Sensitivity (%)	Specificity (%)
RF	81.1	81.7	90.5
SVM	97.7	97.8	98.8
RCNN [16]	92.48	75.24	71.32
HoG-LTP	98.8	98.5	99.4

Table 1 compares the proposed HoG and LTP techniques to the existing techniques which are RF, SVM and RCNN [16]. The proposed method achieves 98.8% accuracy, 98.5% sensitivity, and 99.4% specificity. Whereas the existing RF and SVM respectively obtained an accuracy of 81.1% and 97.7%, sensitivity of 81.7% and 97.8%, and specificity of 90.5% and 98.8%. While the existing RCNN obtains accuracy, sensitivity and specificity at 92.48%, 75.24% and 71.32% respectively. The output of comparison results of the presented HoG with LTP method proves its higher performance when it is compared to previous conventional methods in terms of sensitivity, specificity, and accuracy. Figure 2 compares the proposed HoG with the LTP method to previously existing methods in terms of sensitivity, accuracy, and specificity.

The proposed HoG-LTP method gives higher performance and the error rate is reduced as compared to existing models, which is shown in Figure 3. The retrieved images of Meningioma and Glioma tumors are shown in Fig. 4 and Fig. 5.

Table 2. The comparison of HoG-LTP and existing methods in terms of Error rate, Precision, and F-score for images to retrieve features

Methods	Precision (%)	Error Rate (%)	F-score (%)
RF	90	18	75
SVM	94	2	97
HoG-LTP	100	1	100

Table 2 compares the error rate, F-score, and precision of the proposed HoG-LTP method to those of existing techniques such as RF and SVM. The proposed method achieves a precision of 100%, an error rate of 1%, and an F-score of 100%, whereas, for existing models of RF and SVM, the precision value is 90% and 94%, the rate of error is 18% and 2%, and F-Score is 75% and 97%, correspondingly.

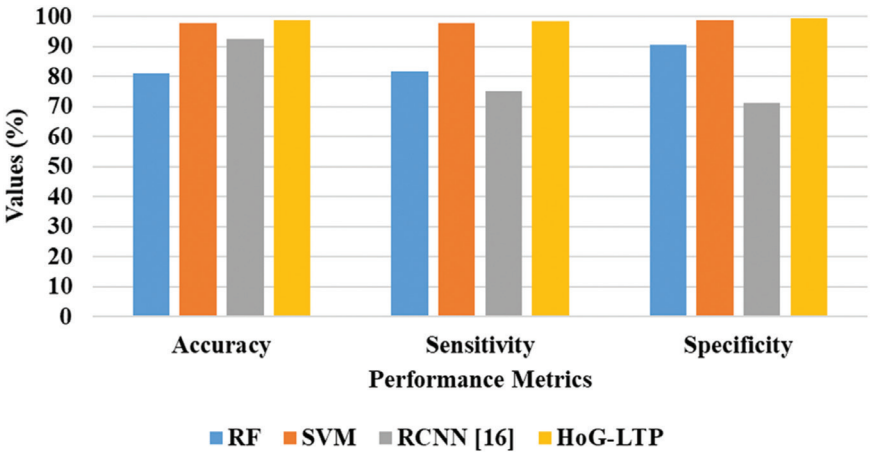


Fig. 2. The comparison of HoG-LTP and existing methods in terms of accuracy, sensitivity, and specificity

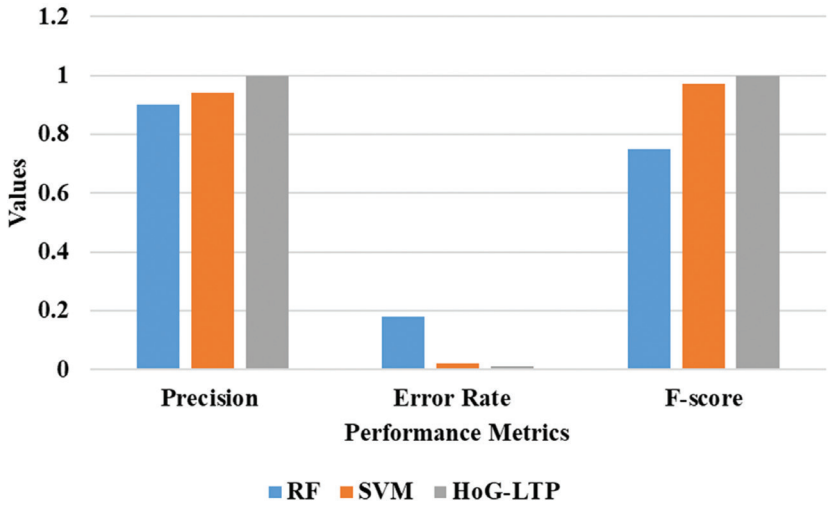


Fig. 3. Comparative graph of HoG-LTP with existing methods in terms of error rate, precision, and f-score

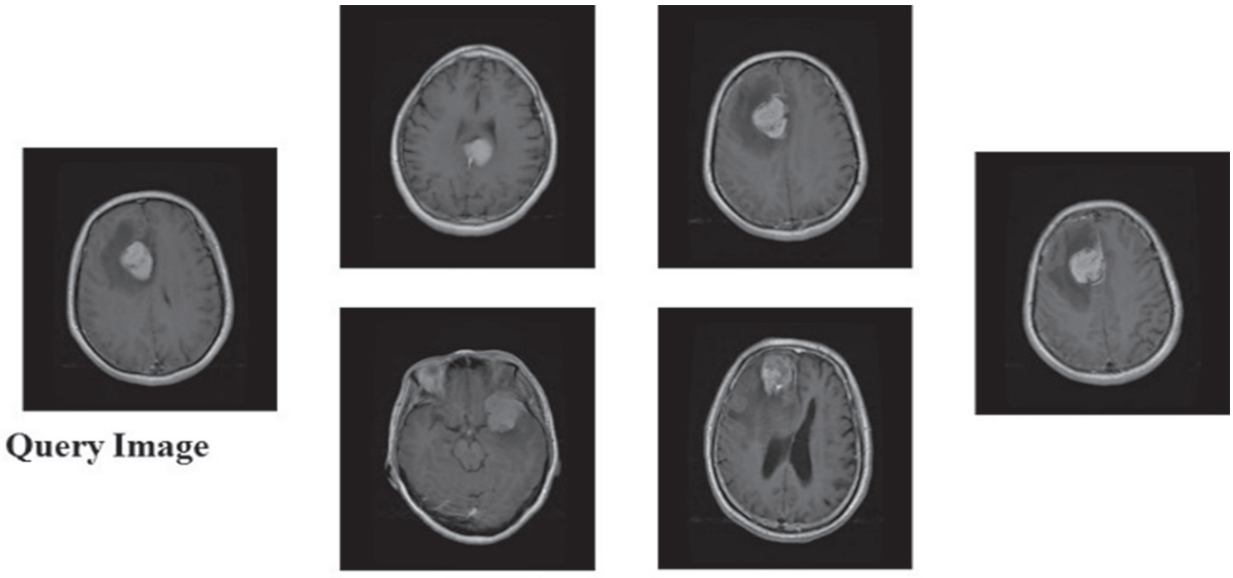


Fig. 4. Meningioma tumor query-based retrieved images

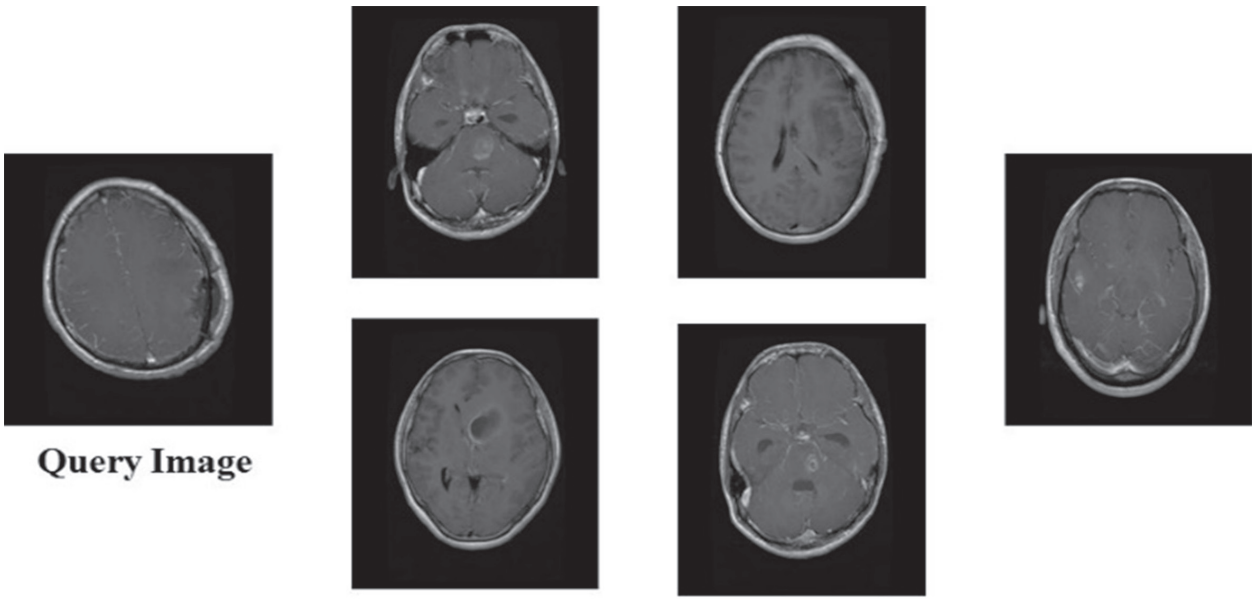


Fig. 5. The images that are retrieved for the glioma tumor image query

5. CONCLUSION

The main objective of the proposed HoG and LTP models using CB-MIR is, to extract relevant features from medical images for accurate diagnosis of disease. To overcome issues persisting in traditional methods, a novel feature selection method is proposed, which combines HoG and LTP to retrieve medical images automatically from CE-MRI dataset. The size of images in CE-MRI is reconstructed to modify the pixels, which results in a reduction of image size. HOG extracts sufficient information on global features from CE-MRI, and LBP provides an appearance of the object and its shape. Adam optimization algorithm is used to select features from CE-MRI dataset and query images. The extracted features are fed into the CNN classifier, which uses them in the query of images to retrieve from the end-user. Euclidean distance is used to compute the similarity measurement between an input of query image feature vectors and the dataset's feature. Experimental results show that the proposed HoG-LTP technique with CNN classifier, achieves higher performance than existing methods. For future works, various machine learning algorithms will be applied for optimized feature extraction in medical images to reduce the redundant data from the given image datasets.

6. REFERENCES

[1] M. Owais, M. Arsalan, J. Choi, K. R. Park, "Effective diagnosis and treatment through content-based medical image retrieval (Cbmir) by using artificial intelligence", Journal of Clinical Medicine, Vol. 8, No. 4, 2019, p. 462.

[2] M. Natarajan, S. Sathiamoorthy, "Heterogeneous Medical Image Retrieval using Multi-Trend Structure Descriptor and Fuzzy SVM Classifier", International Journal of Recent Technology and Engineering, Vol. 8, No. 3, 2019, pp. 3959-3963.

[3] A. Dureja, P. Pahwa, "Medical image retrieval for detecting pneumonia using binary classification with deep convolutional neural networks", Journal of Information and Optimization Sciences, Vol. 41, No. 6, 2020, pp. 1419-1431.

[4] D. B. Renita, C. S. Christopher, "Novel real time content-based medical image retrieval scheme with GWO-SVM", Multimedia Tools and Applications, Vol. 79, No. 23, 2020, pp. 1-17.

[5] M. Garg, G. Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants", Neural Computing and Applications, Vol. 33, No. 4, 2021, pp. 1311-1328.

[6] N. F. Haq, M. Moradi, Z. J. Wang, "A deep community-based approach for large scale content based X-ray image retrieval", Medical Image Analysis, Vol. 68, 2021, p. 101847.

[7] S. Fadaei, A. Rashno, "Content-based Image Retrieval Speedup Based on Optimized Combination of Wavelet and Zernike Features Using Particle Swarm Optimization Algorithm", International Journal of Engineering, Vol. 33, No. 5, 2020, pp. 1000-1009.

[8] M. Kashif, G. Raja, F. Shaukat, "An Efficient Content-Based Image Retrieval System for the Diagnosis of

Lung Diseases", Journal of Digital Imaging, Vol. 33, No. 4, 2020, pp. 971-987.

[9] A. Khatami, M. Babaie, H. R. Tizhoosh, A. Khosravi, T. Nguyen, S. Nahavandi, "A sequential search-space shrinking using CNN transfer learning and a Radon projection pool for medical image retrieval", Expert Systems with Applications, Vol. 100, 2018, pp. 224-233.

[10] L. Tsochatzidis, K. Zagoris, N. Arikidis, A. Karahaliou, L. Costaridou, I. Pratikakis, "Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach", Pattern Recognition, Vol. 71, 2017, pp. 106-117.

[11] Z. N. K. Swati, Q. Zhao, M. Kabir, F. Ali, Z. Ali, S. Ahmed, J. Lu, "Content-Based Brain Tumor Retrieval for MR Images Using Transfer Learning", IEEE Access, Vol. 7, 2019, pp. 17809-7822.

[12] S. Veerashetty, N. B. Patil, "Manhattan distance-based histogram of oriented gradients for content-based medical image retrieval", International Journal of Computers and Applications, Vol. 43, No. 9, 2021, pp. 924-930.

[13] A. Ahmed, "Implementing Relevance Feedback for Content-Based Medical Image Retrieval", IEEE Access, Vol. 8, 2020, pp. 79969-79976.

[14] R. S. Bressan, P. H. Bugatti, P. T. M. Saito, "Breast cancer diagnosis through active learning in content-based image retrieval", Neurocomputing, Vol. 357, 2019, pp. 1-10.

[15] A. Khatami, M. Babaie, A. Khosravi, H. R. Tizhoosh, S. Nahavandi, "Parallel deep solutions for image retrieval from imbalanced medical imaging archives", Applied Soft Computing, Vol. 63, 2018, pp. 197-205.

[16] E. Akanksha, P. R. K. Rao, "A Feature Extraction Approach for Multi-Object Detection Using HoG and LTP", International Journal of Intelligent Engineering and Systems, Vol. 14, No. 5, 2021, pp. 259-268.

Detection of CSR from Blue Wave Fundus Autofluorescence Images using Deep Neural Network Based on Transfer Learning

Original Scientific Paper

Bino Nelson
Faculty of Engineering and Technology, University of Kerala,
Thiruvananthapuram, Kerala, India.
binonelson@gmail.com

Haris Pandiyapallil Abdul Khadir
Department of Electronics and Communication Engineering, College of Engineering Trivandrum,
Thiruvananthapuram, Kerala, India.
haris786pa@gmail.com

Sheeba Odattil
Department of Electronics and Communication Engineering, T. K. M. College of Engineering,
Kollam, Kerala, India.
shb.odattil@gmail.com

Abstract *– Fluid clot below the retinal surface is the root cause of Central Serous Retinopathy (CSR), often referred to as Central Serous Chorioretinopathy (CSC). Delicate tissues that absorb sunlight and enable the brain to recognize images make up the retina. This important organ is vulnerable to damage, which could result in blindness and vision loss for the affected person. Therefore, complete visual loss may be reversed and, in some circumstances, may return to normal with early diagnosis discovery. Therefore, timely and precise CSR detection prevents serious damage to the macula and serves as a foundation for the detection of other retinal disorders. Although CSR has been detected using Blue Wave Fundus Autofluorescence (BWFA) images, developing an accurate and efficient computational system is still difficult. This paper focuses on the use of trained Convolutional Neural Networks (CNN) to implement a framework for accurate and automatic CSR recognition from BWFA images. Transfer Learning has been used in conjunction with pre-trained network architectures (VGG19) for classification. Statistical parameter evaluation has been used to investigate the effectiveness of DCNN. For VGG19, the statistic parameters evaluation revealed a classification accuracy of 97.30%, a precision of 99.56%, an F1 score of 97.25%, and a recall of 95.04% when using a BWFA image dataset collected from a local eye hospital in Cochin, Kerala, India. Identification of CSR from BWFA images is not done before. This paper illustrates how the proposed framework might be applied in clinical situations to assist physicians and clinicians in the identification of retinal diseases.*

Keywords*: Blue Wave Fundus Autofluorescence, Central Serous Retinopathy, Deep Learning, Retinal Pathology, Macular Disease*

1. INTRODUCTION

The retina is located near the optic nerve at the back of the eyeball and is made up of thin layers of ocular tissues [1]. The major job of the retina is to collect light from the focal length of the lens and convert it into message signals that the brain can recognize visually. As a result, the retina is regarded as an important component of the eye and is crucial for identifying diverse nearby objects. Any abnormal or impairment to the retinal layer might cause the patient to lose their vision or have visual impairments. Blindness and visual loss may be caused by Central Serous Retinopathy (CSR), one of the diseases which impacted more people worldwide. The primary reason for CSR is the clotting of the liquid medium on the Retinal Surface Area (RSA), which adversely impairs people's eyesight [2]. To prevent visual loss, early accurate CSR detection can help with treatment choices and diagnostic methods. Several imaging modalities, including Fundus photography, Optical Coherence Tomography Angiography (OCTA), and Blue Wave Fundus Autofluorescence (BWFA), are employed for this CSR detection purpose. Among all BWFA-based methods for detecting a variety of retinal diseases, it is considered the most suitable and cutting-edge imag-

ing tool. Although the BWFA imaging method is similar to ultrasonic imaging, it has a little difference in detection. Unlike ultrasound imaging, which records an ultrasound image using sound waves, BWFA imaging uses blue light rays [3]. Employing a set of endoscopes and catheters, high-definition pictures of the RSA were generated via BWFA. In BWFA images of healthy patients and those with CSR the size of the fovea, the choroid, the Inner Limiting Membrane (ILM), and serous retinal detachment could be detected.

Thanks to the recent advancement of quantitative BWFA, Artificial Intelligence (AI) categorization of eye disorders and computer-aided illness diagnosis are now feasible. Vascular occlusions, Sickle Cell Retinopathy, Age-related Macular Degeneration (AMD), and CSR have all been investigated using quantitative BWFA analysis. Recent research has shown that supervised Machine Learning (ML) is effective for categorizing a range of activities to distinguish between normal and CSR images. Deep Learning (DL) provides a simple means of promoting the clinical application of AI classification of BWFA images. The Convolutional Neural Network (CNN) method, which was designed to mimic how the human brain interprets visual data, is normally referred to as DL [4].

To extract and analyze straightforward characteristics in the early layers and information useful in the layers of the network of the feed-forward processing of visual signals, CNNs employ millions of artificial neurons, often known as parameters. Until a CNN can be trained for a specific classification task, the network parameters need to be adjusted using millions of images. The lack of currently available images for the comparatively recent imaging modality BWFA, however, presents a challenge for the practical use of DL. DL has been proven to be implemented using a Transfer Learning (TL) approach to get around the data size restriction. TL is a technique for optimizing the weights of a pre-trained CNN by adopting some of its weights and correctly retraining parts of its layers (i.e., AI classification of retinal images). TL has been investigated in fundus photography for artery-vein segmentation, glaucoma diagnosis, and diabetic macular thinning assessment. TL has recently been investigated in BWFA for the detection of AMD, diabetic macular edema, and choroidal neovascularization.

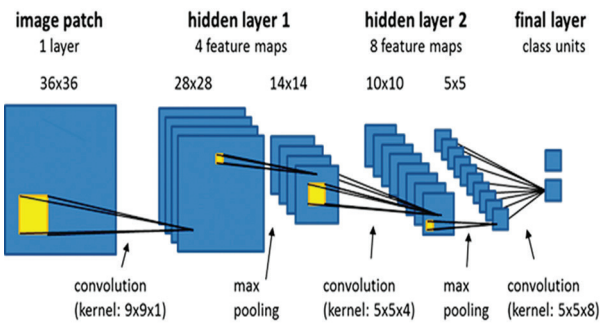


Fig. 1. Convolutional Neural Network for Image Classification

Because the weights in each layer can be changed, TL can theoretically entail a single layer or numerous layers. In a 16-layer CNN, for instance, the precise number of layers needed for retraining may differ based on the dataset at hand and the target job. Since there aren't as many large publicly accessible BWFA datasets, images verified by ophthalmologists have been collected from a private hospital [5]. In this study, we show how BWFA can be used for DL-based automated classification for the first time. We want to train a small dataset using TL to produce a highly accurate CSR classification.

2. RELATED WORKS

To evaluate CSR, Dursun et al. [6] presented the Deep Capillary Plexus (DCP-VD) and Superficial Capillary Plexus (SCP-VD) parameters. The accuracy of detection when utilizing a Support Vector Machine (SVM) is 87%. Pawan et al. [7] introduced a modified version of Capsule Network for the detection and segmentation of CSR. This algorithm is lightweight and reduces the computational overhead. While using 1792 samples this technique provides a Dice Coefficient of 94.04 and an accuracy of 91.67%. Sulzbacher et al [8] categorized neo-vascular and non-neo-vascular CSR. They utilized ML algorithms for the diagnosis of CSR. By using logistic regression, the accuracy obtained is 92.63%. Signal variations at the Chorio Capillaries (CC) level in patients with various phases were examined by Cakir et al. [9]. According to this approach, CSR patients were classified into four categories: acute, persistently atrophic, non-resolving, and inactive. CNN was used for the detection of CSR and the accuracy of detection is 92.13%. Aggarwal et al. [10] conducted an analysis to identify the differences between acute Vogt-Koyanagi-Harada disease (VKH) and acute CSR in terms of image characteristics. Sonoda et al. [11] suggested employing enhanced depth imaging to identify the anatomical changes of the choroid in eyes with CSR. The Inner and Outer Choroid structural changes technique is employed. In the images, the inner and outer choroid's hypo-reflective and hyperreflective regions were measured separately. The results revealed that the outer choroid of the CSR eyes had a significantly larger hyporeflective area (446,549 + 121,214µm2) than the control eyes (235,680 + 97,352µm2, P<0.01).

The diagnosis in CSR, according to Kulikov et al. [12], is based on ML and AI. The values for identifying subretinal fluid were 0.61, 0.99, 0.99, and 0.76, respectively, according to the analysis of test sets. For the detection of anomalies in the retinal pigment epithelium (RPE), the specificity, sensitivity, F1-score, and precision were 0.95, 0.14, 0.24, and 0.94, respectively. For leaking point identification, the sensitivity, specificity, accuracy, and F1-score were 0.06, 1.0, 1.0, and 0.12, respectively. The proposed method concluded that ML had great promise for identifying structural anomalies linked to acute CSR. Narendra et al [25] reported a CNN for the automatic segmentation of Sub-Retinal Fluid (SRF) to

identify CSR. The suggested technique, provided Dice, Recall, and Precision of 0.91, 0.83%, and 0.93%, respectively. An automatic SRF segmentation method is applied in the suggested model. Pre-processing and fluid segmentation are the two stages of the SRF segmentation method. Yoon et al. [13] proposed developing a DL system model to diagnose CSR and differentiate between chronic and acute CSR. This model's accuracy, sensitivity, and specificity for diagnosing CSR were 93.8%, 90.0%, and 99.1%, respectively. Jaisankar et al. [14] identified damaged choroidal areas and associated retinal alterations in acute and recurring CSR. Khalid et al. [15] conducted the first-ever analysis of a decision support system. The proposed system's automated disease identification is based on a multilayered SVM classifier that was trained on 40 labelled scans. With sensitivity, accuracy, and specificity ratings of 98 %, 96.92 %, and 95.86 %, respectively, the suggested approach was successfully diagnosed in 2817/2819 individuals.

Evaluation of the morphological variations in the choroidal vasculature in acute and chronic CSR was suggested by Lee et al. [16]. Different choroidal vascular dilatation patterns between acute and chronic CSR were discovered using the suggested methodology. These discoveries could aid in our understanding of the pathophysiology of CSR. In eyes with acute CSR that had not yet received treatment, Yang et al. [17] recommended using CNN to evaluate the three-dimensional choroidal vascularity index (CVI). This technique makes use of 3D-CVI, which volumetrically evaluates the choroidal vasculature, and is a valuable imaging marker of choroidal disorders, according to the suggested method, which may help better understand the pathophysiology of CSR.

In the evaluation of the eyes, Altnel et al. [18] divided them into three groups: failure, partial remission, and complete remission. The thickness of the Ellipsoid Zone (EZ), Outer Nuclear Layer (ONL), and retinal pigment epithelium (RPE) were all assessed using this procedure. In the group of patients with complete remission, the baseline EZ and RPE were discovered to be intact in 71.4% and 64.3% of the eyes, respectively. However, in the failure group, these rates were 25% and 16.7%, respectively. Tang et al. [19] proposed evaluating quantitative indicators associated with treatment response and the beginning of macular neovascularization (MNV) in CSR using an AI-based methodology. In the experiment, eplerenone produced an outstanding reaction in 40/78 (51%) of the eyes, compared to PDT in 38/78 (49%) of the eyes.

3. METHODOLOGY

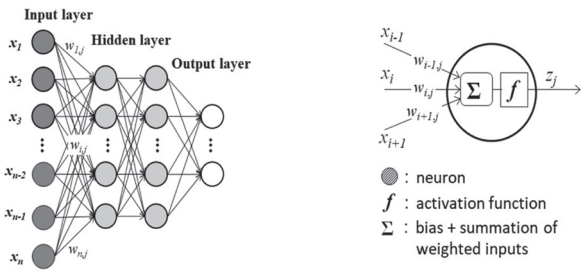
The images collected from a local eye hospital in Cochin, Kerala, India (BWFA dataset) are used to train the classifier (VGG19) used in this research [20]. TL is incorporated to enhance the classifier's performance. In the event of an error, it spreads backwards through the system and modifies the weight of each node.

Iterations of this procedure are repeated over and over, and the distance is continually assessed and improved. Convolutional layers extract the features, and the CNN is trained using these features. A CNN is a DL-based Artificial Neural Network (ANN) architecture. We developed a fairly straightforward method. The convolution layer and max-pooling layer of the CNN, which together form the 2D spatial image (pooling), are mapped to the values in Equations (1) and (2). Every neuron receives an output value z from a function called activation that filters input data x with weight coefficients w. If a layer has an n number of neurons, then a specific neuron named j in the following layer gets an n input value. Bias b_j integrates and adds these coefficient-weighted inputs. $w_{i,j}$ is the weight applied to the input of i^{th} neuron to produce an output at j^{th} neuron. A neuron's activation function f produces the signal z_j. Fig. 2 shows the structure of an artificial neural network.

$$y_j = b_j + \sum_{i=1}^{n} w_{i,j} x_i \tag{1}$$

$$z_j = f(y_j) \tag{2}$$

The input variable x is assumed to have L x L 2D source pixels and to be processed by convolution with a table of weights using H x H window. After extracting values of the same size from the source image, the convolution combines the weighting factors of the window H x H by the filter values of the pixels for the source. By shifting the window, then the process of filtering is applied once more to the source image. It should be mentioned that a zero-padding method was included to make sure that both the input and output are of the same size.



(a) (b)

Fig. 2. Structure of ANN (a) Data Flow in DL (b) Inner Structure of Neuron

The convolution operation is displayed in Fig. 3. As $x_{i,j}$ is then multiplied by $w_{k,l,n}$, a grid at (*i,j*) on the filters is independently shifted. This equation is written as follows, incorporating the bias bn in the expression and the convolution form:

$$y_{i,j,n} = \sum_{k=0}^{H-1} \sum_{l=0}^{H-1} w_{k,l,n} x_{i+k,j+1} + b_n \tag{3}$$

$$z_{i,j,n} = f(y_{i,j,n}) \tag{4}$$

$y_{i,j,n}$ represents the convolution output of the image with a weighing filter and $z_{i,j,n}$ is the final output after applying a rectified linear unit (ReLU) as the activation

function (f), which could select positive input values as a result of an improvement in matrix conversion.

$$f(y_{i,j,n}) = \max(y_{i,j,n}, 0) \quad (5)$$

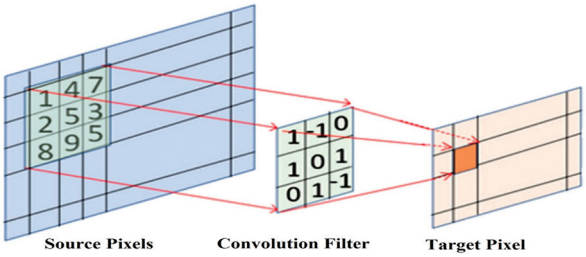


Fig. 3. Convolution Layer with Filter

Equation (6) defines pooling, which is generally used to produce 2D data, but does not include weighted coefficients or activation functions. p, q is the domain's horizontal and vertical parts, and Up, q is the square unit domain with the dimensions R x R.

$$z_{p,q} = \max(x_{i,j}), (i,j) \epsilon U_{p,q} \quad (6)$$

where $z_{p,q}$ represents the max-pooled output of the processed image sub-region $x_{i,j}$, of the whole image $U_{p,q}$. The same purpose as Equations (1) and (2) is served by these processes, which produce relatively small 2D picture data that are subsequently transmitted to a fully linked layer. These 2D picture data are transformed into several number of 1D digital files that offer details about the classification of the original image. The Soft-Max function in Equation (7) transforms outcomes into probabilities in the output layer. It is a normalized exponential function. It evaluates binary categorization.

$$f(y_j) = \frac{e^{y_j}}{\sum_{r=1}^{m} e^{y_r}} \quad (7)$$

f_{yj} represents the softmax function that converts the output score of each neuron y_j to probabilities of being in a class.

This work considers the existence of two distinct domains. Large datasets used to train a model in a specific (source) domain typically require a lot of time before an accurate prediction can be made. When the model is used in a different (target) domain that is unrelated to the source domain, training takes a lot of time. TL is one of the approaches for improving the effectiveness of the prediction in the target domain (e.g., decreasing run time). For the target domain, the TL can repurpose accumulated common knowledge from the source domain. Recently, the CNN combined with a transfer-learning strategy (CNN-TL) is a successful image classification technique. In this work, CNN-TL was used to transform time series data to picture data because our model should be able to forecast time series. CNN was initially carried in the target domain (Fig. 4a). Second, portions of the CNN's hidden layers were modified before being reapplied to the particular domain. Finally, using datasets from the target domain, "the fully connected layer 1" and "the fully connected layer 2" in the deep layers were re-trained (Fig. 4b).

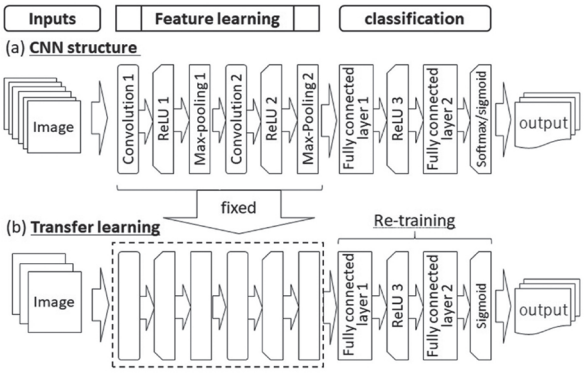
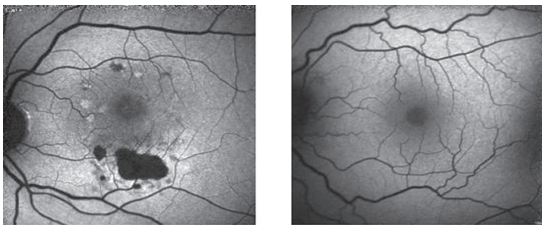


Fig. 4. DL Architectures (a) Typical CNN and (b) CNN with TL

The CNN image must match a CSR value for CNN training and prediction operations to be effective. The dataset's images having size 256x256 were chosen because there weren't many variables. As input data, a square image is utilized. A 1x1 of the data output was linked to the data from the input data at the expected time. We assumed that the quantity of the predicted position in the following time step is produced by the CSR data in a square picture. Take note that the predicted point indicates whether or not CSR is present.

4. RESULTS AND DISCUSSION

The experimental outcomes of the suggested CSR detection algorithm are shown in this section. Utilizing the BWFA dataset, the proposed TL model's implementation is verified. BWFA images of individuals with CSR disease and BWFA images of healthy people have both undergone classifications. The entire dataset of images was divided into two categories by the algorithm, such as training and testing images. The dataset includes 1600 images from the normal (healthy) group and 1608 images from the CSR category. Sample images from both categories are shown in Fig. 5.



(a) (b)

Fig. 5. Sample Images from BWFA Dataset (a) CSR (b) Healthy.

The research addresses the class imbalance problem. The distribution of samples for the training classes is uniform. 70% of the images from each class were taken for training the CNN. The remaining 30% of the images from each class were used for testing the CNN.

Fig. 6 shows the distribution of BWFA images into different groups, which are used for training and testing the classifier.

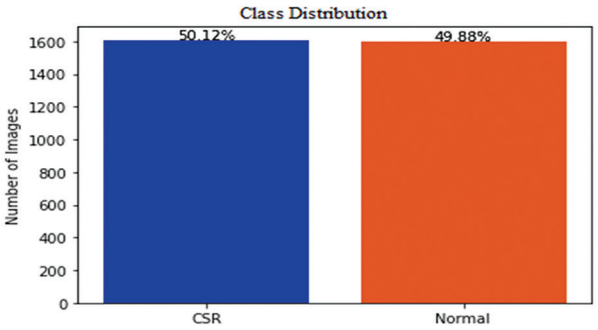


Fig. 6. BWFA Dataset Class Distribution

The input to this network has a fixed-size (256 * 256) RGB image. Hence the matrix had the structure of (256, 256, 3). The mean RGB value of each pixel, calculated throughout the whole training set, was the only preprocessing that was carried out. They made use of kernels with a stride size of 3 × 3 pixels. They were able to completely hide the image because of this. Spatial padding is being used to preserve the spatial resolution of the image. Max pooling was done with stride 2 above a 2x2 pixel window. ReLu was then used to add non-linearity, which improved computational efficiency and helped the model better categorise data than earlier models that employed tanh or sigmoid functions.

VGG19 is a CNN having 19 hidden layers in total. It is possible to load a network that has already been trained using more than a million images from the ImageNet database. The pre-trained network can categorize images into more than 1000 different object categories and has amassed rich visual information for a range of images. After loading the pre-trained model and analyzing the feature representations, it is noticed that there are 20,057,153 parameters in total. Of these parameters, 32,769 parameters are trainable. These parameters are trained using the features obtained from the BWFA dataset. The architecture of the proposed modified VGG19 model is depicted in Fig. 7.

The 32,769 parameters are trained using the proposed model by considering binary cross entropy as the loss function. The usage of loss functions is to calculate the quantity that a model should be used to minimize during training. It computes the cross-entropy loss between true labels and predicted labels. Furthermore, when comparing our results to existing work, we used the same settings by using cross-validation and fixed partitioning techniques. The Python language was used to create the pre-trained models based on TL, which were then run on the Google Colab platform [21]. The rate of learning was 1 x 10-4, the minimum batch size was 32 and the optimum epoch count is 50. A reduction in the rate of learning increases the time for the training network while raising it causes training to become locked at an inadequate result [22].

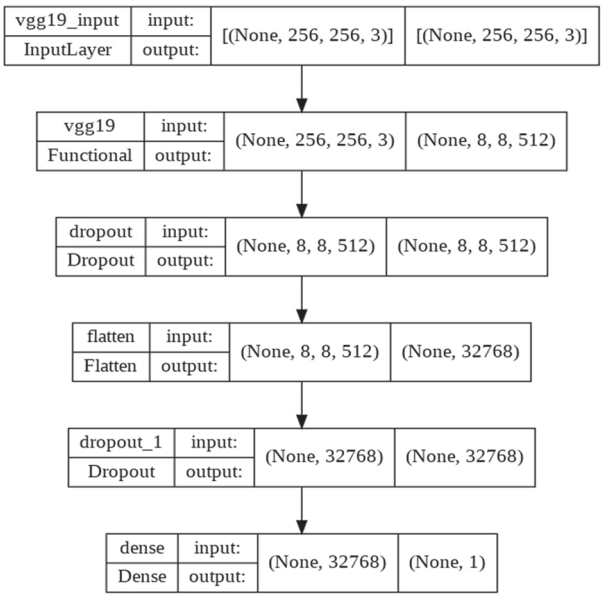


Fig. 7. Proposed VGG19 Model Architecture.

Many performance indicators were considered to validate the effectiveness and efficiency of the pre-trained VGG19 model integrating TL. Precision, F1 Score, recall, ROC AUC, specificity, and accuracy are the measures that are most frequently utilized. The words True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP) are used to describe classification accuracy. The count of CSR images labelled as CSR is TP and the count of healthy images labelled as healthy is TN. The count of healthy images labelled as CSR is FN and the count of CSR images labelled as healthy is FP [23]. The performance parameters are expressed mathematically as,

$$Precision = \frac{TP}{TP+FP} \tag{8}$$

$$Recall = \frac{TP}{TP+FN} \tag{9}$$

$$F1 = 2 * \frac{Precision*Recall}{Precision+Recall} \tag{10}$$

$$Accuracy = \frac{TN+TP}{TP+FP+TN+FN} \tag{11}$$

$F1$ score is the harmonic mean of precision and recall as indicated in equation 10. Fig. 8 shows the visualization charts for the accuracy of BWFA image classification. The number of iterations (epochs) used in the experiment was 50. The values of performance parameters are stable and high after the 40th epoch. This is due to the classification problem being solved using a TL approach. There is a chance that an error will occur during training, which is referred to as a loss. The proposed VGG19 model with TL for CSR detection has a minimum loss of 0.38, which is quite low as illustrated in Fig. 9. The entire set of test images must be used in the evaluation. The proposed VGG19 with TL model for CSR detection has an average accuracy of 97.30%. Precision and recall have average values of 99.56% and 95.04%, respectively. The proposed model provides an

$F1$ score of 97.25% and a ROC AUC of 97.31%. The maximum number of BWFA images used for training can be increased to increase the classification parameters.

Fig. 8. Variation in Accuracy with Epochs

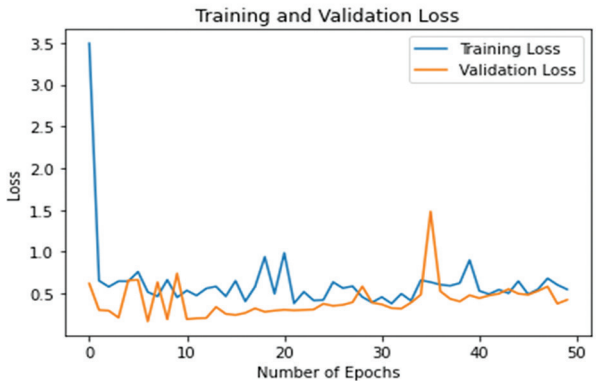


Fig. 9. Variation in Loss with Epochs

Experiments have been conducted on the developed model using a given BWFA dataset. The batch size selected for training the model is 32. An experiment is conducted in various steps by varying the number of epochs. The total number of epochs varied from 10 to 50, and the values of performance metrics were computed. 10 epochs provided lower values for the performance parameters, and 50 epochs provided higher values. There was no change in the performance parameters after 50 epochs. So, the experiments are concluded and the values obtained for various performance metrics are given in Table 1. Fig. 10 shows ROC graphs and AUC displays the performance. The overall AUC for this cross-validation research was 0.9731.

Table 1. Variation in Performance with Epochs

Performance Metrics	Epoch				
	10	20	30	40	50
Accuracy (%)	95.72	96.21	97.78	96.14	97.30
Precision (%)	94.73	96.09	96.56	98.32	99.56
Recall (%)	92.16	91.58	93.57	93.92	95.04
F1 Score (%)	92.37	94.62	94.87	96.56	97.25
Loss	0.72	0.56	0.54	0.45	0.38

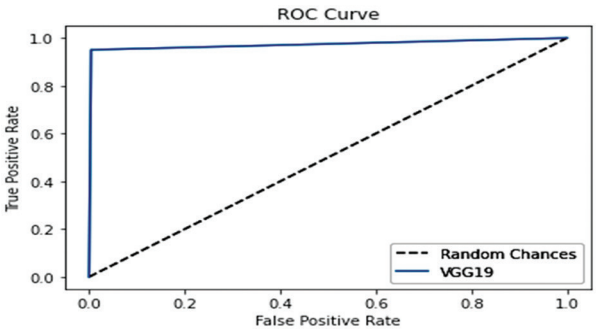


Fig. 10. ROC curves for the cross-validation of the proposed model

A detailed comparison of classification performance is required to assess the effectiveness of the proposed TL approach. On various datasets, the proposed transferred models' classification performance is evaluated. Table 2 compares the efficiency of existing models using the selected performance metrics.

Table 2. Comparison of CSR Detection Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Loss (%)
AlexNet	95.92	96.21	94.78	92.32	0.58
GoogleNet	88.73	89.09	86.56	90.17	0.71
ResNet101	92.16	91.58	90.57	91.53	0.63
VGG16	92.37	92.62	91.87	95.56	0.44
Inception v3	87.92	84.06	85.74	92.45	0.56
Inception ResNet v2	89.95	89.40	91.57	93.52	0.49
Proposed	97.30	99.56	95.04	97.25	0.38

When comparing classification accuracy, the proposed model achieves the best result of 97.30%, which is achieved through TL. In terms of accuracy, pre-trained models like AlexNet (95.92%), ResNet101 (92.16%), and VGG16 (92.37%) performed well. When compared to AlexNet's accuracy, the proposed classification model has an increase of 1.38%. When comparing classification precision, the proposed model achieves 99.56%, which is the highest among all other models. AlexNet's precision was 96.21%, ResNet101's precision was 91.58%, and VGG16's precision was 92.62%. When comparing the proposed classification model to AlexNet's precision, there is a 3.35% increase. The proposed model achieves a recall value of 95.05%. When compared to other models, this has the best recall value. AlexNet had a recall of 95.78%, Inception ResNet v2 had a recall of 91.57%, and VGG16 had a recall of 91.87%. The proposed classification model has a 0.27% increase in recall when compared to AlexNet's recall.

When the training loss is compared, the proposed method and AlexNet have a lower loss. The proposed model has a 0.38 training loss, while the AlexNet model has a 0.58 training loss. When comparing these two models, there is a difference of 0.2. The proposed model achieved the lowest loss and is the best classification

model for CSR detection. The significance of selected parameters demonstrates the value of TL in the reduction of over-fitting and improving classification accuracy. The proposed model and AlexNet were found to be efficient in identifying samples across appropriate datasets. Fig. 11 shows how the performance of the various models compares. In Fig. 12, it is possible to visualise the comparison of training loss.

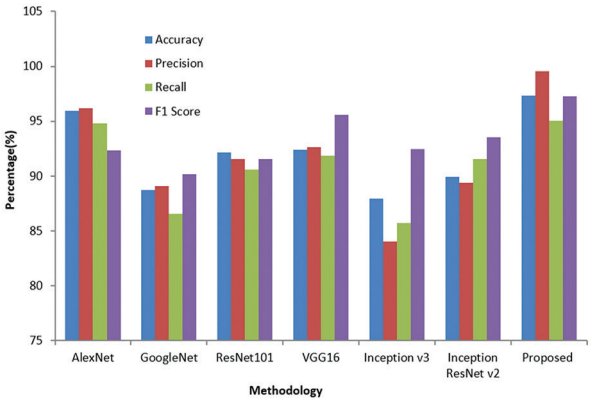


Fig. 11. Comparison of Classification Performance

When it comes to CNN-based models, the real motivation comes from the fact that they are used to solve extremely difficult problems using millions of labelled datasets. To create a dataset for DL models, you'll need years of data collection. There are various advantages to employing pre-trained VGG19 systems with TL for categorization. First and foremost, the coding system is automated. Secondly, noise filtering, ROI delineation, feature extraction, and selection are not required anymore. Thirdly, there are no biases, and the pre-trained CNN models' predictions are repeatable. Finally, in contrast to previous CNN-based studies, a ceiling level of accuracy is obtained. The computation time is reduced because we used the Colab platform's CPU and GPU as hardware resources.

5. CONCLUSION

In this work, we evaluated various pre-trained CNN techniques employing TL for BWFA image categorization. To get the highest recognition rate, pre-trained CNN methods have been successfully combined with TL. The suggested model surpassed every other model that was put to the test, scoring 97.30% accuracy, 99.56% precision, an F1 score of 97.25% and 95.04% recall for the BWFA datasets. This model uses TL to obtain the lowest training loss of 0.38. The results performed better on CSR classification from BWFA images than current classical and DL techniques. It performs better than current methods in eliminating the requirement for pre-processing procedures. Additionally, compared to existing DL-based work, the pre-trained AlexNet model produces better performance metrics. Future studies will concentrate on putting the models into platforms, reducing computational complexity, and investigating additional methods of fine-tuning.

6. REFERENCES:

[1] D. Lu et al. "Deep-learning based multiclass retinal fluid segmentation and detection in optical coherence tomography images using a fully convolutional neural network", Medical Image Analysis, Vol. 54, 2019, pp. 100-110.

[2] J. De Moura, J. Novo, S. Penas, M. Ortega, J. Silva, A. M. Mendonca, "Automatic characterization of the serous retinal detachment associated with the subretinal fluid presence in optical coherence tomography images", Procedia Computer Science, Vol. 126, 2018, pp. 244-253.

[3] H. S. Sandhu et al. "Automated diagnosis of diabetic retinopathy using clinical biomarkers, optical coherence tomography, and optical coherence tomography angiography", American Journal of Ophthalmology, Vol. 216, 2020, pp. 201-206.

[4] A. Miere et al. "Deep learning-based classification of retinal atrophy using fundus autofluorescence imaging", Computers in Biology and Medicine, Vol. 130, 2021, p. 104198.

[5] H. Ullah, M. Faran, Z. Batool, A. Nazir, G. Gilanie, N. Amin, "Diagnosis of Ocular Diseases Using Optical Coherence Tomography (OCT) at λ= 840 nm", Lasers in Engineering (Old City Publishing), 2022, pp. 53.

[6] M. E. Dursun et al. "Evaluation of parafoveal vascular density using optical coherence tomography angiography in patients with central serous chorioretinopathy", Lasers in Medical Science, Vol. 37, No. 2, 2022, pp. 1147-1154.

[7] S. J. Pawan et al. "Capsule Network–based architectures for the segmentation of sub-retinal serous fluid in optical coherence tomography images of central serous chorioretinopathy", Medical & Biological Engineering & Computing, Vol. 59, No. 6, 2021, pp. 1245-1259.

[8] F. Sulzbacher, C. Schütze, M. Burgmüller, P. V. Vécsei-Marlovits, B. Weingessel, "Clinical evaluation of neovascular and non-neovascular chronic central serous chorioretinopathy (CSC) diagnosed by swept-source optical coherence tomography angiography (SS OCTA)", Graefe's Archive for Clinical and Experimental Ophthalmology, Vol. 257, No. 8, 2019, pp. 1581-1590.

[9] E. Borrelli et al. "Optical coherence tomography parameters as predictors of treatment response to eplerenone in central serous chorioretinopathy", Journal of Clinical Medicine, Vol. 8, No. 9, 2019, p. 1271.

[10] K. Aggarwal et al. "Distinguishing features of acute Vogt-Koyanagi-Harada disease and acute central serous chorioretinopathy on optical coherence tomography angiography and en face optical coherence tomography imaging", Journal of Ophthalmic inflammation and infection, Vol. 7, No. 1, 2017, pp. 1-9.

[11] S. Sonoda et al. "Structural changes of inner and outer choroid in central serous chorioretinopathy determined by optical coherence tomography", PLoS One, Vol. 11, No. 6, 2016, p. e0157190.

[12] A. N. Kulikov, E. Y. Malahova, D. S. Maltsev, "Artificial intelligence and machine learning for optical coherence tomography-based diagnosis in central serous chorioretinopathy", Ophthalmology Journal, Vol. 12, No. 1, 2019, pp. 13-20.

[13] J. Yoon et al. "Optical coherence tomography-based deep-learning model for detecting central serous chorioretinopathy", Scientific Reports, Vol. 10, No. 1, 2020, pp. 1-9.

[14] D. Jaisankar, M. Kumar, P. Rishi, S. Singh, R. Raman, "Correlation of retinal changes with choroidal changes in acute and recurrent central serous chorioretinopathy assessed by swept-source optical coherence tomography", Therapeutic Advances in Ophthalmology, Vol. 12, 2020, p. 2515841419899823.

[15] S. Khalid, M. U. Akram, T. Hassan, A. Nasim, A. Jameel, "Fully automated robust system to detect retinal edema, central serous chorioretinopathy, and age-related macular degeneration from optical coherence tomography images", BioMed Research International, 2017.

[16] W. J. Lee, J. W. Lee, S. H. Park, B. R. Lee, "En-face choroidal vascular feature imaging in acute and chronic central serous chorioretinopathy using swept source optical coherence tomography", British Journal of Ophthalmology, Vol. 101, No. 5, 2017, pp. 580-586.

[17] J. Yang, E. Wang, M. Yuan, Y. Chen, "Three-dimensional choroidal vascularity index in acute central serous chorioretinopathy using swept-source optical coherence tomography", Graefe's Archive for Clinical and Experimental Ophthalmology, Vol. 258, No. 2, 2020, pp. 241-247.

[18] M. G. Altınel, B. Acikalin, H. Gunes, G. Demir, "Optical coherence tomography parameters as predictors of treatment response to a 577-nm subthreshold micropulse laser in chronic central serous chorioretinopathy", Lasers in Medical Science, Vol. 36, No. 7, 2021, pp. 1505-1514.

[19] A. A. Tang, A. L. Lin, "Purtscher-like retinopathy in adult-onset Still's disease, complicated by treatment-related central serous chorioretinopathy", American Journal of Ophthalmology Case Reports, Vol. 18, 2020, p. 100631.

[20] P. Gholami, P. Roy, M. K. Parthasarathy, V. Lakshminarayanan, "OCTID: Optical coherence tomography image database", Computers & Electrical Engineering, Vol. 81, 2020, p. 106532.

[21] T. Sekiryu, "Choroidal imaging using optical coherence tomography: techniques and interpretations", Japanese Journal of Ophthalmology, 2022, pp. 1-14.

[22] K. Gao et al. "Double-branched and area-constraint fully convolutional networks for automated serous retinal detachment segmentation in SD-OCT images", Computer methods and programs in biomedicine, Vol. 176, 2019, pp. 69-80.

[23] S. S. Mishra, B. Mandal, N. B. Puhan, "Macular Net: Towards Fully Automated Attention-Based Deep CNN for Macular Disease Classification", SN Computer Science, Vol. 3, No. 2, 2022, pp. 1-16.

[24] M. Treder, J. L. Lauermann, N. Eter, "Deep learning-based detection and classification of geographic atrophy using a deep convolutional neural network classifier", Graefe's Archive for Clinical and Experimental Ophthalmology, Vol. 256, No. 11, 2018, pp. 2053-2060.

[25] R. Narendra, G. N. Girish, A. R. Kothari, J. Rajan, "Deep Learning Based Sub-Retinal Fluid Segmentation in Central Serous Chorioretinopathy Optical Coherence Tomography Scans", Proceedings of the 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 978–981.

Reordering of Source Side for a Factored English to Manipuri SMT System

Original Scientific Paper

Indika Maibam
Department of Computer Science
Indira Gandhi National Tribal University,
Kangpokpi, Imphal, Manipur, India
maibam.indika@igntu.ac.in

Bipul Syam Purkayastha
Department of Computer Science,
Assam University, Silchar, Assam, India
bipul_sh@hotmail.com

Abstract *– Similar languages with massive parallel corpora are readily implemented by large-scale systems using either Statistical Machine Translation (SMT) or Neural Machine Translation (NMT). Translations involving low-resource language pairs with linguistic divergence have always been a challenge. We consider one such pair, English-Manipuri, which shows linguistic divergence and belongs to the low resource category. For such language pairs, SMT gets better acclamation than NMT. However, SMT's more prominent phrase-based model uses groupings of surface word forms treated as phrases for translation. Therefore, without any linguistic knowledge, it fails to learn a proper mapping between the source and target language symbols. Our model adopts a factored model of SMT (FSMT3*) with a part-of-speech (POS) tag as a factor to incorporate linguistic information about the languages followed by hand-coded reordering. The reordering of source sentences makes them similar to the target language allowing better mapping between source and target symbols. The reordering also converts long-distance reordering problems to monotone reordering that SMT models can better handle, thereby reducing the load during decoding time. Additionally, we discover that adding a POS feature data enhances the system's precision. Experimental results using automatic evaluation metrics show that our model improved over phrase-based and other factored models using the lexicalised Moses reordering options. Our FSMT3* model shows an increase in the automatic scores of translation result over the factored model with lexicalised phrase reordering (FSMT2) by an amount of 11.05% (Bilingual Evaluation Understudy), 5.46% (F1), 9.35% (Precision), and 2.56% (Recall), respectively.*

Keywords*: Factored SMT, Reordering, Factoring, English, Manipuri, Automatic Evaluation*

1. INTRODUCTION

Machine Translation (MT) is defined as a "loop consisting of three steps in which, i) a source constituent is detected, ii) required information including syntactic, semantic and other types of information related to the constituent is collected, and iii) finally, it is transferred to a target form which is the end of the translation process for that constituent" [1]. By description, implementing an MT system seems straightforward and uncomplicated. Still, given the variety of languages spoken worldwide, each of which belongs to a separate family and has its unique linguistic structure, MT is not a simple procedure. The major issue is that these difficulties differ depending on the language combination under examination. Some of the challenges in our work are:

- Linguistic differences and their complexity
- Low resource category
- Unavailability of natural language processing (NLP) tools for Manipuri

MT is a very challenging task. The diversity of languages with linguistic differences between them, along with the inherently ambiguous nature, further amplifies the challenges. The language pair English-Manipuri is one such. We highlight a few linguistic differences between English and Manipuri. English has rudimental morphology with Subject-Verb-Object (S-V-O) sentences and non-tonal. The derivation is the most common process of word formation in English. For example, "*un+happy*", "*pre+judge*". In comparison, Manipuri has prolific morphology and agglutinating with Subject-Object-Verb (S-O-V) structure. Manipuri shows a variance from other Tibeto-Burman categories of languages in that it gives prominence to tense rather than mood.

Manipuri also shows tonal contrast, with two levels - high falling and level. For example, The word "*tummi*" can mean "*sleeping*" (high falling tone) or "*pointy*" (level tone). Word formation in Manipuri uses a large number of suffixes with fewer affixes and primarily uses the process of compounding. For example, *lan (war) + mee (person)* - "*lanmee*" (*soldier*). Manipuri has two scripts - Bengali and Meitei-Mayek. We are using the Bengali script for our work.

What is more challenging with our work is that Manipuri is a low-resource language. It is known that Neural Machine Translation (NMT) is data-hungry, and Statistical Machine Translation (SMT) is a better option at low resource conditions [2]. The sufficient training data size for NMT is in the order of millions [3] compared to few thousand for SMT. Above this, NLP tools, such as morphological analysers and part-of-speech (POS) taggers, are unavailable for the Manipuri language. This unavailability greatly restricts the researchers from implementing and trying out different possibilities in their area of research. The two techniques, which are pretty conventional in the area of MT, are SMT and NMT - with NMT getting more acclamation than SMT. Research on translations from more affluent to poorer morphology and vice versa is rarely focussed. The study [4] reports that translating from poorer to richer morphology is more complex and challenging than vice versa. However, for translations that involve two languages which are morphologically and structurally variant, determining which technique is better is still a question of doubt. Our work uses the SMT technique to develop a translation system for English-Manipuri. SMT makes use of parallel corpora and learning algorithms to train a model. Based on the translation model's training process, many SMT approach models are available: baseline model, phrase-based model, factored model and hierarchical phrase-based. Traditional or baseline models use word-level mapping, which produces low translations. An improvement to it is the phrasal one, the most routine SMT. The phrase-based SMT (PBSMT) uses a grouping of surface forms of words treated as phrases. On the other hand, a factored model, an extension of PBSMT, uses surface word forms with additional factors such as lemma, POS tags, morphological information, case, and genders. In contrast, the hierarchical model requires grammar which consists of Synchronous Context Free Grammar rules. In SMT, language models are used to establish the target word order. They are, however, limited by the sparsity of the data caused by larger n-grams. Therefore, a lexicalised reordering model subjects reordering to the PBSMT phrases. One of the most challenging issues in SMT is reordering; it manifests differently depending on the language combinations. Language pairings with far-off syntactic structures, like English and Manipuri, experience long-distance reordering issues that the lexicalised reordering models cannot resolve.

In our paper, we implement a factored model for English-Manipuri along with pre-processing, post-processing, and reordering modules. The following sections give a detailed explanation of the architecture. Unfortunately, unlike high-resource languages such as English with POS taggers available in the open domain, no such tool is available for Manipuri. It is also difficult to find annotated corpora for this language. So, owing to financial reasons, we have developed only a small set of annotated corpora of the entertainment domain to experiment with the system. Additional corpora with more factors, if available, can be incorporated to improve the result. This work is a quest to improve translation quality.

The paper implements a factored model with POS information-based reordering to enhance translation between linguistically disparate language pairs. The goal of the current task is to rearrange the source chunks so that the alignments of the source and destination pieces are more monotonous. The reordering has been done by manually rearranging the source text at the chunk level to replicate the ordering of the target language. The process mitigates the problem of long-range reordering to only short-range, intra-chunk reordering that the lexicalised reordering easily handles. Furthermore, Manipuri's low resources and agglutinating nature produces untranslated words, which are transliterated.

2. FACTORED SMT

The most dominant SMT approach, the PBSMT model, has an extension model called the factored SMT model. The phrase model uses small text chunks and phrases without linguistic information during translation. The PBSMT model uses the noisy channel model. After applying Bayes theorem [5], the translation probability for translating a source sentence (S) to a target sentence (T) is

$$P(T|S)=P(S|T)P(T)=argmax_m\, P(S|T)P(T) \quad (1)$$

In Equation 1, the component P(S|T) represents the translation model. In contrast, $P(T)$ represents the language model. For finding the best translation of a given source sentence, S we use a decoder to find the best probable target sentence, T. The decoder finds the n-best possible translations of S to T, out of which the translation with the highest probability is chosen, specified by $argmax_m$ in Equation 1.

The phrase-based technique provides promising results for language pairs that are structurally and morphologically similar. In the PBSMT framework, the distortion and lexicalised reordering models are widely used to handle reordering. The lexicalised reordering of Moses makes reordering simple for language pairs with analogous syntactic structures. However, the translation is quite bad for language pairs that are structurally and morphologically distant from each other with low resources, due to the problem of long-distance reordering. Although the basic PBSMT model

is expandable to account for long-distance reordering, it often performs worse due to distortion limitations [6]. Mapping at the grouping of surface forms of words does not reflect the pattern between the languages. Therefore, it is difficult for the PBSMT model to learn the translation pattern between the languages. To deal with it, the factored SMT model is adopted. In the factored model, a word is represented not only by its surface form, but by multiple levels called factors - such as lemma, POS, and morphological information. We are incorporating linguistic features into the corpus with this multiple-level representation of words. Incorporating linguistic information through factors will aid in learning a translation model that addresses the linguistic divergences between the languages. Works using factored SMT models significantly improve the translation result of PBSMT. One such implementation is for the Kannada language [7]. The factored model of SMT is mainly helpful in involving language pairs where one is morphologically rich and the other poor. We can feed more factors to improve our translation results better. The contributions of our work are:-

- We develop a small set of POS-tagged Manipuri corpora of the entertainment domain using the IL-POST (Indian Language Parts of Speech Tagset) [8] framework.
- We implement a PBSMT model as the basis for comparison.
- We implement factored SMT models with the in-built reordering option of Moses.
- We perform hand-coded reordering of the POS-tagged English side sentences of the training and testing corpus and implement a factored model of SMT.
- We compare the results of our systems using automatic evaluation metrics and establish that our architecture improves translation results.

3. LITERATURE REVIEW

Works that integrate linguistic information into the more routine PBSMT are limited. Here, we discuss some works that use the factors and factored SMT model in different ways to improve the translation result and address data sparsity, grammatical error, and fluency for morphologically rich languages in MT.

They describe one of the early approaches for translating French to English using linguistic information in work [9]. Utilising factored models for English-Latvian and English-Lithuanian SMT systems is the subject of yet another implementation report [10]. Latvian and Lithuanian languages are highly inflectional. They belong to morphologically rich, have free phrase order and are highly ambiguous, which results in data sparseness in translation. They have addressed this issue by splitting each token into its stem and suffix parts and treating them as separate models for Lithuanian-English translation. While for English-Latvian, morphologic tags are used as an additional language model apart from suffixes. Their work claims a significant improvement over baseline SMT through human evaluation. Similar experiments with phrase-based MT for English-Czech were conducted by [11], demonstrating the benefits of utilising multiple factors. His work involves different models involving combinations of word forms, lemma and morphological tags as factors. His work concludes with the BLEU (Bilingual Evaluation Understudy) score report demonstrating that multi-factor SMT consistently outperforms baseline SMT. In another study, a factored SMT model is implemented by [12] using fixed-length word suffixes that approximate POS tags in some ways. Their work reduces the language model's perplexity and increases the grammatical correctness of the results. Their work shows an improvement over the baseline SMT.

Contrary to translations from morphologically rich languages, translations to them have the limitation of lousy translation quality output. One major issue with morphologically rich languages is the data sparseness problem. Various reports associated with data sparseness are available for morphologically rich languages such as Latvian, Lithuanian, Croatian, Tamil, Malayalam, Mizo, Hindi, Kannada, and Farsi. Their work handles the data sparsity problem in translating a morphologically rich language [13]. They suggested a solution that generates unseen morphological forms fed into the training corpora. Their proposed solution claims to improve the translation quality through translation experiments of English to Hindi and Marathi languages.

Works that perform pre-processing, apart from having factors on the corpus, are available. One such factored SMT system for English to Tamil [14] is available. Their model uses lemma, POS and compound tag as factors on the source side; and lemma, POS tag and morphological information on the target side. They develop a novel pre-processing approach on the source language (English) so that it conforms to the target language (Tamil). The training uses the pre-processed sentences using a factored SMT model. Finally, morphological generators of the Tamil language generate a surface form of words from the factors output by the SMT model. The output result outperforms Google Translate and other systems. Pre-processing to change the source sentences' (English) structure by adding POS tags is another similar effort [15]. They use POS tags to modify the English sentences to be more similar to the richer Spanish and Catalan target sentences. However, some rely upon post-processing rather than pre-processing. In their work [16], they adopt post-processing techniques using syntactic and morphological knowledge of both source and target data to predict inflected forms of a sequence of word stems of the target side.

The work of [7] compares the factored model and the baseline model of the SMT technique for the morphologically rich Kannada translation. They create lan-

guage models based on surface form and POS tag for the factored model. They report that the factored model provides an improvement of 25% in BLEU compared to the baseline model. In another paper [17], morphological information factors comprising word stem, prefix, suffix, and POS tag - handle grammatical error correction. For evaluation, they modelled five individual systems with each factor - stem, prefix, suffix, and POS tag. Evaluation of the systems uses measuring their performance in the grammatical error correction task. Their models show an improvement in BLEU by 32.54% over the phrase-based model. Further, their model was experimented with official test data and compared with thirteen other systems at the "CoNLL 2014 shared task", in which they got the 7th and 5th F0.5 scores. They concluded that POS information is the most effective, out of all factors, and that the model with POS information outperforms others with other factors. Even with the latest technique of NMT, [18] they apply pre-ordering to reduce word order divergence between source and target languages for a few resourced Indic languages. Their approach relatively improves the translation quality even in low-resource scenarios.

For the Manipuri language, [19] they work reports the only factored SMT model implementation. They have used suffix and dependency relations as factors on the source (English) side and case markers on the target (Manipuri) side. Their system was trained on 10,350 sentences and tested on 500 news domain datasets. BLEU score and subjective evaluation claim an improvement in their result. MT systems for the Manipuri language are getting built but are far from perfect. Currently there isn't any Manipuri-language content that makes use of POS tags or sentence restructuring. Their study [20] evaluates the effectiveness of unsupervised MT models for Manipuri-English translation using a comparable corpus of news domain. They use a suffix segmenter using graph-based stemmer and transliteration models to re-score the sentence translation and the lexical probability. They report that the unsupervised SMT model is more successful than the unsupervised NMT models for the language pair. Using data augmentation approaches, they addressed low-resource problems while experimenting with a semi-supervised approach [21]. The data augmentation process uses comparable monolingual corpora from the news domain. They employ a self-training and back-translation approach to produce synthetic parallel data from monolingual data. According to reports, their model outperforms unsupervised, supervised, mBART, and other standard semi-supervised models in quantitative efficacy and can handle sparse data. Additionally, they make empirical claims about how well their models cope with uncommon words and long-term dependencies.

Their study [22] uses the multilingual pre-trained models mBART50 and mT5-base with fine-tuning for transfer learning of low-resource MT involving Assamese, Manipuri, and Bengali languages. Their fine-tuned models outperform the multilingual baseline model indicated by their BLEU scores. On the WAT-2021 test set, their model with mT5-base fine-tuning performs best. Their model, however, predicts a lower score for the Flores-101 test set. Another experiment with a many-to-many NMT model with cross-lingual capabilities is reported [23]. Their work enhances the basic paradigm for many-to-many translation between Manipuri and English and the bilingual model. Additionally, they use zero-shot translation to examine the generalizability of their methodology on language pairs with no direct correspondence and contrast it with pivot-based translation. To translate English to Manipuri, [24] examine the supervised and unsupervised SMT and NMT techniques. They also test out low-resource techniques like self-training and back translation. They examine the difficulties and mistakes made in translating English to Manipuri. Works of [25], [26] and [27] are a few other NMT implementations for the Manipuri language. Recently, Google Translate [28] supported Manipuri using the Meitei-Mayek script, using the concept of Zero-Shot translation [29].

4. PROBLEM STATEMENT

SMT and NMT are the techniques that dominate MT. Dealing with dissimilar language pairs in low-resource settings is still challenging for both techniques. Based on the language pair under consideration, some systems favour SMT [30], while some favour NMT [31]. SMT is good at handling adequacy but trades for fluency [2], while NMT trades adequacy for fluency [32]. NMT has replaced most larger MT systems, which have larger corpora. However, SMT has suggested a better option for low-resource conditions. SMT uses mathematical models to map the source and target language symbols. The mapping is more straightforward for similar languages, and the results are promising. Despite the state-of-the-art PBSMT's usage of reordering models, the inconsistency in word ordering between distant languages leads to subpar translation quality. We pose the divergences between languages and the inability of lexicalised reordering to handle long-distance reordering as a problem. We propose prior source-side reordering to mitigate the syntactic differences and handle long-distance reordering to improve the translation result.

5. PROPOSED METHODOLOGY

As mentioned above, English and Manipuri languages have significant linguistic differences in comparison. Their morphological and structural differences increase the challenges of translation. To treat this, we first employ separate pre-processing modules for both English and Manipuri sentences, followed by training. System training uses our proposed model, followed by a transliteration module, for handling untranslated tokens present in the output. Further explanation is in the following subsections.

5.1. PRE-PROCESSING

5.1.1. English Sentences

The pre-processing of English sentences follows the order of tokenising, factoring and reordering. Tokenising of the English side uses the inbuilt tokeniser of Moses. A factoring module then treats the tokenised output to extract the factors. The factors may be lemma, POS and morphology. Here, we use the maximum entropy-based MXPOST [33] framework to include only POS tags as factors, as the target side has only POS tags as a factor. The MXPOST uses the Penn treebank POS tagset. Using the POS information, we reordered the factored sentences using the linguistic rules of Manipuri. Reordering transforms the structure of English sentences to make them structurally more similar to Manipuri. This transformation will reduce the syntactic divergences between the language pair, thereby better mapping between source and target symbols. The source text's prior reordering has two effects on how well the MT performs [34]. Firstly, it handles long-distance reordering and thus, reduces the workload on the reordering model by prior reordering the source text. Only trivial reordering occurs during decoding, and the translation hypothesis's construction uses a monotone orientation. Secondly, prior source reordering should result in more accurate word alignments, better translation models and higher translation quality since statistical word alignment approaches function efficiently for linguistic groups with analogous grammatical structures. Fig 1. represents the pre-processing module of English sentences. As an example, we have an English sentence below.

Example:

(Before pre-processing)

I am a boy.

(After factoring)

I|PRP am|VBP a|DT boy|NN .|.

(After reordering)

I| PRP boy| NN am| VBP a|DT .|.

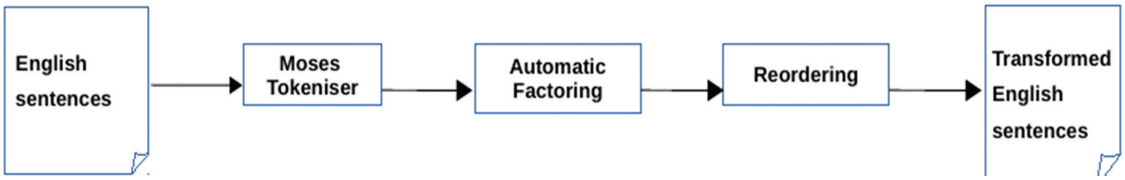


Fig. 1. English side pre-processing

5.1.2. Manipuri Sentences

Fig. 2 shows the pre-processing module of Manipuri sentences. The process is similar to English for Manipuri side pre-processing, except the reordering step is not present. The tokenising step here uses the dedicated indicNLP [29] tokeniser for the Manipuri language.

Tokenising is followed by factoring. Higher the number of factors, the more refined our system is. For our current work, we have considered only the POS tag as the factor due to the high cost of manual preparation, as open-source NLP tools are not available for Manipuri. We manually develop our annotated corpora of the entertainment domain using the ILPOST for the Manipuri language.

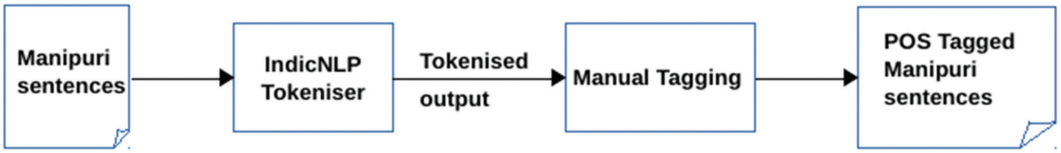


Fig. 2. Manipuri side pre-processing

5.2 Architecture of the System

The overall architecture of our system is in Fig. 3. We use the outputs of pre-processing modules of Fig. 1 and Fig. 2 as the dataset to train our system. There are no POS taggers or annotated corpora for the Manipuri language. Consequently, we developed the dataset for this research with the aid of a linguist. It is essential to specify that the dataset we use to train our system is relatively small owing to financial reasons. We split up the dataset into training, development, and test set using Python's split code. We employ a training set of 8,000 sentences, a development set of 1,000 sentences, and a test set of 2,000 sentences from the entertainment domain. The source side test data undergoes the pre-processing of Fig.1 before feeding for translation. Therefore, the translation step in our system uses the pre-processed text. Adequacy measures the mapping between source and target symbols, which is not an issue for translations involving language pairs with the same structure and morphology. However, when the structure and morphology are different, it is a problem that needs addressing. The language divergence makes it more challenging to map source and target symbols correctly, affecting the adequacy and fluency of translation results, which are the measures of translation quality.

Our architecture uses factoring followed by hand-coded reordering on the source side to address this issue. Through reordering, we attempt to reduce the structural divergences; thus, the translation output will closely resemble the target side language, thereby providing a better mapping mechanism between source and target symbols. Furthermore, when using a factored model, we consider the POS feature in addition to the surface form to produce the translated output. Therefore, incorporating a POS tag on the dataset provides language-specific linguistic knowledge to the training model. Lastly, given the small training data, untranslated words will appear in the translated output, which uses transliteration as a post-processing step.

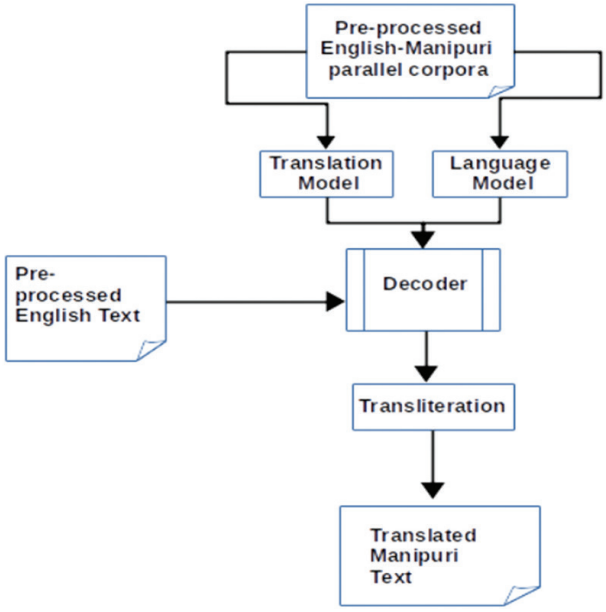


Fig. 3. Architecture of our proposed model of factored English-Manipuri system

6. RESULTS AND DISCUSSION

Our language pair, as mentioned above, has different word order. The distortion limit limits the distance between the following and previously translated phrases, which might seriously impair disparate languages. We, therefore, set the distortion limit to 0.5 instead of the default 0.3 to increase the search space of the decoder. For comparison, we experiment with multiple systems with different parameter settings using the same data size. One system is the general PBSMT which we use as a basis. This system data does not have any factoring or explicit reordering involved.

We also train three different factored models. FSMT1 (factored model with word-based reordering), FSMT2 (factored model with phrase-based reordering) and FSMT3* (factored model with hand-coded reordering on source side). The FSMT3* is the model which uses our proposed architecture model of Fig. 3. It is essential to mention that in FSMT3*, before translation, the source side of the test set first undergoes factoring, followed by hand-coded reordering.

Table 1. Score comparison of the systems

	PBSMT	FSMT1	FSMT2	FSMT3*
BLEU	3.58	3.47	3.71	4.12
F1	0.181	0.181	0.183	0.193
Meteor	0.098	0.093	0.10	0.1
Precision	0.170	0.175	0.171	0.187
Recall	0.194	0.189	0.195	0.2

Table 1 shows the BLEU [35], F1, Meteor [36], Precision and Recall scores of these systems tested using a test set size of 2000 sentences. We find that PBSMT outperforms FSMT1 in terms of BLEU, Meteor, and Recall scores, except Precision - which suggests that phrasal reordering obscures FSMT1's POS feature information. Therefore, the FSMT2 model, which utilises both POS characteristics and lexicalised phrasal reordering, performs better in all scores than PBSMT and FSMT1. However, the proposed model, FSMT3*, proves a further improvement in all scores, even over the FSMT2 model discussed above. FSMT3* outperforms the FSMT2 model by 11.05% (BLEU), 5.46% (F1), 9.35% (Precision), and 2.56% (Recall), even with scant training data. Our result shows that reordering the English sentences as per the Manipuri syntax along with POS features improves the translation quality, even when the dataset is small.

7. CONCLUSION

Our work found that handling linguistic divergences is lucrative in MT. Pre-ordering the source side and adding POS as a linguistic characteristic increased the scores by 11.05% (BLEU), 5.46% (F1), 9.35% (Precision), and 2.56% (Recall), respectively, from that of FSMT2. Despite being small, this gain represents a significant improvement given the dataset and feature limitations. Prior reordering handles long-distance reordering, thus mitigating language divergences. Furthermore, language-specific hand-coded reordering of the source side chunks that match the target language chunks provides better alignment than the lexicalised reordering option of Moses. It thus improves translation quality even under low-resource settings.

Our current work uses a small dataset prepared for experimental purposes, and reordering is also manual. If Manipuri-specific NLP tools are available, we can automate the factoring and reordering process by incorporating other features to refine our results further. Our hand-coded reordering model on the source side with POS as a feature for the low-resource English-Manipuri translation is a bootstrapping strategy towards reducing the linguistic gap for enhancing translation quality. The goal is to find techniques to close the linguistic divergence gap because MT models never produce the best word alignments for languages with far-off linguistic features. Due to our study's usage of different scripts, we exclude a comparison with the Google Translate result.

8. REFERENCES

[1] P. Passban, "Machine Translation of Morphologically Rich Languages Using Deep Neural Networks", PhD diss, Dublin City University, 2017.

[2] S. Sen, M. Hasanuzzaman, A. Ekbal, P. Bhattacharyya, A. Way, "Neural machine translation of low-resource languages using SMT phrase pair injection", Nat. Lang. Eng., Vol. 27, No. 3, 2021, pp. 271–292.

[3] G. Lample, M. Ott, A. Conneau, L. Denoyer, M. Ranzato, "Phrase-Based & Neural Unsupervised Machine Translation", arXivarXiv preprint arXiv:1804.07755, 2018.

[4] P. Koehn, "Europarl: A Parallel Corpus for Statistical Machine Translation", Proceedings of Machine Translation Summit X: Papers, Phuket, Thailand, September 2005, pp. 79–86.

[5] J. V. Stone, "Bayes' Rule: A Tutorial Introduction to Bayesian Analysis", Sebtel Press, 2013.

[6] P. Koehn, H. Hoang, A. Birch, C. Callison-Burch, M. Federico, N. Bertoldi, B. Cowan, W. Shen, C. Moran, R. Zens, C. Dyer, "Moses: Open Source Toolkit for Statistical Machine Translation", Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics Companion Volume Proceedings of the Demo and Poster Sessions, Prague, Czech Republic, June 2007, pp. 177–180.

[7] K. M. Shivakumar, N. Shivaraju, V. Sreekanta, D. Gupta, "Comparative study of factored SMT with baseline SMT for English to Kannada", 2016 International Conference on Inventive Computation Technologies (ICICT), August 2016, vol. 1, pp. 1–6.

[8] LDC-IL, https://www.ldcil.org/standardsTextPOS.aspx (accessed: 2022)

[9] P. F. Brown, S. A. Della Pietra, V. J. Della Pietra, J. D. Lafferty, R. L. Mercer, "Analysis, statistical transfer, and synthesis in machine translation", Proceedings of the Fourth Conference on Theoretical and Methodological Issues in Machine Translation of Natural Languages, Montréal, Canada, June 1992.

[10] A. Utka, J. Vaičenonienė, J. Kovalevskaitė, "Human Language Technologies – The Baltic Perspective: Proceedings of the Ninth International Conference Baltic HLT 2020", IOS Press, 2020.

[11] O. Bojar, "English-to-Czech Factored Machine Translation", Proceedings of the Second Workshop on Statistical Machine Translation, Prague, Czech Republic, June 2007, pp. 232–239.

[12] N. Sharif Razavian, S. Vogel, "Fixed Length Word Suffix for Factored Statistical Machine Translation", Proceedings of the ACL 2010 Conference Short Papers, Uppsala, Sweden, July 2010, pp. 147–150.

[13] P. D. Dungarwal, "Reordering Models for Statistical Machine Translation: A Literature Survey", Indian Institute of Technology, Bombay, India, 2014.

[14] A. Kumar, "Factored Statistical Machine Translation System for English to Tamil Language", Pertanika Journal of Social Sciences & Humanities, Vol. 22, No. 4, 2014.

[15] N. Ueffing, H. Ney, "Using POS Information for SMT into Morphologically Rich Languages", 10th Conference of the European Chapter of the Association for Computational Linguistics, Budapest, Hungary, April 2003.

[16] E. Minkov, K. Toutanova, H. Suzuki, "Generating Complex Morphology for Machine Translation", Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague, Czech Republic, June 2007, pp. 128–135.

[17] R. Wang, C. Ding, M. Utiyama, E. Sumita, "English-Myanmar NMT and SMT with Pre-ordering: NICT's Machine Translation Systems at WAT-2018", Proceedings of the 32nd Pacific Asia Conference on Language, Information and Computation: 5th Workshop on Asian Translation: 5th Workshop on Asian Translation, 2018.

[18] R. Murthy, A. Kunchukuttan, P. Bhattacharyya, "Addressing word-order Divergence in Multilingual Neural Machine Translation for extremely Low Resource Languages", Proceedings of the 2019 Conference of the North, Minneapolis, Minnesota, 2019, pp. 3868–3873.

[19] T. D. Singh, S. Bandyopadhyay, "Manipuri-English Bidirectional Statistical Machine Translation Systems using Morphology and Dependency Relations", Proceedings of the 4th Workshop on Syntax and Structure in Statistical Translation, August 2010, pp. 83-91.

[20] L. Laitonjam, S. Ranbir Singh, "Manipuri-English Machine Translation using Comparable Corpus", Proceedings of the 4th Workshop on Technologies for MT of Low Resource Languages (LoResMT2021), Virtual, August 2021, pp. 78–88.

[21] S. M. Singh, T. D. Singh, "Low resource machine translation of english–manipuri: A semi-supervised approach", Expert Systems with Applications, Vol. 209, 2022, pp. 118-187.

[22] S. M. Singh, L. Sanayai Meetei, A. Singh, T. D. Singh, S. Bandyopadhyay, "On the Transferability of Massively Multilingual Pretrained Models in the Pretext of the Indo-Aryan and Tibeto-Burman Languages", Proceedings of the 18th International Conference on Natural Language Processing (ICON), National Institute of Technology Silchar, Silchar, India, December 2021, pp. 64–74.

[23] S. M. Singh, T. D. Singh, "An empirical study of low-resource neural machine translation of manipuri in multilingual settings", Neural Computing and Applications, Vol. 34, No. 17, 2022, pp. 14823–14844.

[24] T. J. Singh, S. R. Singh, P. Sarmah, "English-Manipuri Machine Translation: An empirical study of different Supervised and Unsupervised Methods", 2021 International Conference on Asian Language Processing (IALP), December 2021, pp. 142–147.

[25] S. M. Singh, T. D. Singh, "Statistical and Neural Machine Translation Systems of English to Manipuri: A Preliminary Study", Soft Computing and Signal Processing, Singapore, 2021, pp. 203–211.

[26] L. Rahul, L. Meetei, H. Jayanna, "Statistical and Neural Machine Translation for Manipuri-English on Intelligence Domain", Advances in Computing and Network Communications, pp. 249–257, Springer Singapore, 2021.

[27] S. M. Singh, T. D. Singh, "Unsupervised Neural Machine Translation for English and Manipuri," Proceedings of the 3rd Workshop on Technologies for MT of Low Resource Languages, December 2020, pp. 69-78.

[28] Google Translate, https://translate.google.co.in (accessed: 2022)

[29] Google Translate adds support for Assamese, Mizo and Manipuri languages - Eastern Mirror, https://easternmirrornagaland.com/google-translate-adds-support-for-assamese-mizo-and-manipuri-languages (accessed: 2022)

[30] M. Dowling, T. Lynn, A. Poncelas, "SMT versus NMT: Preliminary comparisons for Irish", Association for Machine Translation in the Americas (AMTA), 2018.

[31] S. Kinoshita, T. Oshio, T. Mitsuhashi, "Comparison of SMT and NMT trained with large Patent Corpora: Japio at WAT2017", Proceedings of the 4th Workshop on Asian Translation (WAT2017), November 2017, pp. 140–145.

[32] P. Koehn, R. Knowles, "Six Challenges for Neural Machine Translation", arXiv preprint arXiv:1706.03872, 2017.

[33] A. Neumann, nltk-maxent-pos-tagger, https://github.com/arne-cl/nltk-maxent-pos-tagger/blob/023241f9deceeb214cef7304b5a8ebc914024dfd/mxpost.py (accessed: 2022)

[34] M. Holmqvist, S. Stymne, L. Ahrenberg, M. Merkel, "Alignment-based reordering for SMT", Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12), Istanbul, Turkey, May 2012, pp. 3436–3440.

[35] Tilde MT, https://www.letsmt.eu/Bleu.aspx (accessed: 2022)

[36] The METEOR Automatic MT Evaluation Metric, http://www.cs.cmu.edu/~alavie/METEOR/ (accessed: 2022)

Scene Based Text Recognition From Natural Images and Classification Based on Hybrid CNN Models with Performance Evaluation

Original Scientific Paper

Sunil Kumar Dasari
Department of ECE, SOE,
Presidency University, Bangalore, India-560054
sunilkumar.d@presidencyuniversity.in

Shilpa Mehta
Department of ECE, SOE,
Presidency University, Bangalore, India-560054
shilpamehta@presidencyuniversity.in

Abstract *– Similar to the recognition of captions, pictures, or overlapped text that typically appears horizontally, multi-oriented text recognition in video frames is challenging since it has high contrast related to its background. Multi-oriented form of text normally denotes scene text which makes text recognition further stimulating and remarkable owing to the disparaging features of scene text. Hence, predictable text detection approaches might not give virtuous outcomes for multi-oriented scene text detection. Text detection from any such natural image has been challenging since earlier times, and significant enhancement has been made recently to execute this task. While coming to blurred, low-resolution, and small-sized images, most of the previous research conducted doesn't work well; hence, there is a research gap in that area. Scene-based text detection is a key area due to its adverse applications. One such primary reason for the failure of earlier methods is that the existing methods could not generate precise alignments across feature areas and targets for those images. This research focuses on scene-based text detection with the aid of YOLO based object detector and a CNN-based classification approach. The experiments were conducted in MATLAB 2019A, and the packages used were RESNET50, INCEPTIONRESNETV2, and DENSENET201. The efficiency of the proposed methodology - Hybrid resnet -YOLO procured maximum accuracy of 91%, Hybrid inceptionresnetv2 -YOLO of 81.2%, and Hybrid densenet201 -YOLO of 83.1% and was verified by comparing it with the existing research works Resnet50 of 76.9%, ResNet-101 of 79.5%, and ResNet-152 of 82%.*

Keywords*: CNN, Scene based text detection,RESNET50,Text Detection,YOLO*

1. INTRODUCTION

Now day's digital world has increased its digital image or video database sizes, which is used for social media, digital education, commercial purpose, etc. Image-based retrieval has become a dynamic research such as computer vision, artificial intelligence, pattern recognition, etc. The CBIR method is of two types which are 1. Text-based 2. Content-based systems. The text-based method used for Google, Yahoo and Bing, etc., focuses on text keyword searches. The text-based image recovery method has numerous confines. For instance, once the keyword 'Apple' is given, retrieved results for the Apple laptop, company logo, apple fruits, etc. However, this text keyword search does not achieve good performance and does not describe the visual content of an image's properties.

The above problem is rectified by proposing a hybrid method; currently, many academics focus on hybrid features with re-ranking. The hybrid model is defined as the feature level that provides a stabilized order to retrieve similar images. The CBIR technique has resolved this problem, which uses hybrid features centered on low-level aspects like color, texture, shape, etc. Image characteristics are defined by low-level features. Those features that describe visual content might refer to color, texture, shape, etc., and hence, it helps to retrieve similar results and achieves better performance. In today's society, people learn various subject matter aspects that are related to different domains and industries. The subject matter learned consists of varied amounts of knowledge and concepts. With the different types of subjects that are available on a global scale, there exists a requirement to assess the people to

see whether the supplied knowledge is encompassed within the people.

This section will give a brief introduction to the significance of scene-based text recognition and classification techniques. This part will deliver a momentary summary of the state of the art of text classification techniques used for several applications. This section provides the background of the research paper, including the overview of deep learning models for text recognition, along with the challenges associated with it. The section will further elaborate on the objectives of the study and research significance. In the 1950s, the automatic form of text summary instigated an impression, and in the late 1950s, Hans Peter Luhn issued an article named "Automatic creation of literature abstracts" [1] which could deploy features like word and sentence frequency to excerpt elementary stretches from the text for summary purposes.

Conferring to Radev et al. Summary is demarcated as "a text produced from one or more texts that convey the important information in the original text(s), and that is no longer than half the length of the original text(s) and usually significantly less than that." [2] Abstracts assist in realizing the implication of the text. Text summarization supports consumers in accomplishing a huge quantity of data by summarizing documents and integrating additional appropriate evidence.

The text summarization procedure consists of 3 main phases such as analysis, transformation, and combination [3]. The authors of [4] examined and related the performance of 3 diverse processes. The difficulty of scene text recognition compacts with identifying text in natural scene images. In conventional means, text recognition was engrossed in identifying printed text in documents, and such kind of schemes projected images to be black and white and in a document style layout embracing text lines.

In this research, an efficient technique is presented to perceive texts in natural scene images. Also constructs the "CRF model" by relating "CNN scores of MSERs" and multiple neighborhood information. Moreover, missing text mechanisms were additionally recovered using perspective information. Additionally, it incorporates gray and binary features and designs shape-specific classifiers to exactly validate text lines. The suggested system has been estimated on four public benchmarks and succeeds in hopeful performance, representing the efficiency and robustness of this method. The most apparent drawback of CRF is the high computational intricacy of the training phase of the system.

This research work focuses on scene image text detection with the aid of a novel approach. A combination of YOLO and RESNET is the first case, combined INCEPTIONRESNET, AND YOLO and combined DENSENET201 AND YOLO approach. The precision value for all those methodologies will be evaluated for better output.

The other sections of the work are as follows: Section 2 explains the Literature Review that compares the existing research on video-based database systems. Section 3 defines the research methodology - a novel scene-based text recognition and classification model using hybrid CNN. Section 4 explains the analysis of simulation results and performance evaluation of the suggested methodology. Finally, section 5 describes the conclusion of the research.

2. LITERATURE REVIEW

In video-based database systems, every video is tagged manually with the support of several keywords to assist with the searching and retrieval process. This system is found to be protracted and unreliable in nature; for instance, any two users could use various keywords in order to look for a similar video. Another scheme is to extract keywords from text which appears in the frame [5]. Mainly, any video text is divided into two groups such as graphics text and scene text. During the editing section, the graphics text is added to the video content separately. Scene text occurs logically in the prospect captured by the camera. By means of the rapid development of the Internet, there exists an aggregate demand for text detection from video.

Several techniques have been established since earlier research, but text detection still faces several challenges. The major aspects which show the disadvantage is unconstrained colors, size, and alignment of characters. Also, the scene text gets exaggerated by lighting environments and perspective falsifications [6]. Text detection could be mainly categorized into three different means such as "component-based, edge-based and texture-based approaches" [7,8]. Owing to color draining and short divergence of text lines, connected components or elements might not preserve the complete shape of the character, and hence, these approaches cannot be applicable to video images.

There is a vast amount of data found online, so obviously, summarizing the information has special importance. Text summarization generates an overall summary of the given document. There are two different summarization techniques that can be used to generate the summary, extractive and abstractive. But the majority of Indian language summarization works focused on the extractive approach since the grammar rules are a little more complex in such languages to generate an abstract one.

The uses of image processing are developing each day, and some of its applications are: -

1. Pattern recognition - Pattern recognition involves studying and recognizing various patterns like handwriting. It is integrated with artificial intelligence like computers. Assisted diagnostics and handwriting can be done easily.
2. Video Editing - It is also a digital image processing field. A collection of frames or photos is arranged in such a way that it makes the photo flow faster. Includes reduced motion detection of sound movement and color space modification etc.

3. Image sharpening – In this, the appearance and field of an image could be changed. Image sharpening manages the image and gives the desired effect. It includes individual modification, blurring, sharpening, retrieval, and image recognition.
4. Robot Vision - Many robotic machines operate with the use of digital images, where they can see their paths, for example, the roots of obstacle detection and the next line of robots.

Edge-based systems were proposed in order to overcome the low contrast issues. Arithmetical topographies from Sobel edge maps were extracted [9] in 4 directions, and K-means were deployed to determine the pixels into text and non-text clusters. Though this method was found vigorous alongside the composite environs, it could not meet up with detecting low contrast text and text of a lesser font. Moreover, it was more exclusive to accomplish due to the huge feature set. Two filters were deployed in [10] to develop the edges in the text areas, and here, several threshold values were employed to choose whether to progress the edges in a particular area. Thus, it could not simplify in a better way for different datasets. To discover the potential line segments, research [11] analyzed the maximum gradient difference. The multi-orientation problem was addressed in [12] with the support of a system based on the laplacian and skeletonization concepts. This technique ended in a high false positive rate and misdirection ratio for multi-oriented text. This is owing to the heuristics entailed in the segmentation approach [13]. A novel method based on a 2-level classification system and two sets of features specifically developed for capturing equally inherent features of subjectively oriented video texts were suggested in [14]. Gradient Vector Flow based method was proposed to detect the scene text which deeds character gap [15]. This approach did not meet the necessary general symmetry points due to the small font, low contrast, and complex background in the video. The Discrete Cosine Transform (DCT) factors of the intensity images have been extensively deployed as the texture elements for text detection [16-18].

Analysis of visuals like images and videos and their evolving systems is a stimulating and significant task to be enhanced and accomplished on benchmark datasets. This issue is resolved by deploying the STN-OCR model, including deep neural networks and spatial transformer networks. This research network design model consists of 2 phases such as localization and recognition network. In the localization network, it advocates text regions and creates the sampling grid [19]. A novel approach based on wavelet medium moments is presented in [20], along with a new indication of angle projection for perceiving multi-oriented text in the video.

A "deep neural network" for perceiving text in natural scene images was suggested in [21], which uses the drawbacks of arbitrarily oriented texts and intricate background images. This approach mentioned is constrained to natural scene images then not applicable to video as it includes several actions in it. A novel method was suggested in [22] for identifying scene text with the aid of deep reinforcement learning. Here, an agent is provided with a state and learns to predict future return aspects and develops sequential-based decisions to identify scene texts. For reading scene text in the wild centered on scene text proposal, the new method is presented in [23]. For natural scene text detection, [24] leveraged color prior centered MSER which extracts stroke features with support of strokes width distance depending on segmented edges. Several drawbacks still exist, such as multi-type texts, graphics, and scene texts. In [23], "Fourier laplacian filtering and hidden Markov model" is suggested for text and non-text classification. A two-staged approach [25] is suggested by means of a quadrilateral scene text detector, but these structural aspects might not hold for the irregular form of text. An adaptive bezier curve-based network is deployed in [26] for text detection in scene images. Similarity prediction among the textual elements of several views of natural scene images was presented in [27] for attaining advanced performance. A scene text detection-based segmentation system is proposed in [28], which introduced the text-mountain model. An edge descriptor model is used in [29] to discover local binary patterns. CNN model is explored in [30] for multi-lingual based text detection for detecting natural scene images. To enhance the capability of recognition-based methods, research in [31] introduced fractals, wavelet transforms, and optical flow for undertaking the drawback of video and natural scene images. Dependencies among word tokens in a sentence are extracted in [31], which assist 2D spatial dependences among two characters in a scene text image.

3. RESEARCH METHOD

The research methodology includes the proposed research flowchart given in Figure 1 for designing a novel scene-based text recognition and classification model using hybrid CNN. The following steps, such as data collection, data preprocessing, feature extraction, and classification, are discussed as follows:

3.1 DATA COLLECTION

In the object detection phase, "ground truth" states data gathered at a particular location. Ground truth permits image sequence data to be connected to physical features and materials on the ground. The assembly of ground truth data empowers the discovery of the entities in the image or video format.

3.2 DATA PREPROCESSING

Data preprocessing techniques include text filtering, binarization, image segmentation, etc. Image prepossessing are steps taken to format images before they are used by model training and inference systems. This contains re-sizing, orienting, and color corrections. Bounding box on all regions in the image and selects the point with the highest score as the center of the crop.

3.3 FEATURE EXTRACTION

Feature extraction is described as extracting relevant features from data. Here, the function of max pooling could minimize feature map dimension and solves the fitting issue. By depending upon these feature values, the network could be trained and tested.

K-medoids Clustering is an unsupervised means of clustering system that cluster objects in labeled data. In "K-medoids Clustering," as an alternative to taking the centroid of objects in a cluster as a location point as in k-means clustering, this research considers medoid as a reference point.

In ResNet50, layers are unequivocally reformulated as learning residual functions with respect to layers input. The resnet50 could categorize images into 1000 object groups, including keyboard, mouse, pencil, and many animals. As an effect, the network has learned rich feature depictions for an inclusive array of images. The network has an image input size of "224-by-224" for further pre-trained networks in "MATLAB."

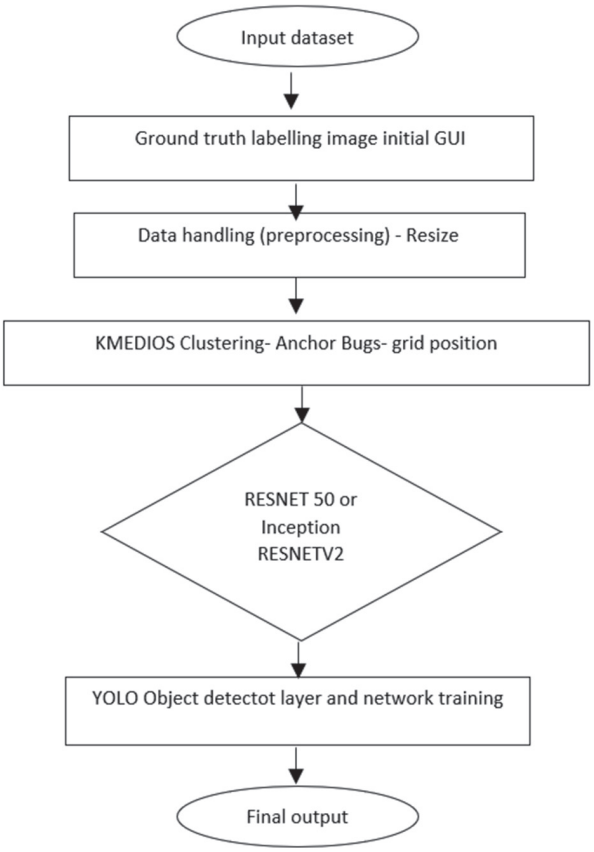


Fig. 1. Proposed Approach

"Inception-ResNet-v2" is a convolutional neural network, and it is qualified on further than a million images from the ImageNet database. The network is 164 layers deep and can categorize images into 1000 object groups like the keyboard, mouse, pencil, and various animals; as a consequence, the network has learned rich feature depictions for a varied series of images and has an image input size of "299-by-299".

DenseNet-201 is a convolutional neural network that is 201 layers deep, and a pre-trained form of network proficient on more than a million images from the "ImageNet database" could be loaded. The pre-trained network could categorize images into "1000 object" types such as keyboard, mouse, pencil, and many animals. As an effect, the network has learned rich feature demonstrations for a comprehensive series of images.

The proposed model encompasses a neural network along with Res Net as a feature extractor and YOLO v2 for classification. The selection of these two systems is to progress the efficacy of the object discovery scheme by increasing accuracy and identifying even smaller objects exactly. A neural network is fashioned by using the ResNet network, which does the feature extraction manner. ResNet is deployed for its enhanced learning and accuracy with deeper networks. The deprivation problem occurs in deeper networks which saturates the accuracy of the model by restating the similar process again and again at greater levels. To overcome this problem, the ResNet model is used, which evades a step that is administered more than twice, so that accuracy may not saturate even with deeper networks. Mainly for classification, a fully connected network is transformed into YOLO v2 network.

YOLO is a single-stage entity-based recognition network, and it has a fast recognition speed. It is qualified by dense and consistent sampling above locations, scales, and an end-to-end flow. YOLO is a system that uses neural networks to afford instantaneous object detection. This system is standard for its speed and accuracy. It has been deployed in innumerable applications to perceive traffic signs, people, parking meters, and animals. In our project, we used to detect the Devanagari images.

Anchor Box:

The "YOLOv2" model fragments the input image into N×N grid cells, and all grid cell has the task of restricting an object if the midpoint of that particular object falls in a grid cell. If the midpoint of 2 objects corresponds with one another, the detection system could merely prefer any one of the objects. To resolve this problem, the conception of "Anchor Boxes" is used.

The YOLO algorithm is significant owing to subsequent details:

This system advances promptness of recognition as it can envisage objects in real time. YOLO is a prognostic system that offers precise effects with insignificant background errors. The system has excellent learning competencies that empower it to learn depictions of objects and relate them to object discovery.

3.4 CLASSIFICATION USING HYBRID CNN MODEL

The classification process is executed by using a hybrid CNN-RNN model, and a detailed analysis of the proposed system architecture is given and evaluates the methodology involved in the text recognition

and classification process. And finally, the network is trained with the help of trainYOLOv2ObjectDetector with SGDM (stochastic gradient descent with momentum) optimizer and detects the output likewise. In a convolution neural network, the convolutional filters remove the noise images. Filters detect spatial patterns such as edges in an image by detecting the changes in intensity values of the image.

Convolutional layers are strong feature extractors in which the convolutional filters are capable of finding features of images. The function of max-pooling layers is to reduce the size of feature maps and solve over fitting problems. Based on these features value, the data is trained and tested. After this PYTHON is integrated in order to recognize the character which is given as input. Here Easy-OCR is used for recognizing the character of the object. Optical Character Recognition is a piece of technology that can identify text in digital images. Text in scanned documents and photos is frequently recognized using this technique.

An easy-to-use Python tool called Easy-OCR makes optical character recognition possible for computer vision engineers. Optical character recognition is by far the easiest to use when it comes to OCR, from that an Easy-OCR is the most popular method. Easy-OCR is a straightforward, lightweight library to use in comparison to others, as the name suggests. It supports a variety of languages. Additionally, it may be accomplished to perform better for particular use cases by adjusting various hyper-parameters. There is also another OCR which is Tesseract-OCR but it will not provide better result based on speed and accuracy when compared to Easy-OCR, so Easy-OCR is chosen. After Object Character Recognition it will be converted to ASCII key and then given to MATLAB GUI to finally identify the text from the natural image provided as input.

4. RESULTS

The analysis of simulation results and performance evaluation of the suggested methodology is described as follows.

1) Initial GUI

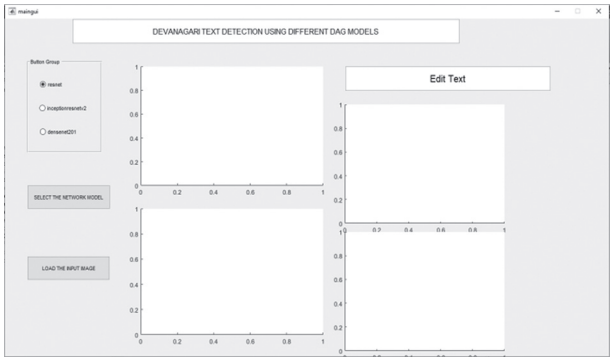


Fig. 2. Initial GUI

Figure 2 shows the initial GUI when the input image is taken. Figure 3 denotes the combined form of RESNET and YOLO output, where the text detected is shown separately. The detected text is shown separately by means of the EC-OCR coding model.

2) Combined RESNET and YOLO output

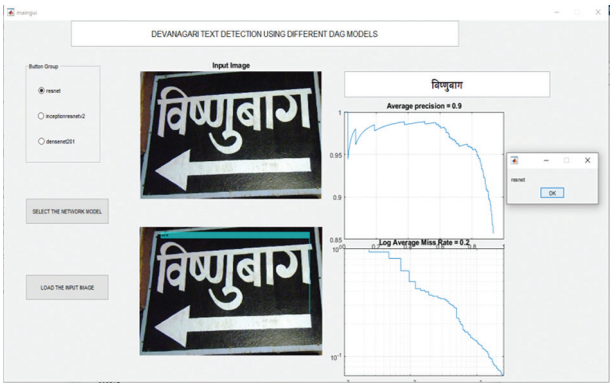


Fig. 3. Combined RESNET and YOLO output

3) Combined INCEPTIONRESNET and YOLO output

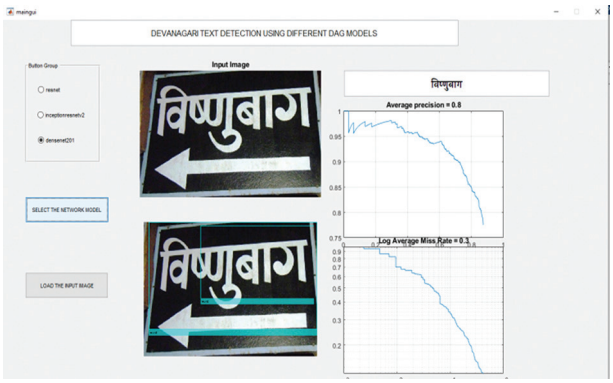


Fig. 4. Combined INCEPTIONRESNET and YOLO output

Figure 4 represents the combined form of INCEPTIONRESNET and YOLO output, where the text detected is shown separately in a white box. Here, the detected text is shown separately by means of the EC-OCR coding model.

4) Combined DENSENET201 and YOLO output

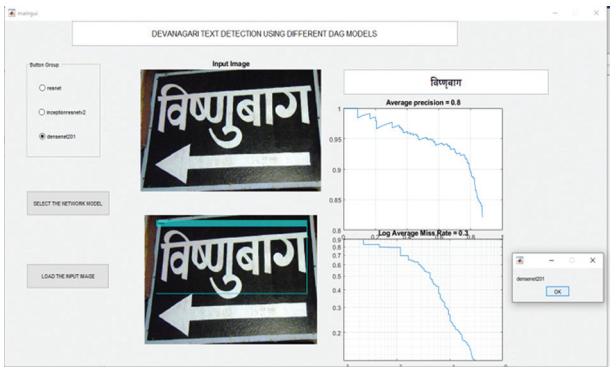


Fig. 5. Combined DENSENET201 and YOLO output

Fig. 5. represents the combined form of DENSENET 201 and YOLO output, where the text detected is shown separately in a white box. Here, the detected text is shown separately by means of the EC-OCR coding model and, the accuracy graph is extracted for each model.

Accuracy is a fraction of true positive occurrences to all positive examples of objects in the detector, centered on ground truth. For a multiclass indicator, average precision is a vector of average precision scores for all object classes.

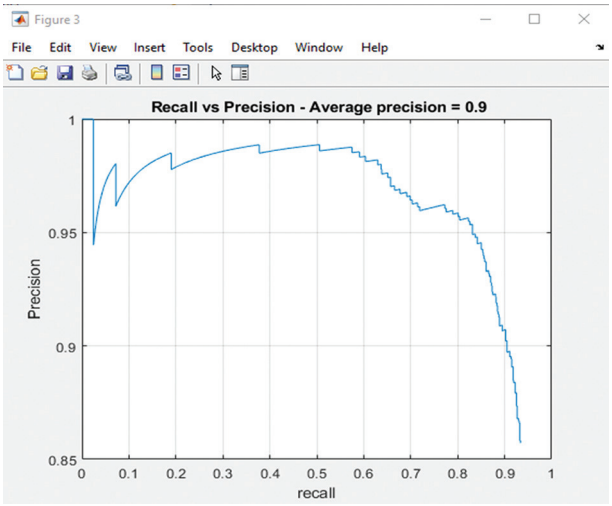


Fig. 6. Representation of the total word accurately detection rate in each and every images

Log miss rate, resumed as either one vector of numeric scalars or as cell arrangement. For the multiclass indicator system, FPPI and log miss rate were cell arrays, where every cell covers data points for every object class.

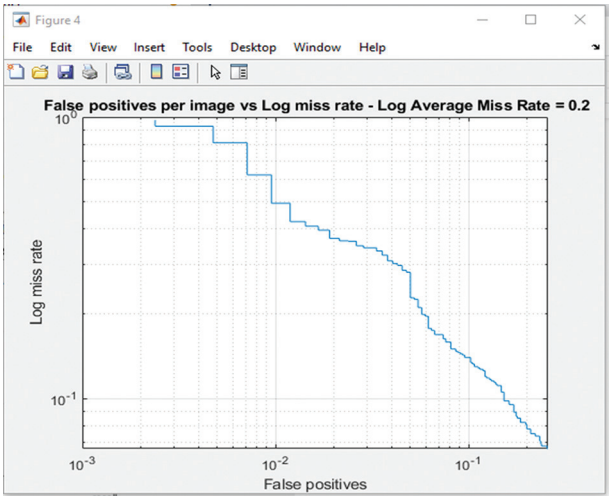


Fig. 7. Representation of the total word detection miss rate in each and every images

4.1. PERFORMANCE EVALUATION

The performance of the anticipated system is estimated with the aid of diverse performance metrics like "accuracy, precision, recall, f1 score, and support", and respective outcomes are discussed.

Table 1. Proposed model

S. No	Models name	Precision	Detection miss rate
1	Hybrid resnet - YOLO	0.907	0.152
2	Hybrid inceptionresnetv2 -YOLO	0.819	0.307
3	Hybrid densenet201 -YOLO	0.831	0.283

4.2 PERFORMANCE COMPARISON

A comparative analysis is provided by comparing the efficiency of the anticipated method with the methodology presented in the existing works. Maximum accuracy of 0.907 is attained from the hybrid resnet –YOLO approach, which shows the efficiency of the proposed method.

Results attained while considering the existing research are given as follows:

Model names resnet50 and Region Proposal Network (RPN) with feature pyramid network to generate bounding boxes for the input image are presented.

Table 2. Existing research work

S. No	Models name	Precision
1	Resnet50	0.769
2	ResNet-101	0.795
3	ResNet-152	0.82

5. CONCLUSION

The current research work focused on the scene-based text detection from any such natural images with the aid of comparative methodology, which includes a combination of YOLO and RESNET as the first case, combined INCEPTIONRESNET AND YOLO and combined DENSENET201 AND YOLO approach as a third case. In the first case, the model comprehends a neural network with Res Net as a feature extractor and YOLO v2 for the classification stage. The choice of these two algorithms is to advance the efficacy of the object detection system by increasing accuracy through detecting even smaller objects in an accurate manner. This approach could improve the speed of detection as it could foresee objects in real time. The research was implemented and tested using MATLAB 2019A, and the packages used were RESNET50, INCEPTIONRESNETV2, and DENSENET201. The effectiveness of the suggested technique was verified by relating accuracy value with the existing research works. Maximum accuracy of 0.907 is attained from the hybrid resnet –YOLO approach, which shows the efficiency of the proposed method.

6. REFERENCES

[1] H. Butt, M. R Raza, M. J. Ramzan, M.J. Ali, M. Haris, “Attention-based CNN-RNN Arabic text recognition from natural scene images", Forecasting, Vol. 3, No. 3, 2021, pp. 520-540.

[2] Y. S. Arafat, M. J. Iqbal, "Urdu-text detection and recognition in natural scene images using deep learning", IEEE Access, Vol. 8, 2020, pp. 96787-96803.

[3] A. N. Joseph, C. Junmin, R. Nersisson, V.G. Mahesh, Z. Zhuang, "Bilingual text detection from natural scene images using faster R-CNN and extended histogram of oriented gradients", Pattern Analysis and Applications, 2022, pp. 1-13.

[4] H. Lin, P. Yang, F. Zhang, "Review of scene text detection and recognition", Archives of Computational Methods in Engineering, Vol. 27, No. 2, 2020, pp. 433-454.

[5] L. T. Akin, T. Jaya, "Improved firefly algorithm-based optimized convolution neural network for scene character recognition Signal", Image and Video Processing, Vol. 15, No. 5, 2021, pp. 885-893.

[6] T. Khan, R. Sarkar, A. F. Mollah, "Deep learning approaches to scene text detection: a comprehensive review", Artificial Intelligence Review, Vol. 54, No. 5, 2021, pp. 3239-3298.

[7] T. Khan, A. F. Mollah, "AUTNT-A component level dataset for text non-text classification and benchmarking with novel script invariant feature descriptors and D-CNN", Multimedia Tools and Applications, Vol. 78, No. 22, 2019, pp. 32159-32186.

[8] Z. Zhong, L. Sun, Q. Huo, "Improved localization accuracy by LocNet for Faster R-CNN based text detection in natural scene images", Pattern Recognition, Vol. 96, 2019, p. 106986.

[9] X. Chen, L. Jin, Y. Zhu, C. Luo, T. Wang, "Text recognition in the wild: A survey", ACM Computing Surveys, Vol. 54, No. 2, 2021, pp. 1-35.

[10] J. Wang, H. Zhang, C. Zhang, W. Yang, L. Shao, J. Wang, "An effective scheme for generating an overview report over a very large corpus of documents", Proceedings of the ACM Symposium on Document Engineering, 2019, pp. 1-11.

[11] D. Kumar, S. Bhalekar, H. Disle, S. Gorule, K. Gotarane, "Automatic Text Summarization Using Local Scoring and Ranking", International Journal for Research Trends and Innovation, Vol. 4, No. 5, 2019.

[12] X. Zhang, X. Gho, C. Tian, "Text detection in natural scene images based on color prior guided MSER", Neurocomputing, Vol. 307, 2018, pp. 61-71.

[13] D. Van Nguyen, S. Lu, S. Tian, N. Ouarti, M. Mokhtari, "A pooling-based scene text proposal technique for scene text reading in the wild", Patten Recognition, Vol. 87, 2019, pp. 118-129.

[14] A. Sain, A. K. Bhunia, P. P. Roy, U. Pal, "Multi-oriented text detection and verification in video frames and scene images", Neurocomputing Vol. 1549, No. 275, 2020, p. 531.

[15] S. Wang, Y. Liu, Z. He, Y. Wang, Z. Tang, "A quadrilateral scene text detector with two-stage network architecture", Pattern Recognition, Vol. 102, 2020, p. 107230.

[16] Y. Liu, H. Chen, C. Shen, T. He, L. Jin, L. Wang, "ABCNet: real-time scene text spotting with adaptive Bezier curve network", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13-19 June 2020.

[17] S. H. Katper, A. R. Gilal, A. Alshanqiti, A. Waqas, A. Alsughayyir, J. & Jaafar, "Deep neural networks combined with STN for multi-oriented text detection and recognition", International Journal of Advanced Computer Science and Applications, Vol. 11, No. 4, 2020.

[18] H. Wang, S. Huang, L. Jin, "Focus on scene text using deep reinforcement learning", Proceedings of the 24th International Conference on Pattern Recognition, Beijing, China, 2018, pp. 3759-3765.

[19] S. Albahli, M. Nawaz, A. Javed, A. Irtaza, "An improved faster-RCNN model for handwritten character recognition", Arabian Journal for Science and Engineering, Vol. 46, No. 9, 2021, pp. 8509-8523.

[20] J. Diaz-Escobar, V. Kober, "Natural scene text detection and segmentation using phase-based regions and character retrieval", Mathematical Problems in Engineering, Vol. 2020, 2020.

[21] X. Liu, G. Meng, C. Pan, "Scene text detection and recognition with advances in deep learning: a sur-

vey", International Journal on Document Analysis and Recognition, Vol. 22, No. 2, 2019, pp. 143-162.

[22] M. Vidhyalakshmi, S. Sudha, "Text detection in natural images with hybrid stroke feature transform and high performance deep Convnet computing", Concurrency and Computation: Practice and Experience, Vol. 33, No. 3, 2021, p. e5271.

[23] Y. Tang, X. Wu, "Scene text detection using superpixel-based stroke feature transform and deep learning based region classification", IEEE Transactions on Multimedia, Vol. 20, No. 9, 2018, pp. 2276-2288.

[24] F. Cong, W. Hu, Q. Huo, L. Guo, "A comparative study of attention-based encoder-decoder approaches to natural scene text recognition", Proceedings of the International Conference on Document Analysis and Recognition, 2019, pp. 916-921.

[25] D. Pandey, B. K. Pandey, S. Wairya, "Hybrid deep neural network with adaptive galactic swarm optimization for text extraction from scene images", Soft Computing, Vol. 25, No. 2, 2021, pp. 1563-1580.

[26] Y. Xu, Y. Wang, W. Zhou, Y. Wang, Z. Yang, X. Bai, "Textfield: Learning a deep direction field for irregular scene text detection", IEEE Transactions on Image Processing, Vol. 28, No. 11, 2019, pp. 5566-5579.

[27] L. T. Sherly, T. Jaya, "An efficient indoor scene character recognition using Bayesian interactive search algorithm-based adaboost-CNN classifier", Neural Computing and Applications, Vol. 33, No. 22, 2021, pp. 15345-15356.

[28] F. Jiang, Z. Hao, X. Liu, "Deep scene text detection with connected component proposals", arXiv:1708.05133, 2017.

[29] A. Ali, M. Pickering, "A hybrid deep neural network for Urdu text recognition in natural images", Proceedings of the IEEE 4th International Conference on Image, Vision and Computing, 2019, pp. 321-325.

[30] N. Gupta, A. S. Jalal, "Text or non-text image classification using fully convolution network (FCN)", Proceedings of the International Conference on Contemporary Computing and Applications, Lucknow, India, 2020, pp. 150-153.

[31] X. Wang, X. Feng, Z. Xia, "Scene video text tracking based on hybrid deep text detection and layout constraint", Neurocomputing, Vol. 363, 2019, pp. 223-235.

Enhancement in Speaker Identification through Feature Fusion using Advanced Dilated Convolution Neural Network

Original Scientific Paper

Hema Kumar Pentapati
Department of Electrical Electronics and Communication Engineering,
GITAM School of Technology
Visakhapatnam-530045, India
hpentapa@gitam.in | https://orcid.org/0000-0002-9373-9132

Sridevi K
Department of Electrical Electronics and Communication Engineering
GITAM School of Technology
Visakhapatnam-530045, India
skataman@gitam.edu | https://orcid.org/0000-0002-6716-6705

Abstract *– There are various challenges in identifying the speakers accurately. The Extraction of discriminative features is a vital task for accurate identification in the speaker identification task. Nowadays, speaker identification is widely investigated using deep learning. The complex and noisy speech data affects the performance of Mel Frequency Cepstral Coefficients (MFCC); hence, MFCC fails to represent the speaker characteristics accurately. In this proposed work, a novel text-independent speaker identification system is developed to enhance the performance by fusion of Log-MelSpectrum and excitation features. The excitation information is obtained due to the vibration of vocal folds, and it is represented using Linear Prediction (LP) residual. The various types of features extracted from the excitation are residual phase, sharpness, Energy of Excitation (EoE), and Strength of Excitation (SoE). The extracted features were processed with the dilated convolution neural network (dilated CNN) to fulfill the identification task. The extensive evaluation showed that the fusion of excitation features gives better results than the existing methods. The accuracy reaches 94.12% for 11 complex classes and 91.34% for 80 speakers, and Equal Error Rate (EER) is reduced to 1.16% for the proposed model. The proposed model is tested with the Librispeech corpus using Matlab 2021b tool, outperforming the existing baseline models. The proposed model achieves an accuracy improvement of 1.34% compared to the baseline system.*

Keywords*: Log-MelSpectrum, MFCC, Speaker Identification, excitation features, Convolution Neural Network(CNN), LP Residual, deep learning, Deep Neural Network*

1. INTRODUCTION

Speaker recognition is one of the most prominent bio-feature based methods. Identifying the speaker depends on the speaker's voice signal [1]. The naturally produced voice is one of the significant human biometric characteristics. The vocal tract, larynx sizes, and various organs responsible for speech generation are unique for each individual [2]. It contains rich information which provides many things about the speaker, including age, gender, and emotions [3,1]. The unique characteristic of speech enables us to focus on speaker recognition. Advanced speaker recognition systems are used in many fields, including online banking, security control, voice access, forensics, and military applications [4,2,5]. Speaker recognition may be mainly classified into two classes: speaker identification (SI) and verification. The SI identifies the unknown speaker's identity by comparing their speech characteristics with those of recognized speakers. The speaker with the highest utterance score is identified as an actual speaker [1]. Thus, it is considered a 1:N match where each utterance is compared against many sets of utterances [4,3]. The speaker verification verifies the claimed identity of the person by accepting /rejecting the speaker [5,3]. It is the 1:1 match where the claimed speaker is compared with their characteristics [3]. Speaker recognition can be divided into text-dependent and Text-independent systems. Text-dependent modules work with the same set of phrases in training and testing [6]. Whereas in text-independent systems, there are no such limitations on text phrases in training and testing.

This paper aims to focus on a text-independent speaker identification system.

Artificial intelligence (AI) approaches can be helpful to Speaker recognition since it is considered a pattern recognition approach [7]. Deep learning addresses complex recognition tasks and effectively helps implement speaker identification systems [8]. Various factors affect the performance of speaker identification. The noisy environment may lead to the misclassification of speakers. The quality of speech signal is severely degraded in forensic applications, leading to difficulty in identifying the correct speaker [9]. Thus, extracting more discriminative features for each individual is the challenge in speaker identification [10]. The researchers found various feature engineering techniques in their study on speaker identification, such as MFCC, LPCC, and several time-domain and spectral-domain features. However, these features cannot represent the unique characteristics under complex and noisy data such as LibriSpeech [11,10]. Moreover, dependence on only a single type of feature set restricts the performance and reliability of the system. In this regard, we propose the model in relation to the researcher's work in this area.

1.1. RATIONALE BEHIND USING EXCITATION FEATURES IN SPEAKER IDENTIFICATION

The MFCC projects the useful speaker's characteristics for speaker identification. However, it is necessary to understand the acoustic cues that describe the detailed speaker characteristics, including different voice qualities and emotions. Then, utilize this information to identify the speaker [12]. Hence, it is required to capture the features describing excitation and the vocal filter in speech production. Thus, excitation features provide supportive information to the frequently used vocal tract features of various speakers. The different methods are well established to describe the vocal tract filter, but the researchers showed less interest in excitation features [12]. The study of [13] demonstrated the methods to capture the excitation features effectively and mentioned the future scope to combine excitation and vocal tract features. This motivates us to combine the derived excitation features from the LP Residual method and vocal tract filter characteristics.

1.2. CONTRIBUTIONS

The significant contributions of the paper are as follows: 1) proposed an efficient feature engineering technique by combining the log-MelSpectrum (vocal tract filter feature) and the various excitation-based features (source features) to improve the identification accuracy. 2) proposed a deep learning-based model referred to as an advanced dilated convolution network that takes the combined feature set of each sample as input to reduce memory consumption and training time. The reason for choosing the dilation convolution network for model construction is it allows us to learn sparse relationships. Primarily, it helps learn speaker-dependent features, as elaborated in the following sections. 3) Evaluate the performance of the system on standard complex LibriSpeech corpus and compare it with baseline methods. 4) Evaluate the proposed model performance by comparing its accuracy with baseline models and analyze the performance with the different number of speakers.

Different sections of the paper are structured as follows. The recent trends in speaker identification research and the performance of the well-established speaker identification methods are presented in section II. Section III presents the analysis of various features and identification models. Section IV presents the dataset corpus, training, and experiments. Section V reports the results of different experiments and discusses the importance of obtained results. Section VI concludes with overall findings and improvements in the proposed model.

2. LITERATURE REVIEW

The human speech signal is a powerful medium for communication. Due to the unique characteristics of each individual, it is widely used in biometric systems such as speaker recognition and speech recognition. A speaker recognition system should handle various variations, such as environmental and background noises, speaker-based and technology-based variations. It extracts the speaker characteristics effectively from the speech signal [1] and it includes the features commonly used, such as Mel-FrequencyCepstralCoefficients (MFCC) and Linear Prediction Cepstral Coefficients (LPCC). The various feature extraction methods, the variations in MFCC, and fusion [7] have been adopted to address complexity and noise-related issues in speaker identification. Significant progress has been made in the first efforts to build the speaker identification system.

The non-parametric approaches called vector quantization [14] and Dynamic time warping have been popular in the research. Later on, researchers shifted towards parametric methods such as Hidden-Markov Model and Gaussian Mixture Model (GMM) for text-independent based systems. The approaches on MFCC and GMM-based text-independent speaker identification systems [15] were popular. The expectation-maximization algorithm was used for parameter estimation and Maximum likelihood(ML) to train the model, while maximum a posteriori(MAP) was used to train the model in [15]. But it requires long utterances, and the MFCC fails to classify correctly under noisy audio data. In some other approaches, the support vector machine(SVM) trained with a large amount of data and Neural Networks were widely used in SI [3].

Jahangir et al. [10] proposed a feature combination of MFCC and time base features to enhance the overall accuracy of their proposed identification system and

fed to the deep neural network (DNN). They obtained an accuracy of 89% on LibriSpeech for 100 speakers. Although they achieved a good performance, the accuracy is still affected due to similar voice patterns. Also, this is due to their simpler deep learning architecture. Chowdhury et al. [9] recently proposed a speaker recognition system in different degraded audio conditions. The study fused MFCC and Linear Predictive Coefficients (LPC) features to classify speakers using their proposed architecture with dilated convolution. The four other datasets were used to evaluate the model. The total match rate was improved by 12% on degraded TIMIT data compared to the existing method but still failed to verify 14% of samples under this dataset.

Hao Meng et al. [16] proposed dilated CNN-based novel architecture for speech emotion recognition. It was evaluated on two popular emotional databases. They picked up the log-MelSpectrum of the speech signal and obtained a notable improvement in accuracy. In [12], the authors stated that it is required to analyze the generation of acoustical cues in the process of speech production. Extracting the features of both the excitation and vocal tract system is required. Various approaches are used to extract excitation information [17,18]. The authors of [13,17] elaborated on the three excitation features based on strength, energy and derived frequency around Glottal Closure Instant (GCI) locations. The authors [17] concluded that the performance of recognition systems might be improved by combining excitation- based features with the features that describe the vocal tract systems.

Ali et al. [19] proposed an SVM-based speaker identification using an Urdu dataset to identify ten speakers. They fused deep belief network features and MFCC features. The model achieved an accuracy of about 92.6%. But only ten speakers were used for evaluation, and each utterance contained only one word. Nainan Kulkarni et al. [20] evaluated 1D CNN, SVM and GMM based on dynamic MFCC features. The 1D CNN-based model achieved a validation accuracy of about 73.25% on the VidTimit dataset. But it was evaluated for only 43 speakers. Samia Abd El-Moneim et al. [21] proposed an LSTM-RNN-based text-independent speaker recognition system to identify five female speakers using the Chinese mandarin dataset. They extracted the log spectrum from the speech signal and used it as the feature set for the model. The experimental results achieved 98.7% with undistorted data. The spectral subtraction method was used to reduce the noise from the speech signal and improve the recognition performance. Although the model improved accuracy, it was evaluated on only five speakers.

Ting Lin et al. [22] proposed a long-term acoustic features-based DNN to identify ten speakers from the LibriSpeech corpus. The authors extracted MFCC and super MFCC features as LTA features and DNN is employed as a classifier. The experimental results achieved the accuracy of 90%. V Srinivas et al. [23] proposed an efficient adaptive fraction bat-based support vector neural network for speaker recognition. They employed frequency-dependent features like spectral kurtosis. It was evaluated on the ELSDSR dataset and achieved an accuracy of 95%. It can be further extended to test on a large dataset since ELSDSR consists of only 22 speakers. Soleymanpour et al. [24] proposed a text-independent speaker identification model based on an artificial neural network and clustering of MFCC. It was evaluated on the ELSDSR dataset and achieved an accuracy of 93%.

3. PROPOSED METHODOLOGY

This part of the paper deals with the detailed methodology of speaker identification and it is shown in Fig. 2. Firstly, the male and female voices from the Librispeech corpus were collected to conduct the experiments [10]. Second, the speech signals from the database are processed, extracting the various useful features to form the fused feature set. This feature set was given as input to the dilated CNN architecture to design the speaker identification model. The developed model was evaluated with the different number of speakers and complex data and compared with existing baseline systems. The next sections will discuss in detail the preferences and parameters considered for the model.

3.1. SPEECH DATA ORGANIZATION

First, split the speech samples into several frames with a window of 25ms. The log MelSpectrum and residual phase MelSpectrum are derived with the exact window sizes. The log-MelSpectrum of each frame comprises 13 coefficients. For the fixed sampling rate of 16KHz and duration of 2s, we have 198 feature frames from each speech sample in the database. Hence, each speech sample is of size 13 x 198 for the log-Mel spectrum. The extraction of excitation features enriches the feature extraction phase. The residual phase MelSpectrum comprises 13 coefficients and two excitation features per frame.

3.2. FEATURE ENGINEERING

The feature set plays a prominent role in the system's overall performance. In deep learning, the speaker identification task depends on the relevant features extracted [10]. Hence, several valuable features are required to extract from the speech samples and these features are appropriate to accomplish the speaker identification mechanism. Therefore, extracting the discriminative features and preparing the feature set through feature engineering is essential. These features are used by the deep learning model and make learning and developing the speaker identification model easier. Thus, the proposed work adopts the extraction and fusion of novel features known as log-MelSpectrum and excitation features to build an accurate system for speaker identification. The functionality and usage of these features are discussed in the following sections.

The vocal tract information can be best depicted by log-MelSpectrum and is used to capture the individual speech characteristics. First, each speech utterance is split into overlapping segments called frames. Each frame of the same length passed through the hamming window. Thus, we obtain the overlapping frames with a step size of 10ms. The number of frames can be calculated as shown in (1).

$$\text{Number of frames N} = \frac{\text{Total size of speech sample - window size}}{\text{Step size}} \quad (1)$$

After frame blocking and windowing of speech signal, the log-Mel spectrum can be extracted by computing the discrete Fourier transform. Then apply the mel-scale filter bank and take the logarithm of the obtained values.

3.3. PROPOSED EXCITATION FEATURES

The speech signal provides acoustical information about the speaker's state. It includes the emotion and health state of the speaker. The speech production system consists of two functional parts [12]. One is excitation produced at the larynx and another is filtering, which is excited by the excitation e(n). Thus, the speech signal s(n) can be produced by exciting the filter v(n) using an excitation.

$$s(n) = e(n) * v(n) \quad (2)$$

The excitation component contains relevant individual speakers' data as the vocal fold vibrations are distinct for a specific individual speaker [25]. The glottal vibrations, the shape of the glottal wave and the SoE of each glottal cycle are the characteristics of the speaker [25]. The variation of physical acoustical movement of excitation leads to changing essential features such as pitch and quality of voice [12]. The excitation features extracted from speech signals are Instantaneous Frequency (Fo), GCI and Glottal Opening Instant (GOI). The glottal flow shows a sudden negative peak in the excitation during the closing phase. This instant where the negative peak occurs is called GCI [12]. This peak is an important excitation to the vocal tract. LP analysis derives the excitation component from the speech signal [13]. Due to the abrupt variations in the movements of the vocal fold, the impulse-like event is produced which gives the prominent characteristic of the excitation [13]. In the LP residual, the instant where the maximum error value occurs refers to the GCI location. In linear prediction, each present sample is predicted from the past *l* samples [25]. This mathematical representation of prediction is given by (3).

$$\hat{s(n)} = -\sum_{k=1}^{l} a_k \cdot s(n-k) \quad (3)$$

The obtained difference between true and predicted sample is said to be the error signal as given in (4).

$$e(n) = s(n) - \hat{s}(n) = s(n) + \sum_{k=1}^{l} a_k . s(n-k) \quad (4)$$

Where a_k indicates prediction coefficients and l indicates the order of prediction. The obtained LP residual signal or glottal excitation suppresses vocal tract characteristics. The glottal excitation phase, or the residual phase, is unique for each individual, and thus the feature is helpful in recognition applications [18]. This feature can be extracted by taking the cosine phase of an analytic signal using the Hilbert transform $e_h(n)$. The analytic signal $e_a(n)$, cosine phase and Hilbert envelope $h_e(n)$ representations are given in (5), (6), (7), respectively.

$$e_a(n) = e(n) + je_h(n) \quad (5)$$

$$\text{cosinephas} = \frac{e(n)}{h_e(n)} \quad (6)$$

$$h_e(n) = \sqrt{(e^2(n) + e_h^2(n))} \quad (7)$$

The detailed flow chart of computation residual phase melspectrum is given in Fig. 1. Various Glottal Closure Instant(GCI) parameters which reflects the residual signal [13], [17] are discussed below.

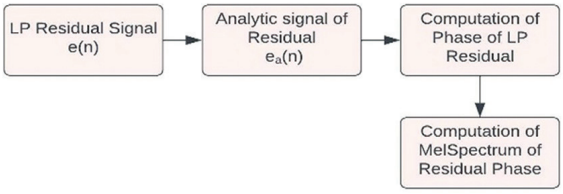


Fig. 1. The Block diagram for computation of Residual Phase based MelSpectrum

- **Strength of Excitation**: The strength is given by the substantial residual error at the prominent excitation and it varies by the rate of glottal closure [25]. The Strength of Excitation is referred to the magnitude of the impulse-like excitation [13]. It is computed as the amplitude of excitation at GCI. From the LP Residual signal, the SoE can be given by (8).

$$SoE = |h_e(n)| \quad \text{for all N} \quad (8)$$

- **Energy of Excitation:** LP residual is closely related to the excitation of the speech sample. EoE given in (9) reflects the vocal effort. This can be calculated from the Hilbert envelope around each GCI of LP Residual.

$$EoE = \frac{1}{(2k+1)} \sum_{i=0}^{2k} h_e^2(i) \quad (9)$$

- **Sharpness of Excitation:** The sharpness of excitation can be computed as the ratio of standard deviation (ŋ) to the mean (μ) [18]. The ŋ and μ are computed for the Hilbert envelope of excitation at the closure instance. The mathematical expression is given in (10).

$$sharpness = \frac{\eta}{\mu} \quad (10)$$

3.4. FUSION OF LOG-MELSPECTRUM AND EXCITATION FEATURES

The four sets of features are obtained as given in (11). The first vector represents the log-Mel spectrum of the

speech sample with 13 coefficients and the second vector represents the LP residual phase MelSpectrum with 13 coefficients for each frame. The third and fourth feature vector represents the sharpness and strength of the excitation. Thus, it forms the 28 features for each frame of the speech segment. Finally, all the feature vectors are fused to create a 28 x 198 size matrix, shown in (11) and fed to the dilated convolution neural network to accomplish the speaker identification task.

$$SS1 = \begin{bmatrix} M_{11} & M_{12} & \dots & M_{1n} \\ M_{21} & M_{22} & \dots & M_{2n} \\ \vdots & \vdots & \dots & \vdots \\ M_{m1} & M_{m2} & \dots & M_{mn} \\ R_{11} & R_{12} & \dots & R_{1n} \\ R_{21} & R_{22} & \dots & R_{2n} \\ \vdots & \vdots & \dots & \vdots \\ R_{r1} & R_{r2} & \dots & R_{rn} \\ P_1 & P_2 & \dots & P_n \\ S_1 & S_2 & \dots & S_n \end{bmatrix} \quad (11)$$

Where:

SS1= first speech sample

M= log-melspectrum of speech sample 'SS1'

m= number of coefficients of M

R = melspectrum of residual phase for speech

r = number of coefficients of R

P = Sharpness of Excitation

S = Strength of Excitation

n = Number of frames

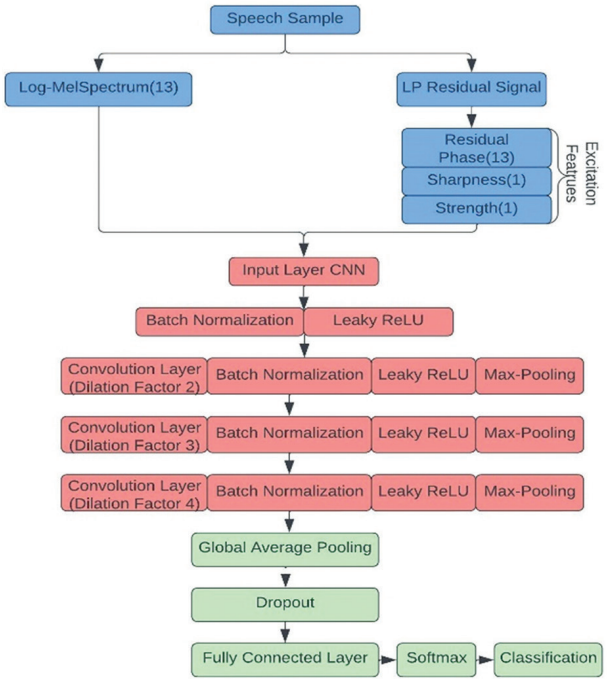


Fig. 2. Proposed Dilated Convolution Network with Excitation Features

3.5. DILATED CONVOLUTION

The learning capability of CNN depends on the design and depth of the convolution layers. Each layer learns and then gives the transformed data to the layers in the network and has sparse connectivity [26]. Since the speech signal changes continuously due to articulatory movements, it is considered non-stationary. So, working on short time segments of the speech signal is desired to get stable speech characteristics [9]. We used dilated convolution instead of traditional convolution due to the following reasons. The pooling layer imposes the data loss in the network by reducing the dimensionality. Without imposing a computational burden, dilated convolution increases the receptive field extensively. It learns the sparse relationship between the features. For example, a one-dimensional convolution kernel of size 4 x 1learns the sparse filter of 7 x1 with a dilation factor of 2. The dilation convolution kernel populated alternative positions when the dilation factor is two and this reduces the data loss. Also, learning the sparse relationship between the features is helpful in degraded speech as the frequency bands are degraded in dense regions [9]. Hence, dilated convolution prevents it from learning local dense regions due to sparse filters.

4. DATASET AND EXPERIMENTS

4.1. SPEECH DATASET

The LibriSpeech dataset in the experiments is used for evaluating certain aspects of speaker identification. The class or category with speech utterances of various speakers is named an unknown class. The experiments on the LibriSpeech corpus aim to analyze the proposed algorithm under the different number of speakers and different classes, including the unknown category. The publicly available LibriSpeech corpus provides the clean speech data of nearly 100 speakers. Each speaker's training and testing sets are disjoint to make it text-independent. It consists of 250 utterances by each speaker in the English language. We split the available data into the training and test with 80% and 20% proportions, respectively.

4.2. TRAINING

The proposed network schemes are implemented using Matlab 2021b and the deep learning toolbox is used to build the network. The model is trained for 25 epochs for all the experiments with a minibatch of 128. The number of iterations is changed for the different number of speakers. The validation data is used to tune hyper-parameters. The proposed dilated convolution network shown in Fig. 2 is trained using Adam optimization with an initial learning rate of 0.00001. The convolution kernel size and learning rate are selected based on a trial and error approach. The weights are updated throughout the learning process.

The tendency of the gradient of the sigmoid activation function is to shrink at every step throughout the depth of the network [1]. It could lead to a vanishing gradient problem. The ReLU creates the dead neuron

when the gradient becomes zero. Hence, in our experiments, the Leaky ReLU activation function was used after each convolution layer.

4.3. EXPERIMENTAL SETUP

The experimental overview of the proposed feature fusion and dilated convolution network is discussed in this section. Many experiments are conducted to measure the performance of the developed system. The following sets of experiments on the proposed model are performed to evaluate and compare its performance with baseline systems.

Firstly, the proposed excitation features are extracted from speech utterances to identify 11 classes, including one unknown class. The excitation features such as residual phase, sharpness of excitation, EoE and SoE are combined with log-MelSpectrum and given as input to the dilation convolution neural network for the classification of speakers. The four sets of features (four experiments) are given as input to Dilated CNN separately to evaluate the performance.

In the second set of experiments, the performance of the different number of speakers (say 100, 70, 56, 30,10) for the proposed model is evaluated. This setting allows us to evaluate the classification performance in classifying the large number of speakers.

Lastly, our model is compared with the baseline system and evaluates the performance of five female speakers. The improvement of the proposed feature fusion-based system is compared with several existing baseline- models by computing accuracy, EER and ROC.

5. EXPERIMENTAL RESULTS

5.1. EVALUATION METRICS

We compare the Accuracy, Equal Error Rate (EER), Receiver Operating Characteristics (ROC) curves and confusion matrices for both baseline and proposed systems of a varying number of speakers and various classes of speakers. These are chosen as performance metrics of the proposed method. We also evaluated and reported the accuracies of male and female speaker identification for five speakers of the proposed system. Accuracy is computed as the number of samples that are correctly identified divided by the total samples in the test subset. Equation (12) gives the mathematical expression for accuracy.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{12}$$

Equal Error Rate (EER) is computed by finding the same value of FAR and FRR. The system with low EER indicates good performance.

ROC is used to analyze the classifier's performance for each speaker class and gives a precise performance overview by plotting curves of each class. The Area Under the Curve(AUC) measures how efficiently the architecture separates classes [4]. A value of AUC nearly equal to one indicates the good conduct of the classifier, while this value of less than 0.5 indicates poor performance [8,10] . To understand and plot the ROC, the mathematical expressions of true positive (TP) and negative (TN) rates are useful as shown in (13), (14).

$$TPR = \frac{TP}{TP + FN} \tag{13}$$

$$FPR = \frac{FP}{FP + TN} \tag{14}$$

5.2. MODEL RESULTS

In this section, the results of various experiments are discussed as follows. We trained the dilated convolution neural network using Librispeech corpus with an available speaker training set to obtain the results and analyze the performance. Then, evaluate the trained model for identification of speakers' test data. First, the results of the proposed model with excitation and log-MelSpectrum with various classes, including unknown class are obtained. The results of the proposed method with the different number of speakers are also obtained. Lastly, the results of the comparison of the proposed network with baseline systems and with respect to gender are obtained. All these results are discussed in the following sections.

In the first set of experiments, the proposed model's overall accuracies and Equal Error rate with various excitation features are evaluated and compared with the previously developed models in the research. The reported accuracies ranged from 76.61% to 94.12% with the librispeech dataset shown in Table 1. As shown in Table 1, the proposed deep learning model with 28 features outperformed the existing models by obtaining an accuracy of 94.12% and an EER of 1.16% for speaker identification. In the other three methods, the advanced dilated convolution network, when trained and tested on the fusion of log-MelSpectrum and excitation features such as sharpness of excitation and EoE, obtained the highest accuracy of about 91.57% with EER 1.16%. The plot of accuracies of various methods is shown in Fig. 3.

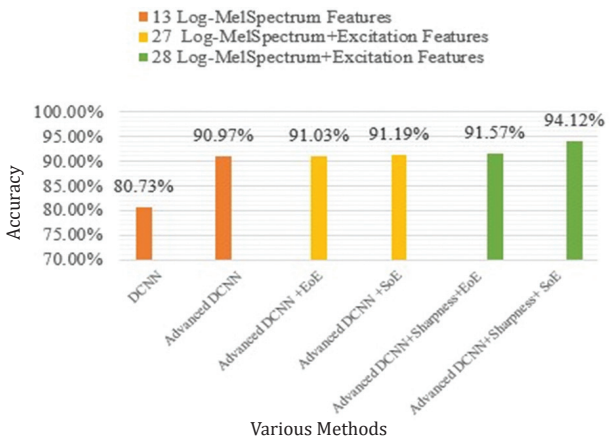


Fig. 3. Plot of Accuracies for various methods

Table 1. The performance of the identification models with 11 classes of speakers (including unknown class) using LibriSpeech data

Type of Model	Features	Dilation factor	Accuracy (%)	EER
CNN	MFCC	1	76.61%	5.90 %
Dilated CNN	Log-MelSpectrum	2	80.73%	4.82 %
AlexNet	Log-MelSpectrum	1	82.40%	-
Advanced Dilated CNN	Log-MelSpectrum	4	90.97%	2.38%
Advanced Dilated CNN	Log-MelSpectrum + LP Residual Phase MelSpectrum+EoE	4	91.03%	2.28%
Advanced Dilated CNN	Log-MelSpectrum + LP Residual Phase MelSpectrum + Sharpness	4	91.19%	2.27%
Advanced Dilated CNN(Proposed)	Log-MelSpectrum + LP Residual Phase MelSpectrum + Sharpness of Excitation + EoE	4	91.57%	1.16%
Advanced Dilated CNN (Proposed)	**Log-MelSpectrum + LP Residual Phase MelSpectrum + Sharpness of Excitation + SoE**	**4**	**94.12%**	**1.16%**

Across the eight experiments in Table 1, the proposed method correctly identifies an extra 2.59% of the test samples over advanced dilated CNN with sharpness and EoE. Also, an additional 3.15% of the test samples were over the Log-MelSpectrum model. Fig. 4 shows the confusion matrix that analyzes the number of test samples over false acceptance and rejection for the proposed method. The proposed system also improved the EER from 5.90% to 1.16%.

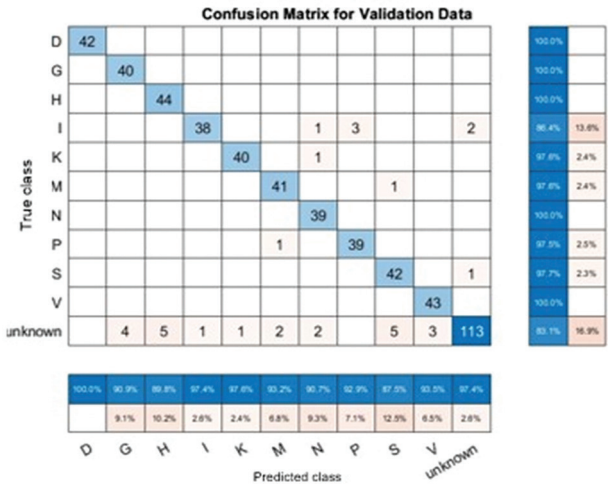


Fig. 4. Confusion matrix for validation data of the proposed method

The efficiency of the proposed method is revealed by comparing its performance with baseline systems. The experimental results of the baseline and proposed methods are shown in Table 2. The ratio of train and test sets is the same (80:20) for all the methods mentioned in Table 2. In general, the performance of the classification model decreases when the number of speakers increases. As presented in the Table 2, the proposed excitation features based dilated convolution network outperformed the baseline methods. The plot of training and validation accuracy is given in Fig. 5.

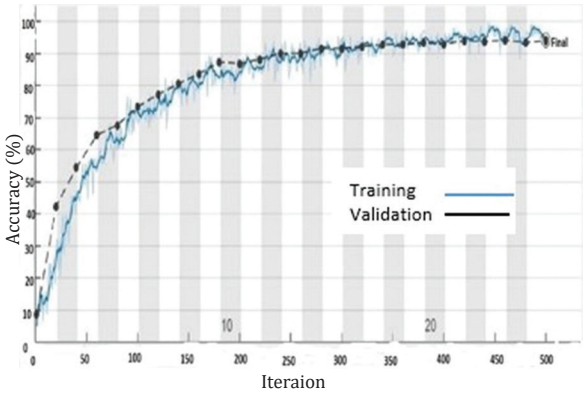


Fig. 5. Accuracy Vs Iterations plot of proposed method

Table 2. The performance comparison of proposed model with existing baseline models

Literature	Methodology	Features	Database	Speakers	Accuracy
Jahangir et.al. [10]	DNN	MFCC + MFCCT	Libri Speech	50	90 %
Ting Lin et.al. [22]	DNN	MFCC and Super MFCC	Libri Speech	10	90 %
S. Chakraborthy et. al. [27]	GMM-UBM	MFCC	Libri Speech	40	86 %
Pentapati et.al.[6]	Dilated CNN (dilation factor 2)	Log-Mel Spectrum	Libri Speech	11	80.6 %
Proposed	**Dilated CNN (dilation factor 4)**	**Excitation features and Log-Mel Spectrum**	**Libri Speech**	**80** **11 (with Unknown)**	**91.34 %** **94.12 %**

The baseline method[10] obtained 90% and 89% accuracy with 50 and 100 speakers, respectively. Compared with the baseline method [10], the proposed model shows improvement in performance by obtaining an accuracy of 91.34% with 80 speakers. To confirm this, the number of speakers gradually increased from 10 to 80. The experiments are conducted to evaluate the performance of each set of speakers as shown in Table 3. In all these four experiments, the proposed method outperformed the baseline method.

Table 3. Performance comparison of the proposed method with different number of speakers

Number of Speakers	Accuracy (%)	EER	Training error (%)	Validation Error (%)
10	97.61%	1.02%	0%	2.3923%
26	96.26%	1.16%	0.1348%	3.7409%
56	93.54%	2.08%	0.6827%	6.4551%
80	91.34%	3.11%	0.9093%	8.6585%

The comparative analysis of the proposed and baseline method [10] can be observed in Fig. 6. The accuracy of the proposed system is slightly lower than that of the baseline method [10] by 1.9% for ten speakers. Despite that lower value, the accuracy of our approach is much better on 26, 56 and 80 speakers compared to the method proposed in [10].

Across the several methods in Table 2, the proposed method improves the accuracy by 7.61% compared to [22] baseline method for 10 speakers. The accuracy of the proposed method further improves by nearly 8% and 14% compared to the baseline methods [27] and [6] respectively. The proposed work does not introduce computational complexity since extracting the log-melspectrum and excitation features have been done offline. Thus, the proposed model achieved improved accuracy with training time of 115 seconds per epoch.

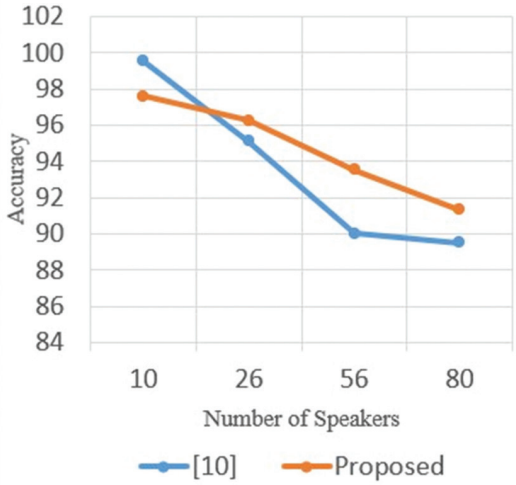


Fig. 6. Comparative analysis of proposed and baseline method

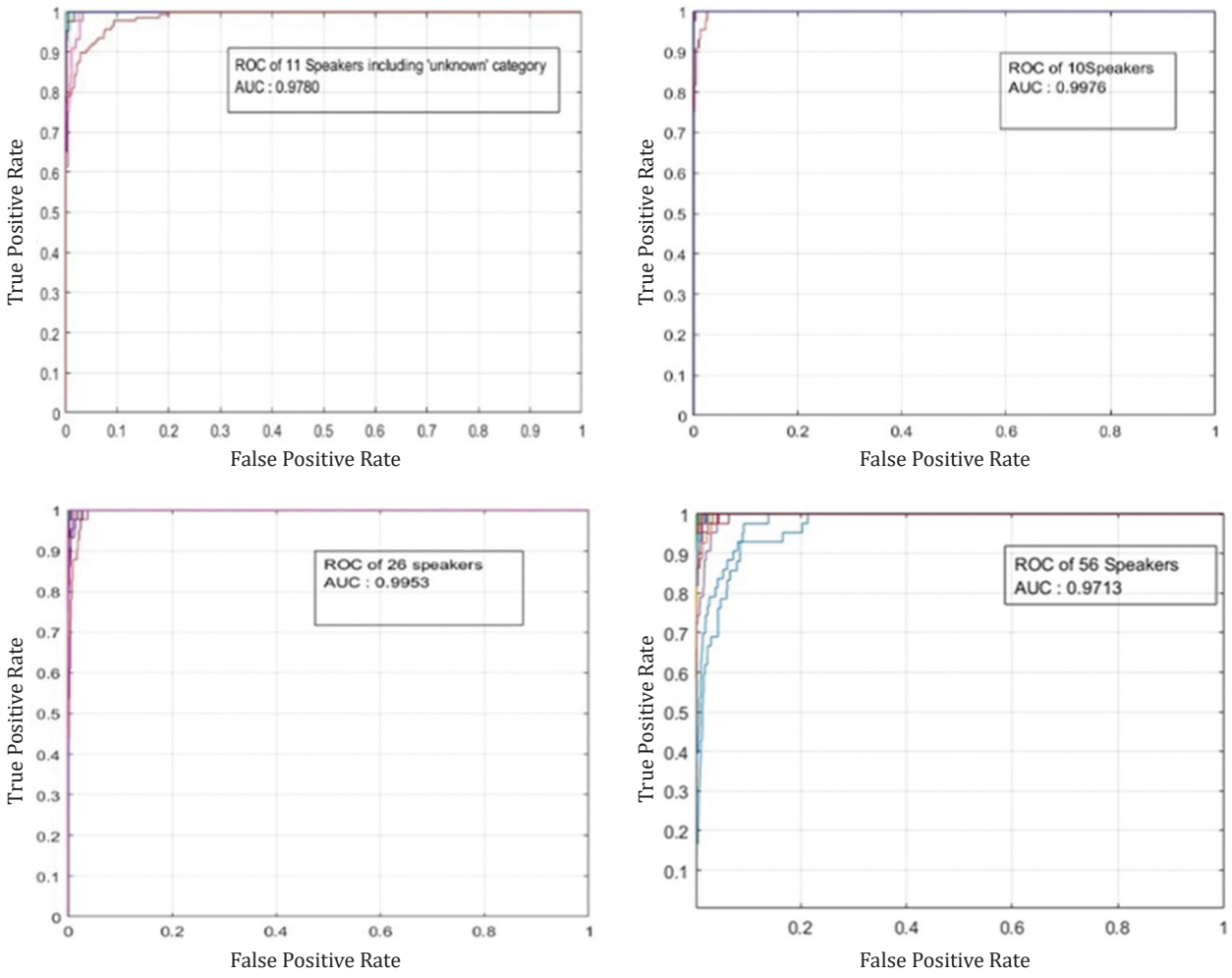


Fig. 7. ROC and AUC of the proposed method for different number of speakers

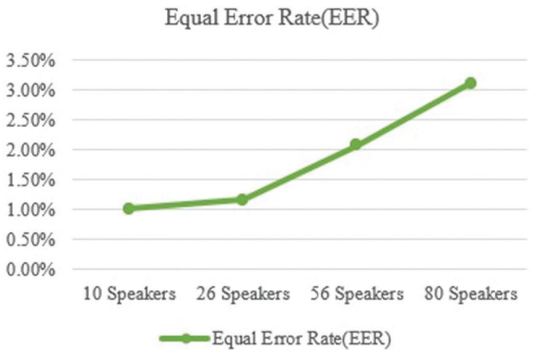


Fig. 8. Variation of EER for different number of speakers

In Table 2, Jahangir et al. [10] were achieved an accuracy of 90% for 50 speakers since an efficient MFCCT-based feature fusion was employed. Although they used novel feature fusion, the model didn't reach the desired accuracy due to simpler deep learning architecture. Thereby, misclassification happens due to similar voice patterns. Ting Lin et al. [22] were also extracted novel LTA-based features along with MFCC. They used simple DNN and tested only ten speakers on the Librispeech corpus and achieved 90% accuracy. In contrast, the proposed method used an advanced dilated convolution network with feature fusion of excitation features. Hence, enhancing the accuracy to 91.34% for 80 speakers. In Table 1, the previous model with advanced dilated CNN achieved only 90.97% accuracy since only log-melspectrum was used in the feature extraction phase.

Across the four experiments shown in Table 3, the proposed model with 10 speakers correctly identifies an additional 1.35% of the test samples over 26 speakers, 4.07% over 56 speakers, 6.27% over 80 speakers. The Fig. 7 shows ROC and AUC plots of proposed method with 10,11,26 and 56 speakers. This shows AUC of 10 speakers achieved the highest value of 0.9976 when compared with other sets. As shown in Fig.8, It also improved the EER from 3.11% to 1.02%.

The log-spectrum based baseline method [21] obtained the highest accuracy of 98.7% with five female speakers on the Chinese mandarin database. The Five female speakers with only 100 utterances per speaker, a frame size of 256 samples and 128 features were considered in the model proposed by [21]. In contrast, the proposed model considered the 250 utterances of each speaker with a frame size of 400 samples.

The Librispeech and Chinese Mandarin datasets considered the undistorted samples with the same sampling rates. Also, the training and testing sets are disjoint and collected under the same environmental conditions in both datasets. As the number of utterances to classify increases, the model performance decreases. To confirm the improvement in the proposed design, the experiment is conducted to test the proposed method with five male and five female speakers. As shown in Table 4, the proposed model obtained an accuracy of 99.52% for female speakers. It correctly identifies an additional 0.82% of the test samples over the baseline [21].

Table 4. Performance comparison of the proposed method and baseline method [21] for five Speakers

Method	Dataset	Gender	Features	Frame size in samples	Accuracy in %
Samia Abd El-Moneiet. al. [21]	Chinese mandarin Corpus	Female	Log-Spectrum	256	98.7%
Proposed	Libri Speech Corpus	Female	Log-Melspectrum + Residual Phase + Sharpness +SoE	400	99.52%
Proposed	Libri Speech Corpus	Male	Log-Melspectrum + Residual Phase + Sharpness + SoE	400	99.49%

6. CONCLUSION

In this paper, the efficient excitation-based features are considered and combined with Log-MelSpectrum of the speech sample. The excitation-based features such as the Residual phase, Sharpness of excitation, EoE and SoE are extracted using Linear prediction analysis. Extensive experiments are conducted on the proposed advanced dilated convolution network with excitation features using the LibriSpeech corpus. It showed that the accuracy and EER are 94.12% and 1.16%, respectively, with 11 complex classes including the unknown category. For 80 speakers, the proposed method achieved an accuracy of 91.34% and the accuracy reached 99.52% for five female speakers. The experimental results proved that the proposed speaker identification system is effective with respect to the number of speakers and computational complexity compared to existing baseline systems. In the future, we aim to develop a speaker identification system with a more profound architecture to handle a much more extensive database and reduce the misclassified samples. The hyper-parameter tuning process can be enhanced to improve the system's performance further.

7. REFERENCES:

[1] F. Ye, J. Yang, "A deep neural network model for speaker identification", Applied Sciences, Vol. 11, No. 8, 2021, pp. 1-18.

[2] X. Wang, F. Xue, W. Wang, A. Liu, "A network model of speaker identification with new feature extraction methods and asymmetric BLSTM", Neurocomputing, Vol. 403, 2020, pp. 167-181.

[3] S. Farsiani, H. Izadkhah, S. Lotfi, "An optimum end-to-end text-independent speaker identification system using convolutional neural network", Computers and Electrical Engineering, Vol. 100, 2022, p. 107882.

[4] M. M. Kabir, M. F. Mridha, J. Shin, I. Jahan, A. Q. Ohi, "A Survey of Speaker Recognition: Fundamental Theories, Recognition Methods and Opportunities", IEEE Access, Vol. 9, 2021, pp. 79236-79263.

[5] S. Hourri, J. Kharroubi, "A deep learning approach for speaker recognition", International Journal of Speech Technology, Vol. 23, No. 1, 2020, pp. 123-131.

[6] H. K. Pentapati, Sridevi K, "Dilated Convolution and MelSpectrum for Speaker Identification using Simple Deep Network", Proceedings of the 8th International Conference on Advanced Computing and Communication Systems, Coimbatore, India, 25-26 March 2022, pp. 1169-1173.

[7] S. S. Tirumala, S. R. Shahamiri, A. S. Garhwal, R. Wang, "Speaker identification features extraction methods: A systematic review", Expert Systems with Applications, Vol. 90, 2017, pp. 250-271.

[8] R. Jahangir, Y. W. Teh, F. Hanif, G. Mujtaba, "Deep learning approaches for speech emotion recognition: state of the art and research challenges", Multimedia Tools and Applications, Vol. 80, No. 16, 2021, pp. 23745-23812.

[9] A. Chowdhury, A. Ross, "Fusing MFCC and LPC Features Using 1D Triplet CNN for Speaker Recognition in Severely Degraded Audio Signals", IEEE Transactions on Information Forensics and Security, Vol. 15, 2020, pp. 1616-1629.

[10] R. Jahangir et al., "Text-Independent Speaker Identification through Feature Fusion and Deep Neural Network", IEEE Access, Vol. 8, 2020, pp. 32187-32202.

[11] V. Panayotov, G. Chen, D. Povey, S. Khudanpur, "Librispeech: An ASR corpus based on public domain audio books", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 19-24 April 2015, pp. 5206-5210.

[12] S. R. Kadiri, P. Alku, B. Yegnanarayana, "Extraction and Utilization of Excitation Information of Speech: A Review", Proceedings of the IEEE, Vol. 109, No. 12, 2021, pp. 1920-1941.

[13] S. R. Kadiri, P. Gangamohan, S. V. Gangashetty, P. Alku, B. Yegnanarayana, "Excitation Features of Speech for Emotion Recognition Using Neutral Speech as Reference", Circuits, Systems and Signal Processing, Vol. 39, No. 9, 2020, pp. 4459-4481.

[14] M. K. Gill, R. Kaur, J. Kaur, "Vector Quantization based Speaker Identification", International Journal of Computer Applications, Vol. 4, No. 2, 2010, pp. 1-4.

[15] D. A. Reynolds, T. F. Quatieri, R. B. Dunn, "Speaker verification using adapted Gaussian mixture models", Digital Signal Processing, Vol. 10, No. 1, 2000, pp. 19-41.

[16] H. Meng, T. Yan, F. Yuan, H. Wei, "Speech Emotion Recognition from 3D Log-Mel Spectrograms with Deep Learning Network", IEEE Access, Vol. 7, 2019, pp. 125868-125881

[17] S. R. Kadiri, P. Alku, "Excitation Features of Speech for Speaker-Specific Emotion Detection", IEEE Access, Vol. 8, 2020, pp. 60382-60391.

[18] T. Thomas, Spoorthy, N. V. Sobhana, S. G. Koolagudi, "Speaker Recognition in Emotional Environment using Excitation Features", Proceedings of the Third International Conference on Advances in Electronics, Computers and Communications, Bengaluru, India, 11-12 December 2020.

[19] H. Ali, S. N. Tran, E. Benetos, S. Artur, A. Garcez, "Speaker recognition with hybrid features from a deep belief network", Neural Computing and Applications, Vol. 29, No. 6, 2018, pp. 13-19.

[20] S. Nainan, V. Kulkarni, "Enhancement in speaker recognition for optimized speech features using GMM, SVM and 1-D CNN", International Journal of Speech Technology, Vol. 24, No. 4, 2021, pp. 809-822.

[21] S. A. El-Moneim, M. A. Nassar, M. I. Dessouky, N. A. Ismail, A. S. El-Fishawy, and F. E. Abd El-Samie, "Text-independent speaker recognition using LSTM-RNN and speech enhancement", Multimedia Tools and Applications, Vol. 79, No. 33-34, 2020, pp. 24013-24028.

[22] T. Lin, Y. Zhang, "Speaker recognition based on long-term acoustic features with analysis sparse representation", IEEE Access, Vol. 7, 2019, pp. 87439-87447.

[23] Vasamsetti Srinivas, Ch Santhi Rani, "Optimization-Based Support Vector Neural network for Speaker Recognition", The Computer Journal, Vol. 63, No. 1, 2020, pp. 151-167.

[24] M. Soleymanpour, H. Marvi, "Text-independent speaker identification based on selection of the most similar feature vectors", International Journal of Speech Technology, Vol. 20, No. 1, 2017, pp. 99-108.

[25] S. R. Mahadeva Prasanna, C. S. Gupta, and B. Yegnanarayana, "Extraction of speaker-specific excitation information from linear prediction residual of speech", Speech Communication, Vol. 48, No. 10, 2006, pp. 1243-1261.

[26] Z. Liu, Z. Wu, T. Li, J. Li, C. Shen, "GMM and CNN Hybrid Method for Short Utterance Speaker Recognition", IEEE Transactions on Industrial Informatics, Vol. 14, No. 7, 2018, pp. 3244-3252.

[27] S. Chakraborty, R. Parekh, "An improved approach to open set text-independent speaker identification (OSTI-SI)", Proceedings of the Third International Conference on Research in Computational Intelligence and Communication Networks, 3-5 November 2017, pp. 51-56.

Fuzzy controller hardware implementation for an EV's HESS energy management

Original Scientific Paper

JBARI Hatim
Moulay Ismaïl University,
National High School of Arts and Crafts of Meknes (ENSAM-Meknès),
Team of Modelling, Information Processing and Control Systems (MPICS)
B.P. 15290 EL Mansour Meknes 50500, Meknès, Morocco
ha.jbari@edu.umi.ac.ma

ASKOUR Rachid
Moulay Ismaïl University,
National High School of Arts and Crafts of Meknes (ENSAM-Meknès),
Team of Modelling, Information Processing and Control Systems (MPICS)
B.P. 15290 EL Mansour Meknes 50500, Meknès, Morocco
r.askour@umi.ac.ma

BOUOULID IDRISSI Badr
Moulay Ismaïl University,
National High School of Arts and Crafts of Meknes (ENSAM-Meknès),
Team of Modelling, Information Processing and Control Systems (MPICS)
B.P. 15290 EL Mansour Meknes 50500, Meknès, Morocco
b.bououlid@umi.ac.ma

Abstract *– The recent technological advances related to embedded systems, and the increased requirements of the Electric Vehicle (EV) industry, lead to the evolution of design and validation methodologies applied to complex systems, in order to design a product that respects the requirements defined according to its performance, safety, and reliability. This research paper presents a design and validation methodology, based on a hardware-in-the-loop (HIL) approach, including a software platform represented by Matlab/Simulink and a real-time STM32 microcontroller used as a hardware platform. The objective of this work is to evaluate and validate an Energy Management System (EMS) based on Fuzzy Logic Controller (FLC), developed in C code and embedded on an STM32 microcontroller. The developed EMS is designed to control, in real-time, the energy flow in a hybrid energy storage system (HESS), designed in an active topology, made of a Li-ion battery and Super-Capacitors (SC). The proposed HESS model was organized using the Energetic Macroscopic Representation (EMR) and constructed on Matlab/Simulink software platform. The evaluation and validation of the developed algorithm were performed by comparing the HIL and simulation results under the New European Driving Cycle (NEDC).*

Keywords*: Fuzzy Logic Controller, Hardware-In-the-Loop, Hybrid Energy Storage System, STM32 microcontroller*

1. INTRODUCTION

All around the world, the transportation sector is the main contributor to urban pollution [1]. Public health studies in different countries have demonstrated that the increase in air pollution concentrations is associated with an increase in the mortality rate, particularly in urban areas. The electric vehicle (EV) seems to be an efficient solution to overcome this public health problem, a technology that is encouraged by government policies, to support and generalize this means of transport that ensures zero-CO2 emissions. Meanwhile, several manufacturers have focused a large part of their development program on this industrial segment, with the aim of accelerating the total insertion of the EV in the car fleet during the next years.

Therefore, EVs are currently facing several challenges, especially those related to the range and lifetime of the main source of energy provided by batteries [2]. In fact, industrial companies and scientists are confronted with the challenge of finding solutions to EV's

limitations, taking into account the constraints related to its performance required by the users. In this case, the adoption of a *HESS*-based architecture with the integration of *SCs* [3,4], and the implementation of an *EMS* ensuring efficient energy management between energy sources, allowing optimization of the efficiency of the battery lifespan and the vehicle autonomy.

According to the literature, *EMS* is a key element to optimize the EV battery lifespan and autonomy. Indeed, there are three main approaches to energy management algorithms: *EMS*-based optimization algorithms, algorithms based on artificial intelligence, and rule-based algorithms [5,6,7]. The authors in [8,9] propose a rule-based *EMS*, applied to the management of the energy flow in a *HESS* in a full-active topology [10], ensuring the supply of the EV's powertrain. The *EMS* presented in this work is based on a Fuzzy Logic Controller (FLC), which supports two inputs represented by the SC's state of charge and the power ratio related to the maximum battery power. Thus, the objective of this study is to control the *SCs* as an energy buffer, and to support the power pulses, in order to improve the battery performance and optimize its lifespan. Another work presented in [11], uses a fuzzy controller that supports the same inputs mentioned previously. This work ensures the same objectives with the criterion of minimizing the Root Mean Square (RMS) of the battery current. In [12], the authors present a work that is based on two fuzzy controllers, managing the distribution of the power demanded by the powertrain, between the battery and the *SCs* in a *HESS*'s full-active topology. The first fuzzy controller supports the estimated source's states of charge using the Kalman filter, the energy sources, and the required power. The second one determines the power required by the battery, where the inputs are based on the power calculated by the first block of the fuzzy controller, and the power estimated via the Kalman filter. The objective of this work is to minimize the power pulses that the battery should support.

On the other hand, the evolution of EVs has to take into account constraints related to costs and development times, in order to optimize the time to market of this technology. Hence, the rise of new methodologies and platforms, designed to satisfy these requirements, has been able to optimize the design of complex systems such as the EV, for instance, the Hardware-In-the-Loop (*HIL*) applied to *EMS*s. Thus, some researchers in [13] have proposed a *HIL* architecture for the emulation and real-time control of a boost converter, designed in the PSIM software, with a PI cascaded controller on the STM32 microcontroller, to regulate the output voltage. The objective of this work is to evaluate the control model through the comparison of the *HIL* results, and the simulation model established in the PSIM software. In [14] the authors present another *HIL* architecture based on a real-time simulator of an EV, built using the Simulink platform, and a vehicle control unit (VCU) designed on a hardware platform. The EV model is made of a BLDC engine with a motor control unit (MCU), and a *HESS* powering the vehicle. The VCU is connected to the simulator through inputs/outputs (CAN bus, RS232, analog/digital I/O) using a DSP interface. Another *HIL* simulation platform applied to the electric vehicle powertrain used to emulate the driving cycle is presented in [15]. The proposed architecture is based on two coupled electrical machines, the first one is an induction machine that emulates the vehicle's traction machine and the second one is a DC machine that provides the load simulation. A dSPACE DS1102 board is used as an interface between the control established on the Simulink platform, and the power converter assures the control of both machines. The authors in [16,17] present a solution of *HIL* that aims at the validation of an FLC-based *EMS* algorithm, designed for the control of the energy flow in a *HESS* in a full-active topology, built on the Simulink platform. The Atmega microcontroller was used in the first work, while the FPGA hardware platform was used in the second work for the same architecture.

Thus, this paper presents a *HIL* architecture based on an active-topology *HESS* model, organized using the Energetic Macroscopic Representation (EMR) and implemented on the Matlab/Simulink software platform. The proposed *HESS* is designed to supply an EV to ensure a maximum speed of 120km/h with a propulsion power of 70kW. The electric vehicle can provide a total range of 150km. The *HESS*'s *EMS* is used for the management and optimization of the energy flow using an FLC. The proposed FLC was developed in C language, and embedded in the STM32 microcontroller. The STM32 is connected in a closed loop with the *HESS* simulation model, to control in real-time, the contribution of both sources: *SCs* and battery.

In a previous work [16], an architecture of a *HESS* is presented. The proposed *EMS* based on FLC is implemented on an Atmega microcontroller. The output of the used FLC is defined by a simplified method for computing the Center of Gravity (*CoG*) abscissa of the resulting polygon. The FLC is based on a 3x3 rule base matrix. In this work, the rule-based *EMS*-FLC is embedded in the STM32 low-power microcontroller, which is a part of the family of ARM-based 32-bit microcontrollers and has some major hardware advantages over the Atmega microcontrollers 8-bit. It's also popular and widely used in the field of EVs. Moreover, the *CoG* defuzzification method is based on a precise value depending on the fuzzy set's center of gravity of the resulting area. In addition, the proposed FLC is based on a 4x5 rule base matrix.

Indeed, the objective of this work is to evaluate on the one hand, the performance of the developed *EMS*, on the other hand to validate the algorithm of the *EMS*-FLC, via the comparison between the results obtained from the *HIL* and the simulation.

The rest of this paper is organized as follows: section 2 presents the nomenclature; the description of the model and the control of the developed *HESS* is detailed in section 3. Section 4 presents the proposed *HESS*. Section 5 details the *EMS* implementation methodology in the STM32 microcontroller. Section 6 presents and discusses the experimental and simulation results. Finally, Section 7 concludes the paper.

2. NOMENCLATURE

C_{SC}	SC Capacitance (F)
CoG	Center Of Gravity
E_{bat}	Battery energy
EMS	Energy Management System
$HESS$	Hybrid Energy Storage System
HIL	Hardware-In-the-Loop
i_{bat}	battery current (A)
i_{bat_ref}	battery current reference (A)
i_{ch_bat}	DC–DC battery Current converter (A)
i_{ch_sc}	DC–DC SC Current converter (A)
i_{DC}	DC BUS Current (A)
i_{req_ref}	HESS Current reference (A)
i_{sc}	SC current (A)
i_{sc_ref}	SC current reference (A)
k_{cont}	Battery contribution factor
P_{bat_Fil}	Filtered battery power reference
P_{bat_lim}	Battery power limit
P_{req}	Required drivetrain power (W)
R_{ESR}	SC Equivalent Series Resistance
$RMSE$	Root Mean Square Error
SC	Supercapacitors
SoC_{sc}	SC State of Charge (%)
U_{bat}	battery voltage (V)
U_{bat_meas}	Measured battery voltage (V)
U_{DC}	DC BUS voltage (V)
U_{ch_bat}	DC–DC battery voltage converter (V)
U_{ch_SC}	DC–DC SC voltage converter (V)
U_{sc}	SC voltage (V)
α_{bat}	DC–DC battery converter duty cycle
α_{sc}	DC–DC SC converter duty cycle
η_{ch_bat}	DC–DC Battery converter efficiency
η_{ch_sc}	DC–DC SC converter efficiency

3. *HESS* MODELING AND CONTROL

The developed system is a 70 kW *HESS* in full-active topology [12], shown in Fig. 1. It includes two energy sources: the primary one is represented by the Li-Ion battery and the secondary source uses the *SCs*. The 100Ah battery and the 5000F *SCs* are interfaced with the DC bus via two bidirectional buck-boost DC/DC converters and a principal coupling capacitor. The characteristics of the power sources are shown in Table 1 and Table 2. The adopted topology has the advantage of decoupling the battery and the *SCs* control, in order to control separately the power flow between the electrical sources and the DC bus voltage.

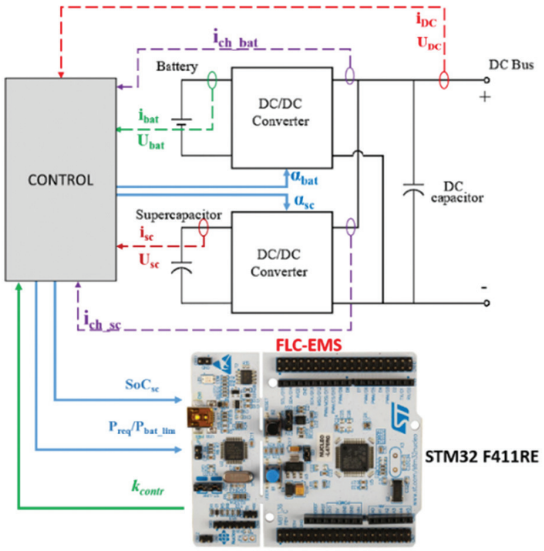


Fig. 1. Scheme of the proposed *HESS*

2.1. *HESS* MODELING

- Battery:

The complex electrochemical phenomena involved in a battery make its modeling highly complicated [18]. The simplified model adopted in this study shown in equation 1, is characterized by an internal capacity C_{bat} and an internal impedance r_{bat} of the battery pack.

$$U_{bat}(t) = r_{bat} i_{bat} + \frac{1}{C_{bat}} \int_{t_0}^{t} i_{bat}(t)\, dt \quad (1)$$

$$SOC_{bat}(t) = SOC_{bat}(t_0) + \frac{1}{Q_{bat}} \int_{t_0}^{t} i_{bat}(t)\, dt \quad (2)$$

The battery state of charge (equation 2) is the ratio of the battery's capacity at any given time to its total capacity. The state of charge is an important parameter to manage efficiently the battery control:

Table 1. Characteristics of the battery source

Battery Li-Ion cell (3.7V)	Symbol	Value	Unit
SoC Limits	SoC_{batt}	[30, 100]	%
Energy density	ρ_{batt}	160	Wh/kg
total mass	M_{batt}	314	Kg
energy	E_{batt}	654	Ah

- Supercapacitors (SC):

Characterized by a fast dynamic behavior, *SCs* have the advantage of being able to support frequent power variations during charging or discharging phases. There are several models in the literature, with variable complexity depending on the purpose of using these models. However, a simplified model has been adopted in this study, shown in equation 3, based on an internal capacitance C_{SC} and the equivalent series resistance R_{ESR}.

$$U_{SC}(t) = R_{ESR} i_{SC} + \frac{1}{C_{SC}} \int_{t_0}^{t} i_{SC}(t)\, dt \quad (3)$$

The calculation model of the *SCs* state of charge in a given instant t, is presented in equation 4. The calculation model is based on the ratio of the voltage of the *SCs* at this moment t to its initial voltage.

$$SOC_{SC}(t) = \frac{U_{SC}(t)}{U_{SC\_init}} \quad (4)$$

Table 2. Characteristics of the SC's source (GTCAP GTSPP-2R7-508UT)

SC (2.7V)	Symbol	Value	Unit
Capacitance	Csc	5000	F
ESR	RESR	0.21	mΩ
SoC Limits	SoCsc	[50, 100]	%

- DC-DC converter:

The technology of the DC-DC converter adopted in this work is a Buck-Boost chopper. This converter allows for ensuring a bidirectional energy flow, in order to ensure the power required by the powertrain on one side, the recharging of the energy sources on the other side, in particular during the braking phases, and the exchanges between both energy sources.

Thus, the first part of the chopper model is presented in equation 5, represented by the current smoothing inductor. This element is characterized by a capacity to store energy as magnetic energy, in order to supply it when it is required. Thus, providing a quasi-constant current profile. The inductor model is characterized by an inductance L_{source} and an internal resistance r_{source}.

$$\begin{cases} L_{bat} \frac{d}{dt} i_{bat} = U_{bat} + U_{L\_bat} + r_{L\_bat} i_{bat} \\ L_{SC} \frac{d}{dt} i_{SC} = U_{SC} + U_{L\_SC} + r_{L\_SC} i_{L\_sc} \end{cases} \quad (5)$$

On the other hand, the switch model is presented in equations (6) and (7). The adopted approach in this case uses an average model with α_{bat} and α_{sc} representing the duty cycles of both choppers, η_{ch_bat}, and η_{ch_bat} are the efficiencies of the DC-DC converters, used in the proposed *HESS*.

$$\begin{cases} U_{L\_bat} = \alpha_{bat}.\ U_{DC} \\ i_{ch\_bat} = \eta_{ch\_bat}.\alpha_{bat}.\ i_{bat} \end{cases} \quad (6)$$

$$\begin{cases} U_{L\_SC} = \alpha_{SC}.\ U_{DC} \\ i_{ch\_SC} = \eta_{ch\_SC}.\alpha_{SC}.\ i_{SC} \end{cases} \quad (7)$$

- Sources Coupling and DC-bus capacitor:

In order to ensure continuous power supplied by both sources, the electrical coupling is required. Indeed, equation 8 shows the electrical model adopted in this case, which is based on the nodes law between both choppers controlling the power sources.

$$i_{hess} = i_{ch\_SC} + i_{ch\_bat} \quad (8)$$

One of the most important criteria for evaluating the *HESS* performance is the DC bus voltage, which should be constant. Thus, integrating a capacitor interfacing the *HESS* and the powertrain allows control and minimizes the U_{DC} voltage fluctuations. Thus, the model of the capacitor is presented in equation 9 characterized by the internal capacitance C_{cpl}.

$$C_{cpl} \frac{d}{dt} U_{DC} = i_{hess} - i_{DC} \quad (9)$$

- Load equivalent model:

In order to simplify the simulation model, and optimize the computation time. The equivalent load model is considered as a controllable current source, where the power profile is presented in Fig. 2. The required power used in the simulation is extracted from a powertrain model, with a maximum power of 70 kW. The power profile depends on the NEDC cycle speed, with a maximum value of 120 km/h.

$$i_{DC} = \frac{P_{req}}{U_{DC}} \quad (10)$$

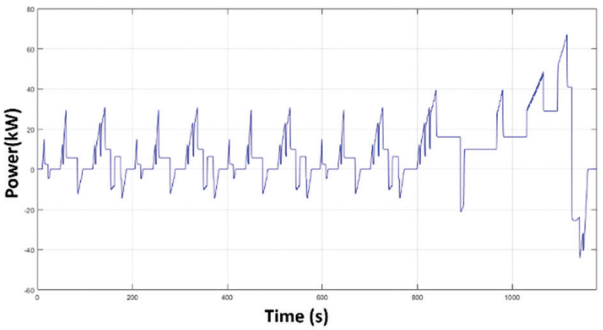


Fig. 2. Power requirement profile of the powertrain

2.2. *HESS* CONTROL

In order to design and establish the control layer, the EMR approach was adopted. The Energetic Macroscopic Representation (EMR) was developed by the Laboratory of Electrical Engineering and Power Electronics (L2EP) of Lille (France) [18, 19, 20] in the 2000s. This formalism allows to the definition of a synthetic representation for complex Multiphysics systems, respecting a functional description. It is based on integral causality and power conservation. Thus, three elements that represent the subsystems can be distinguished: source, energy accumulation, and conversion without energy accumulation. Indeed, the establishment of the control layer has been carried out based on the EMR inversion rules, in order to deduce the maximum control structure (MCS) presented in Fig. 3.

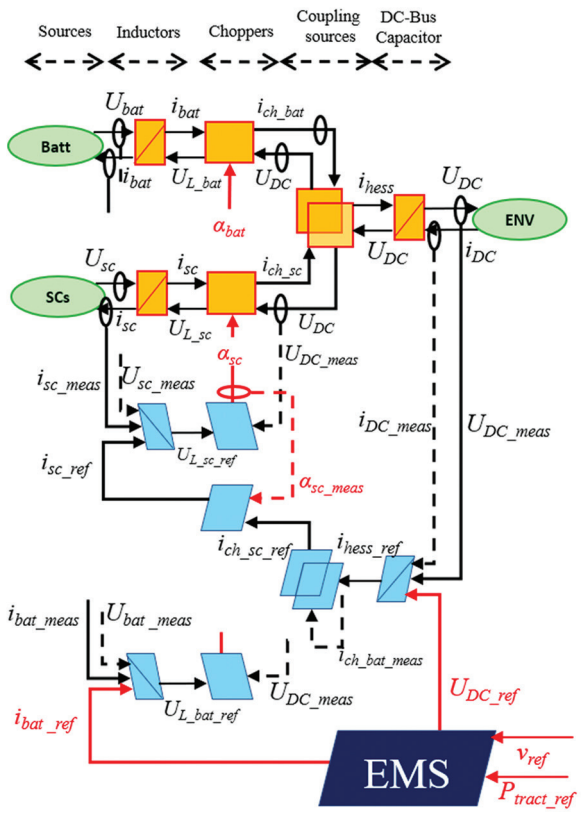


Fig. 3. EMR of the proposed *HESS*

4. ENERGY MANAGEMENT SYSTEM

The *EMS* or Energy Management System in an EV *HESS* refers to the algorithm implemented in an on-board electronic control unit, designed to ensure power management and vehicle propulsion. Hence, *EMS* represents the control strategy, which takes into account the state and model of the vehicle. The *EMS* should control the power distribution between the different energy sources, in order to minimize the solicitations which are harmful to the battery performance. It ensures the required power response respecting vehicle performance requirements (speed, stability, precision, and robustness).

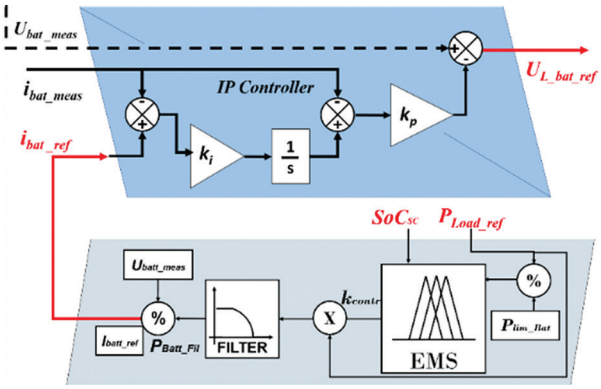


Fig. 4. *EMS* strategy scheme

In fact, this work proposes an FLC-based *EMS* (Fig. 4), providing the same objectives mentioned above. The designed *EMS* operates according to the different phases of the vehicle driving cycle, which are summarized as follows:

- During acceleration:

In this phase, power pulses are generated, which can cause battery degradation. Indeed, the *EMS* should engage the *SCs* to support the power demand of the powertrain, since *SCs* are characterized by their fast dynamics to ensure this function, which allows to preserve the battery and optimize its lifetime.

- During quasi-constant speed:

During this phase, the *EMS* is engaging the battery progressively, to provide the powertrain demand. The *SCs* are activated to support high-frequency harmonics. This distribution is provided by the low-pass filter as shown in Fig. 4.

- During the braking phase:

In order to optimize *EV* autonomy, the braking phase is used as a source of recovered electrical energy. Thus, the *EMS* should activate the chopper of the *SCs* in order to ensure their recharge and restore them to their initial state of charge. This operation allows to the *SCs* control as an energy buffer and ensures their optimal functioning in the range of their optimal energy efficiency $SoC_{SC} \in [50,100]$ %.

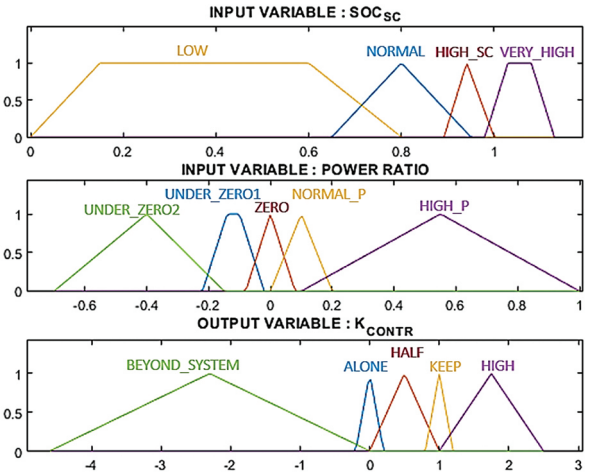


Fig. 5. Membership functions of the proposed Fuzzy Logic *EMS*

- During the total stop of the vehicle:

The architecture adopted in the proposed *HESS* allows the energetic exchange between both energy sources. Thus, during a total vehicle stop, the *EMS* should activate the battery chopper in order to recharge the *SCs*, and restore them to their initial state (when $SoC_{sc}(t) < SoC_{sc\_Init}$).

Indeed, the proposed *EMS* used to satisfy the previously mentioned objectives, is based on an *FLC-EMS*. Thus, the *EMS* output function, shown in equation 11, represents the battery contribution depending on

two input parameters that reflect the vehicle state. Therefore, the first input parameter represented by SoC_{SC} is based on four *MFs* defining the charging level of *SCs*: LOW, NORMAL, HIGH_SC, and VERY_HIGH. The second input parameter is the ratio of the required power of the *EV* powertrain to the battery power limit: $P_{Load\_ref} / P_{lim\_bat}$. This parameter allows defining the *EV* driving phases (acceleration/constant speed, deceleration/braking, vehicle total stop) according to five *MFs*: Under_ZERO2, UNDER_ZERO1, ZERO, NORMAL_P, HIGH_P.

$$k_{CONTR} = (SoC_{SC}, \frac{P_{Load\_ref}}{P_{\lim\_Bat}}) \quad (11)$$

Table 3. Inference Rules-base of the proposed FLC

	SOC_SC			
POWER_ RATIO	**LOW**	**NORMAL**	**HIGH_SC**	**VERY_ HIGH**
UNDER_ ZERO2	BEYOND-SYSTEM	BEYOND-SYSTEM	BEYOND-SYSTEM	BEYOND-SYSTEM
UNDER_ ZERO1	BEYOND-SYSTEM	ALONE	ALONE	ALONE
ZERO	HIGH	HIGH	HALF	ALONE
NORMAL_P	HIGH	KEEP	ALONE	ALONE
HIGH_P	KEEP	HALF	ALONE	ALONE

In fact, according to the literature, FLC processing includes three main steps:

- The first step is related to the input fuzzification (SoC_{SC} and $P_{Load\_ref} / P_{lim\_bat}$), represented by the membership functions (MF) of the inputs shown in Fig. 5.
- The second step is based on the inference rules (Table 1), which are constructed according to the objectives defined at the beginning of this section.
- The last step consists of defuzzification based on Mamdani's Center of Gravity (*CoG*) approach, in order to calculate the k_{CONTR} value which defines the battery's contribution. This part will be detailed in the next section.

5. *HESS HIL* DEVELOPMENT

The real-time *HIL* simulator is a validation approach that combines a hardware platform, in this case, the STM32 microcontroller, and a software simulation platform represented by the proposed Simulink model of *HESS*. The objective of this technique is the evaluation of the energy management algorithm embedded in the microcontroller. Therefore, the validation criteria are based on the control and analysis of parameters that reflect the EV performance. Fig. 6 shows the proposed *HIL* architecture, including the following elements:

- Two Simulink Models of *HESS*, organized in EMR subsystems. The first model uses the Fuzzy library block based on MFs and inference rules, which are discussed in the previous section. The second model uses the Serial Communication blocks with a Simulink block scheme, designed to concatenate the input parameters, in order to ensure their transfer in a single frame.
- The STM32 F411RE microcontroller with 100 MHz clock frequency, connected with the computer through a USB cable via the ST-LINK debugger/ programmer interface.

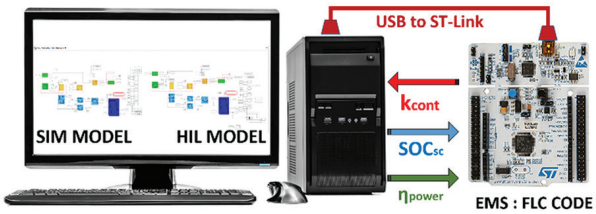


Fig. 6. The proposed *HESS*'s *EMS* Hardware-in-the-loop (*HIL*) architecture

In the first step of the *HIL* processing, the Simulink simulation platform of *HESS* sends both input parameters (SoC_{SC}, $P_{Load\_ref} / P_{lim\_bat}$) via the serial port. Simultaneously, the microcontroller proceeds to load the Membership functions in the predefined array structure. Once the frame is received, the first part of the program performs the extraction, assignment, and cast in the "double" data type of the input variables (SOC_{SC} et $P_{Load\_ref} / P_{lim\_bat}$).

After completing the variables loading, the algorithm performs the fuzzification function. This function allows to determine the degree of membership of each input, the mathematical model of this operation is presented in equation 12.

$$f_{dm} = \frac{(Y_{i+1}-Y_i)}{(X_{i+1}-X_i)} x + (Y_i - \frac{(Y_{i+1}-Y_i)}{(X_{i+1}-X_i)} X_i) \quad (12)$$

Where, (X_i, Y_i) and (X_{i+1}, Y_{i+1}) are the coordinates of two points of the segments which define the MFs triangle. The input parameters (SoC_{SC} or $P_{Load\_ref} / P_{lim\_bat}$) should satisfy the condition of being in the interval defined by this segment.

Thereafter, the function of the inference rules is executed, to define the MFs value of the output parameters, based on the decision matrix which is detailed in previous work in [17]. The inference rules, in this case, are based on the logical operator "and". Each value of the MFs-outputs is defined by the minimum value of the related input values.

Once the degrees of membership of the MFs-outputs are defined, the defuzzification function is activated. This function is designed on the principle of the Center of Gravity (*CoG*) calculation. Indeed, the adopted approach is based on the decomposition of the MFs-output surfaces into independent sub-surfaces S_i (Fig. 7).

The definition of the possible configurations is performed according to the degree of membership value calculated for each MFs γ_i (Fig. 7 and Fig. 8), compared

with the intersection points of the triangle P_i representing the MFs.

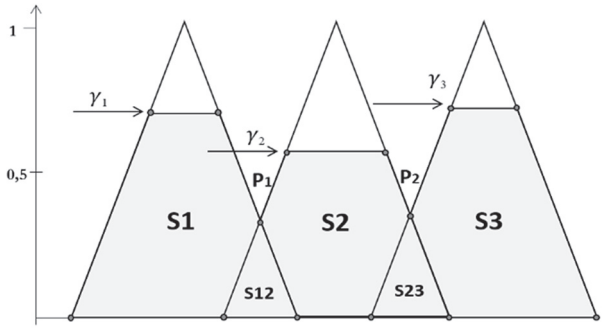


Fig. 7. Vertices of the resulting polygons ($\gamma_i > Y_P$)

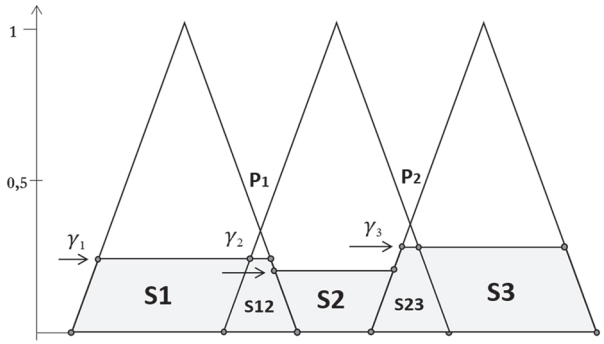


Fig. 8. Vertices of the resulting polygons ($\gamma_i < Y_P$)

Then, this operation allows for defining the configuration and coordinates of each polygon's vertices, and then the calculation of the sub-surface's (S_i) *CoG* which is based on equation 13:

$$\begin{cases} X_{S_i} = \frac{1}{6S_i}\sum_{j=0}^{n-1}(x_j + x_{j+1})(x_j y_{j+1} - x_{j+1}y_j) \\ S_i = \frac{1}{2}\sum_{j=0}^{n-1}(x_j y_{j+1} - x_{j+1}y_j) \end{cases} \quad (13)$$

After calculating the *CoG* abscissae X_i and sub-surfaces S_i, the resulting polygon *CoG* is calculated as follows (equation 14):

$$k_{CONTR} = \frac{1}{\sum_{i=1}^{n} S_i}\left(\sum_{i=1}^{n}(S_i X_{S_i})\right) \quad (14)$$

Thus, the resulting area *CoG* abscissa defines the value of the *FLC* output. It represents the battery contribution, corresponding to the input value and the inference rules established for the proposed solution. Once this value is calculated, it is submitted to the *HESS* simulation platform running on MATLAB/Simulink via the serial port, before starting a new iteration.

6. RESULTS AND DISCUSSIONS

In order to evaluate the *HESS* model and verify the proposed *FLC-EMS* algorithm, both simulation and *HIL* models were built and run in the same software platform (MATLAB/Simulink), aiming to ensure the same number of computational samples (Fig. 6). The simulations were conducted using the NEDC driving cycle with a calculation step of 10-5. The verification of the developed algorithm was done by comparing the results of both models.

Indeed, according to the analysis of Fig. 9, the curves of SoC_{SC} and SoC_{bat} demonstrate that the *EMS* allows controlling the *SCs* as an energy buffer, and the *SCs* recover their initial SoC_{SC} at the end of each cycle $SoC_{SC}(t_{ec}) = SoC_{SC}(t_{init}) = 85\%$. Moreover, the *SCs* operate in the range that is prescribed by the *EMS*: $SoC_{SC} \in [65,86]$ %, which ensures optimal energy efficiency of the *SCs*. Whereas, the SoC_{bat} is generally smoothed and shows a limited variation. This behavior demonstrates that the battery stresses are reduced and therefore an optimization of its lifetime. Thus, according to the evaluation of the *SCs* and battery behavior, as is demonstrated by SoC_{SC} and SoC_{bat} curves in Fig. 10, the *EMS* allows the control of both sources considering the vehicle driving phases:

- *SCs* operate during acceleration phases (e.g., $t \in [10, 14]$ s, $t \in [48, 60]$ s and $t \in [117, 142.5]$ s) when power demand is fast.
- The battery supports the power demand of the powertrain during ($t \in [14, 22]$ s, $t \in [60, 84]$ s, $t \in [142.5, 154]$ s and $t \in [162, 175]$ s) when the power profile is approximately constant.
- During the braking and deceleration phases ($t \in [22, 27]$ s, $t \in [84, 95]$ s, $t \in [154, 162]$ s and $t \in [175, 187]$ s), the *EMS* engages the *SCs* to ensure their recharging from the recovered energy so that they can satisfy the next cycle demand.
- During the total stop of the vehicle ($t \in [24, 27]$ s, $t \in [84, 94]$ s and $t \in [175, 187]$ s), the SoC_{bat} decreases gradually while the SoC_{SC} increases to reach $SoC_{SC} = 85\%$. This demonstrates the possibility of the energy exchange between both sources, ensured by the adopted architecture and the implemented *EMS*.

In addition, Fig. 9 shows the behavior of the DC bus voltage U_{DC}. The analysis of the U_{DC} curve shows that the value of this parameter is relatively constant during the entire NEDC cycle: U_{DC} =163,25 V with ΔU_{DC_max} = 0,5V, which demonstrates the robustness of the implemented control and the developed *EMS*. In fact, the *EMS* ensures a continuous and relatively stable power supply, without compromising the driving cycle performed by the vehicle.

- In addition, the power ratio curve shown in the same figure demonstrates that the value of the power demanded by the powertrain, is kept below the battery power limit ($P_{Load\_ref} / P_{lim\_bat} < 1$). This characteristic guarantees the reduction of battery degradation and the validity of the proposed *HESS* architecture.
- On the other hand, Fig. 9 presents the results obtained from the *HIL* simulations, allowing to validate the developed code of the *EMS-FLC* embedded in the STM32 microcontroller. Indeed, according to the comparison of the curves of the energetic

parameters presented in Fig. 9, both types of results (*HIL* and simulation) are in perfect agreement. Thus, Table 4 presents the differences calculated between both platforms, using the Root Mean Square Error *RMSE*, which presents acceptable differences, due to the defined precision (10-4) in the transferred data between Matlab/Simulink and the STM32 microcontroller.

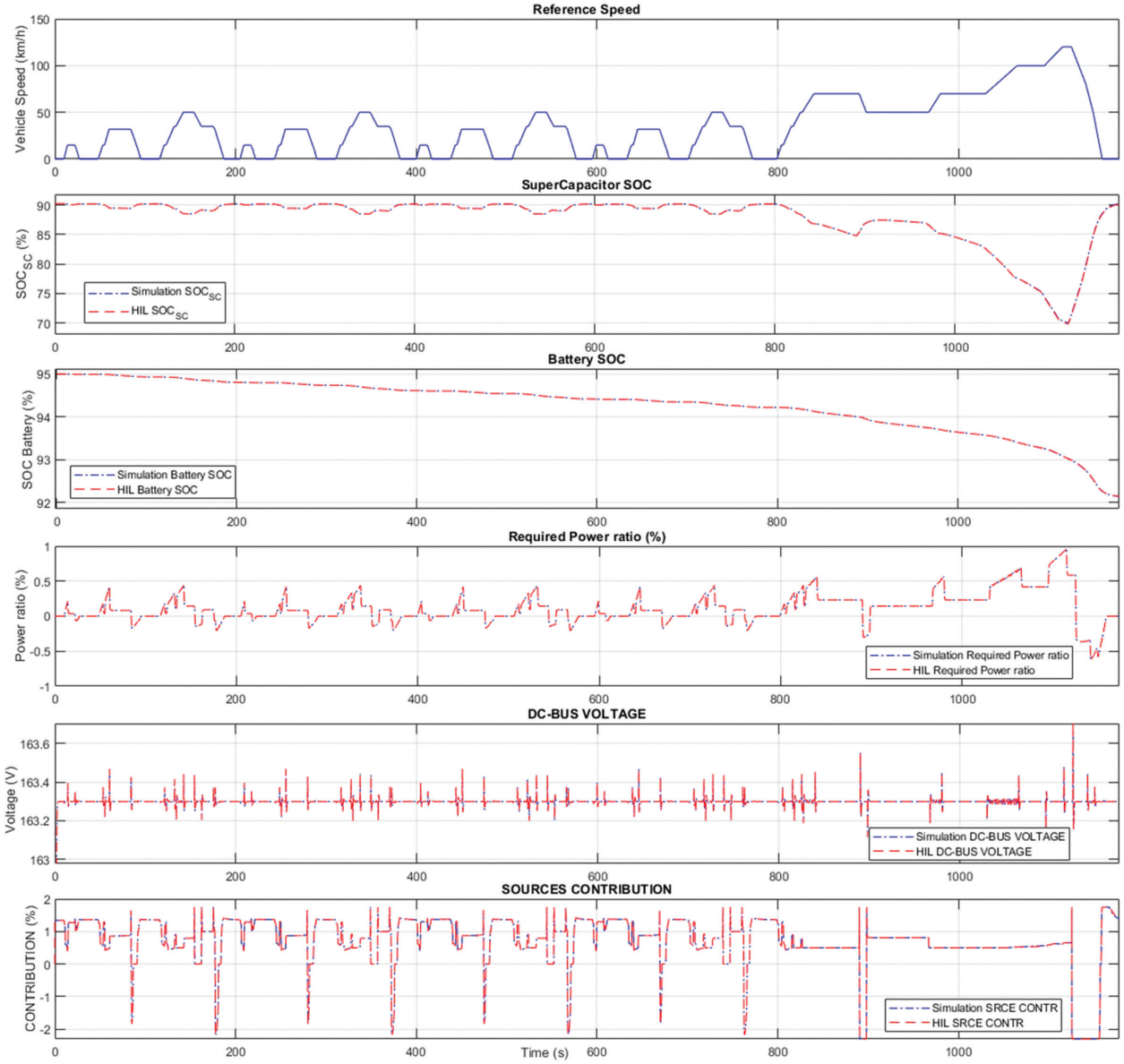


Fig. 9. Simulation results of the EV *HESS* for NEDC driving cycle [0,1180 s]

In addition, the power ratio curve shown in the same figure demonstrates that the value of the power demanded by the powertrain, is kept below the battery power limit ($P_{Load\_ref} / P_{lim\_bat} < 1$). This characteristic guarantees the reduction of battery degradation and the validity of the proposed *HESS* architecture.

On the other hand, Fig. 9 presents the results obtained from the *HIL* simulations, allowing to validate the developed code of the *EMS-FLC* embedded in the STM32 microcontroller. Indeed, according to the comparison of the curves of the energetic parameters presented in Fig. 9, both types of results (*HIL* and simulation) are in perfect agreement. Thus, Table 4 presents the differences calculated between both platforms, using the Root Mean Square Error *RMSE*, which presents acceptable differences, due to the defined precision (10^{-4}) in the transferred data between Matlab/Simulink and the STM32 microcontroller.

Table 4. *RMSE* of *HIL* and simulation energy parameters

Parameters	*RMSE*
SoC_{SC} (%)	3.57 10^{-2}
SoC_{Bat} (%)	1.8 10^{-3}
U_{DC} (V)	2.433 10^{-5}
P_{req} / P_{lim_bat}	2.1085 10^{-17}
Sources Contribution (%)	12.1 10^{-3}

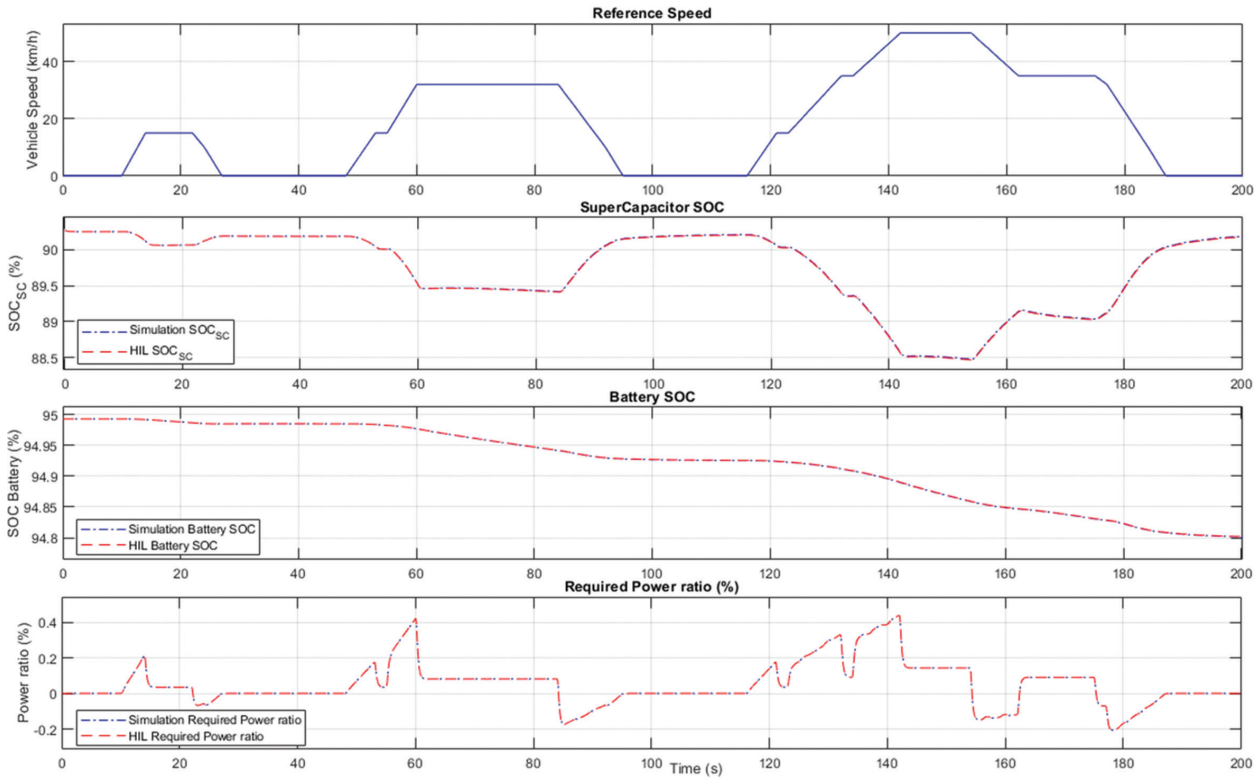


Fig. 10. Simulation results of the EV *HESS* for ECE-15 driving cycle [0,200s]

7. CONCLUSION

In this paper, a real-time *HIL* simulation for a 70kW full-active topology *HESS* is proposed. The energy management system is provided by a rule-based *EMS-FLC*, validated on MATLAB/Simulink simulation platform, developed in C code, and embedded on an STM32 microcontroller. Thus, according to the performance evaluation based on the obtained results, the developed *EMS* satisfied the conditions set for the control of both sources (*SCs* and battery), under the NEDC driving phases on the one hand. On the other hand, the comparison of the results of the energy parameters of both *HIL* and simulation platforms allowed validation of the approach adopted during the development of the proposed C code algorithm, embedded in the STM32 microcontroller.

Therefore, future research will be conducted on a real *HESS* with the same characteristics adopted in this study, using the same approach for the design of an electronic control unit ECU, integrating the control layer and communication interfaces, using the protocols applied in the EVs embedded systems, namely the Control Area Network (*CAN*).

8. REFERENCES:

[1] S. Anenberg, "A global snapshot of the air pollution-related health impacts of transportation sector emissions in 2010 and 2015", International Council on Clean Transportation, George Washington University Milken Institute School of Public Health, and the University of Colorado, Boulder, USA, p. 55.

[2] X. Zhang et al. "A novel quantitative electrochemical aging model considering side reactions for lithium-ion batteries", Electrochimica Acta, Vol. 343, 2020, p. 136070.

[3] Z. Song, J. Li, J. Hou, H. Hofmann, M. Ouyang, J. Du, "The battery-supercapacitor hybrid energy storage system in electric vehicle applications: A case study", Energy, Vol. 154, 2018, pp. 433-441.

[4] B. Yang et al. "Applications of battery/supercapacitor hybrid energy storage systems for electric vehicles using perturbation observer based robust control", Journal of Power Sources, Vol. 448, 2020, p. 227444.

[5] A. Biswas, A. Emadi, "Energy Management Systems for Electrified Powertrains: State-of-the-Art Review and Future Trends", IEEE Transactions on Vehicular Technology, Vol. 68, No. 7, 2019, pp. 6453-6467.

[6] D.-D. Tran, M. Vafaeipour, M. El Baghdadi, R. Barrero, J. Van Mierlo, O. Hegazy, "Thorough state-

of-the-art analysis of electric and hybrid vehicle powertrains: Topologies and integrated energy management strategies", Renewable and Sustainable Energy Reviews, Vol. 119, 2020, p. 109596.

[7] R. Xiong, H. Chen, C. Wang, F. Sun, "Towards a smarter hybrid energy storage system based on battery and ultracapacitor - A critical review on topology and energy management", Journal of Cleaner Production, Vol. 202, 2018, pp. 1228-1240.

[8] M. A. Silva, H. N. de Melo, J. P. Trovao, P. G. Pereirinha, H. M. Jorge, "An integrated fuzzy logic energy management for a dual-source electric vehicle", Proceedings of the 39th Annual Conference of the IEEE Industrial Electronics Society, Vienna, Austria, November 2013, pp. 4564-4569.

[9] J. P. Trovao, M. R. Dubois, M.-A. Roux, E. Menard, A. Desrochers, "Battery and SuperCapacitor Hybridization for a Pure Electric Three-Wheel Roadster", Proceedings of the IEEE Vehicle Power and Propulsion Conference, Montreal, QC, Canada, October 2015, pp. 1-6.

[10] R. S. Sankarkumar, R. Natarajan, "Energy management techniques and topologies suitable for hybrid energy storage system powered electric vehicles: An overview", International Transactions on Electrical Energy Systems, Vol. 31, No. 4, 2021.

[11] H. Jbari, M. Haidoury, R. Askour, B. Bououlid Idrissi, "Fuzzy Logic Controller for an EV's Dual-Source Hybridization", Proceedings of the 4th International Conference of Computer Science and Renewable Energies, Vol. 297, 2021, p. 01039.

[12] Z. Shengzhe, W. Kai, X. Wen, "Fuzzy logic-based control strategy for a battery/supercapacitor hybrid energy storage system in electric vehicles", Proceedings of the Chinese Automation Congress, Jinan, China, October 2017, pp. 5598-5601.

[13] D. S. Castro, R. F. Q. Magossi, R. F. Bastos, V. A. Oliveira, R. Q. Machado, "Low-Cost Hardware in the Loop Implementation of a Boost Converter", Proceedings of the 18th European Control Conference, Naples, Italy, June 2019, pp. 423-428.

[14] C.-H. Wu, Wu-Yang Sean, Y.-H. Hung, "Development of a real-time simulator for electric vehicle with a Dual Energy Storage System", Proceedings of the 2nd International Conference on Power Electronics and Intelligent Transportation System, Shenzhen, China, December 2009, pp. 96-100.

[15] B. Tabbache, Y. Aboub, K. Marouani, A. Kheloui, M. E. H. Benbouzid, "A Simple and Effective Hardware-in-the-Loop Simulation Platform for Urban Electric Vehicles", p. 5.

[16] H. Jbari, R. Askour, B. B. Idrissi, "Real-time aTmega microcontroller-based simulator enabled hardware-in-the-Loop for fuzzy control dual-sources HESS", E3S Web Conf., Vol. 351, 2022, p. 01022.

[17] H. Jbari, R. Askour, B. B. Idrissi, "Real-time FPGA based simulator enabled Hardware-In-the-Loop for fuzzy control dual-sources HESS", International Journal of Renewable Energy Research, Vol. 12, No. 2, 2022.

[18] T. Transi, P. G. Pereirinha, A. Bouscayrol, M. Degano, "Study of Regenerative Braking Effects in a Small Electric Race Car using Energetic Macroscopic Representation", Proceedings of the International Young Engineers Forum, Costa da Caparica, Portugal, May 2019, pp. 106-111.

[19] W. Lhomme, F. Verbelen, M. N. Ibrahim, K. Stockman, "Energetic Macroscopic Representation of Scalable PMSM for Electric Vehicles", Proceedings of the IEEE Vehicle Power and Propulsion Conference, Gijon, Spain, November 2020, pp. 1-6.

[20] B.-H. Nguyen, J. P. F. Trovao, S. Jemei, L. Boulon, A. Bouscayrol, "IEEE VTS Motor Vehicles Challenge 2021 - Energy Management of A Dual-Motor All-Wheel Drive Electric Vehicle", Proceedings of the IEEE Vehicle Power and Propulsion Conference, Gijon, Spain, November 2020, pp. 1-6.

Effective Memory Diversification in Legacy Systems

Original Scientific Paper

Heesun Yun
Sungshin Women's University
Department of Convergence Security Engineering
02844, Seoul, South Korea
yunheesun718@gmail.com

Daehee Jang
Sungshin Women's University
Department of Convergence Security Engineering
02844, Seoul, South Korea
djang@sungshin.ac.kr

Abstract *– Memory corruption error is one of the critical security attack vectors against a wide range of software. Addressing this problem, modern compilers provide multiple features to fortify the software against such errors. However, applying compiler-based memory defense is problematic in legacy systems we often encounter in industry or military environments because source codes are unavailable. In this study, we propose memory diversification techniques tailored for legacy binaries to which we cannot apply state-of-the-art compiler-based solutions. The basic idea of our approach is to automatically patch the machine code instructions of each legacy system differently (e.g., a drone, or a vehicle firmware) without altering any semantic behavior of the software logic. As a result of our system, attackers must create a specific attack payload for each target by analyzing the particular firmware, thus significantly increasing exploit development time and cost. Our approach is evaluated by applying it to a stack and heap of multiple binaries, including PX4 drone firmware and other Linux utilities.*

Keywords*: Legacy System, Diversification, Memory Layout Randomization, UAV Firmware*

1. INTRODUCTION

Memory corruption vulnerabilities are caused by unexpected changes or reference of memory values, and there are various categories of them [1,2]. Generally, the existence and detailed structures of specific memory regions vary by application, but the stack and heap memory regions of most applications, including legacy binaries, are composed based on a general and common memory layout; thus, its exploitation steps are fairly similar from one to another [3-5].

For instance, memory corruption vulnerabilities in a stack such as buffer overflow vulnerability often target to corrupt a specific/common data so-called return address. Defending such memory corruption-based exploitation is also well established by security researchers: use stack canary and check if the canary value has changed [6]. However, such defense (stack canary) is applied by compilers; thus, it is difficult to apply such defense to legacy binary which cannot be re-compiled.

Various vulnerabilities are also likely to exist in the heap memory area as the heap is dynamically allocated at runtime and it is usually more complicatedly constructed than the stack memory buffers [7]. Similarly, to stack situations, multiple defense systems including address space layout randomization (ASLR) are proposed as a part of the compiler feature or heap-allocator's runtime verification logic [8].

All such state-of-the-art defenses are inapplicable to legacy binaries as they cannot be re-compiled [9]. However, we claim that applying the ASLR effects (even partially) to stack and heap memory regions would apply to some extent solely based on binary analysis and in-place instruction patching [10,11]. This is the part where our idea comes in.

To apply the memory diversification effect to legacy binaries, we use binary analysis techniques and in-place instruction patching to change the runtime memory layout of binary without re-compiling the source code.

The goal is to create different memory layouts for each physical device, without altering the original functionality. More specifically, we patch the binary/firmware to differentiate the size/distance of unused memory pad-

dings in stack and heap. As a result, when our diversification technology is applied to drones, even if the software system is prone to memory corruption, the attacker must adjust the attack payload differently each time when attacking an instance because diversification renders the memory layout of each physical instance differently. As a result, an attacker must create multiple attack payloads by analyzing the instance of each machine, which significantly increases attack time and cost.

Several related works also consider the lack-of-source-code situation. For example, [9] also fortifies the binary to prevent exploitation. The key idea of such work is to patch the ROP-gadget instruction instead of changing the memory layout as our system does. We have surveyed previous studies to this end and summarized more details in section 2.

In short, we proposed binary-patch-based memory layout randomization technique that does not require source code and is thus suitable for legacy systems. The main contributions of this paper are as follows.

1. Our system provides a way to change the memory layout without re-compiling the binary. This is useful to fortify outdated system binaries when their source code is intentionally or unintentionally unavailable.
2. Our system effectively raises the bar of applying memory corruption attacks against multiple physical devices using the same firmware. Because our approach diversifies the memory layout of each target, the attacker must develop different exploit payloads (e.g., ROP payload for buffer overflow vulnerability) for each target device.
3. We implemented and deployed our idea against commercial-off-the-shelf (COTS) software such as Intel PX4 aero drone firmware and popular Binutils binaries in the Linux system.
4. We provide academic analysis and discussion that our approach must consider data alignment issues and code/data interleaving problems in ARM-like CPU architectures. This is a thought-provoking discussion for additional future research to this end.

The rest of the paper is structured as follows. Section 2 covers various related works regarding our research direction. Section 3 proposes a system theory about the legacy system and memory safety. Next, we describe the design of our framework in section 4, and the implementation of algorithms in section 5. Afterward, we present the evaluation results in section 6. Finally, we summarize our paper and conclude in section 7.

2. RELATED WORK

2.1. BINARY PATCH

Because many commercial software applications are developed based on closed-source code, it is difficult to apply security patches to their vulnerable parts. Various researches are addressing this limitation. [12] conducted a study to prevent return-oriented programming (ROP) attacks based on the return instruction without based on debugging symbol information in a Linux 64-bit environment without source code. [13] conducted a patch for a buffer overflow vulnerability by automating the security patch for the Windows x86 binary without the use of source code, debugging information or human intervention. [14] considers a closed-source code patching environment in PowerPC-based binaries. [9] achieves security effectiveness by patching the binary to prevent code reuse attacks. [15] identified vulnerabilities in closed-source code software based on the bug signature. [16] performed a static binary patch to prevent vulnerabilities in ARM-based Internet of Things (IoT) device firmware.

Additionally, many studies aim to better identify the causes of security vulnerabilities in binary. [17], [18] enabled finding similar patches or vulnerabilities in different binaries by identifying code portions changed by a binary patch through basic block analysis. [19] detected software vulnerabilities in the binary codes of patched and unpatched programs using the patch diffing technology. Previous studies primarily aimed at preventing specific security vulnerabilities in an environment with binary-only approach without source code. To some extent, our study is similar to these studies in that patch is performed in a binary-only environment, but our framework is mainly focused on diversifying the memory layout and hindering memory corruption exploitation for physical instances (e.g., each drone machine).

2.2. MEMORY RANDOMIZATION

As attacks that exploit memory corruption vulnerabilities have been launched against several commercial software applications, the latest security technologies for memory protection, such as ASLR, have been implemented in most software applications. ASLR is a technology that can defend against attacks using a fixed address by randomly changing the address of the data area whenever a binary is executed. Thus, research to prevent external attacks by randomizing part of the memory is being actively conducted. [20] proposed a technology to randomize the stack layout based on the LLVM compiler to prevent memory corruption vulnerabilities. The position of each object was randomized using source code information so that attackers cannot predict the stack layout. [8] performed randomization for the entire heap memory area by proposing a random memory block allocation algorithm. It is similar to our study in that randomization is performed in the heap memory area, but the detailed process to obtain the randomization effect is different. [21] proposed a solution to prevent just-in-time (JIT) code reuse attacks using memory randomization technology that is subdivided for each process so that code sharing is not disturbed. [22] conducted a study on the kernel defense mechanism by performing device driver randomiza-

tion and proposed a solution to prevent ROP attacks on the Linux kernel by increasing the KASLR entropy. [11] proposed a method for randomizing the memory layout in user-end machines to prevent buffer overflow attacks. As shown in the above examples, memory randomization to reinforce security is being extensively researched. Although some studies do not use source code and symbols, many studies perform memory randomization using the source code information. Our study performed effective memory randomization only for binaries without source code information.

2.3. LEGACY SYSTEM

A legacy system is one that uses software applications from older versions or works in uncommon ways. Currently, the latest software updates and distributions are done quickly based on open sources, but many legacy systems are still being used in military or space exploration systems. Because these legacy systems are mostly based on closed-source code, it is difficult to apply security patches to them. In this situation, research into the reinforcement of the security of legacy systems is indispensable. [23] generates randomized instruction addresses whenever a legacy x86 binary is executed without source code or symbol information. [9] proposed a binary-based rewriting technology for patching legacy binary to prevent code reuse attacks. [24] proposed a solution for detecting control flow integrity (CFI) attacks, such as buffer overflow and ROP attacks. It detected CFI attacks with high accuracy by inserting performance counters and instrumentation hooks through binary editing in the ELF file, which is legacy software.

3. PROPOSED SYSTEM THEORY

3.1. LEGACY SYSTEM

A legacy computer system in this paper refers to a computer device with old software application/firmware that differs from the recent standard. Currently, the development and distribution of software are done quickly based on state-of-the-art compilers, yet many legacy systems are still used as old version binaries in various fields. Because most legacy systems are closed-sourced, and not properly updated, they can be a good target for attackers to abuse system vulnerabilities such as memory errors.

3.2. ASLR

ASLR is a memory layout diversification technology that changes the address of the data whenever a program is executed by randomly placing memory layouts such as stack, heap, and library codes in the unpredicted address space to make memory exploitation attacks difficult. As memory exploitation technologies such as return oriented programming (ROP) are becoming more popular, memory protection techniques such as ASLR are now widely used in most software systems.

3.3. BASIC BLOCK

A basic block in computing is a straight-line code sequence with no branches other than the entry and exit. Because of this characteristic, computer science researchers often analyze binaries based on basic blocks as a unit of algorithm testing/measure.

3.4. MEMORY ALIGNMENT

When data structures or classes are stored in memory in programming languages such as C++, paddings are sometimes inserted in between variables. These paddings are dummy values added by the compiler. The memory is aligned in 4-byte units to optimize the performance when the CPU accesses the memory. Because of the data bus structure between the CPU and memory, all variables that enter the memory must be placed in consideration of memory alignment to improve system performance [25]. In some cases, if data is accessed to an address without considering memory alignment, it can cause an alignment fault rather than causing a performance issue [26]. In our paper, we consider memory alignment problems while applying our patch for memory diversification.

4. DESIGN

4.1. OVERVIEW

The overall framework of this study is shown in Fig. 1. The firmware binary of the legacy system is ①disassembled using the Capstone [27] library and ②functions are extracted based on binary analysis. Afterward, ③ we analyze basic blocks in the function to run our code detection algorithm for proper diversification patch. For the next step, ④ our patch tool performs a diversification patch to the corresponding assembly instructions to adjust the memory layout. After our algorithm is applied, ⑤ all the basic blocks are reassembled into a diversified version of binary. We applied this approach to PX4 drone firmware to confirm its effectiveness and stability. However, in the case of Windows CE binaries, we could not apply the final re-assembly step because the underlying tool only considers Linux-based binary. In such a case, we applied in-place binary patching directly against machine codes with additional heuristics.

4.2. STACK DIVERSIFICATION

The overall methodology for stack memory diversification is to insert a random-sized dummy padding inside each stack frame. Because each binary might have unique compiler options, the details of the stack memory layout may differ from one instance to another. However, the main structure of the stack frame and its uses are mostly similar. To launch a successful attack via buffer overflow vulnerability in the stack; the return address, buffer position, layouts, and offsets among local variable data must be precisely calculated. A general

defense technology against such attacks is to place a stack canary between a local variable and the return address, then check whether the stack canary has been changed before exiting the function. If the canary value has been changed, it indicates that unintended data overwrites occurred during the execution of the function.

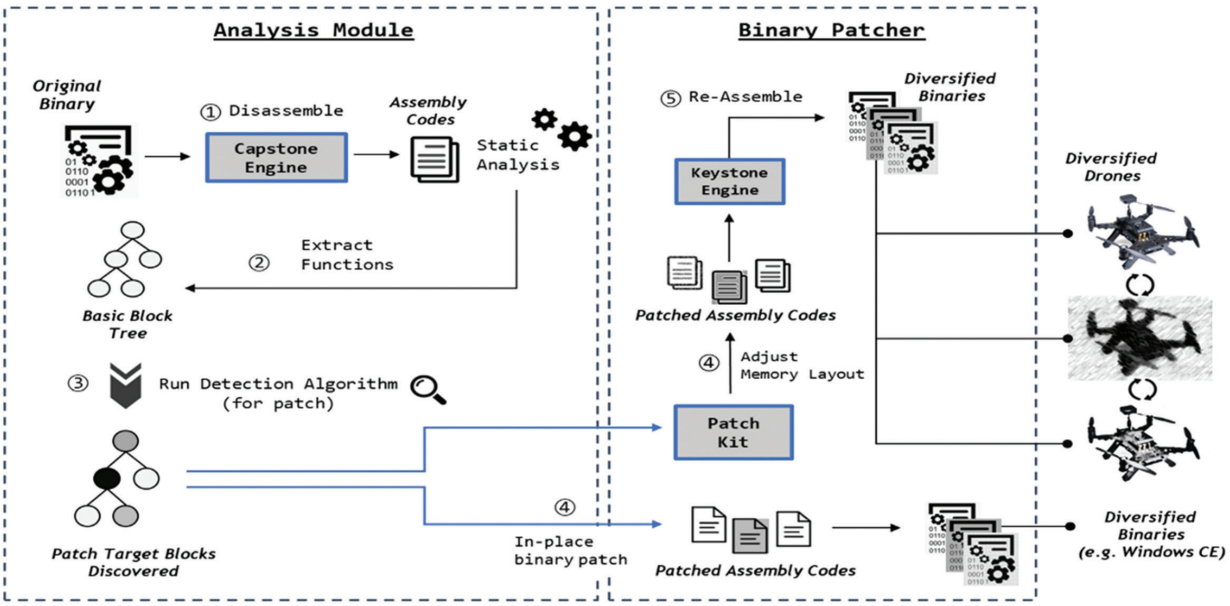


Fig. 1. System Overview. Analysis module spots patch targets based on various pattern detection algorithms. Binary patcher is responsible for modifying and re-assembling the target binary.

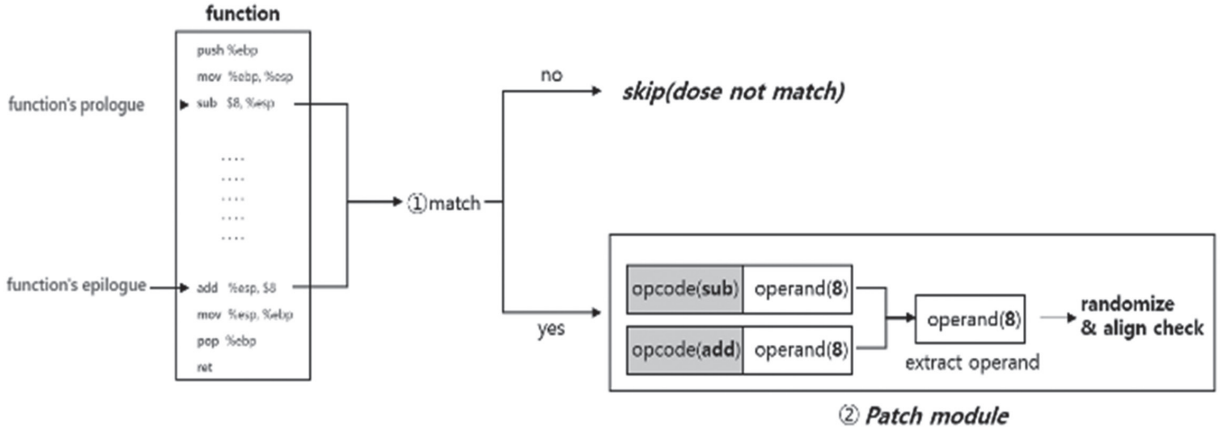


Fig. 2. demonstrates how our patch algorithm checks whether the function's prologue and epilogue match each other. If the condition is satisfied, the patch module performs randomization patch. The patch module extracts the operand from the assembly

However, without re-compiling, the source code, adding stack canary features to legacy systems is difficult. Fig. 3 and 4 show the codes before and after applying the stack canary. Lines 81 to 90 in Fig. 4 represent the stack canary part. The figure shows the same program, but with/without additional assembly codes as the stack canary defense feature. The problem with legacy systems is that it is difficult to *insert* additional assembly code without breaking the other codes. However, it is feasible to in-place *modify* the assembly codes without breaking other codes. Therefore, instead of using a stack canary; in this study, we diversify the distance between the local variable of the stack and the *return* address via inserting a random-sized padding by modifying the stack-offset related variables embedded in the assembly instruction. This can be done relatively simply based on a binary patch, considering how the compiler allocates the stack frame.

The compiler adds the stack allocation and deallocation codes as function prologue and epilogue. Our algorithm detects this code via binary analysis and randomly modifies the size of the stack frame dummy padding differently for each legacy instance. The point of our idea is that we can make this modification solely based on in-place binary patching. To build our system, our patch framework analyzes each binary architecture of the target binary, such as ARM and Intel, and applies

a tailored algorithm to each CPU architecture. For example, ARM uses fixed-length instruction encoding, while Intel uses variable-length instruction encoding.

Fig. 3. Typical disassembly result w/o stack canary

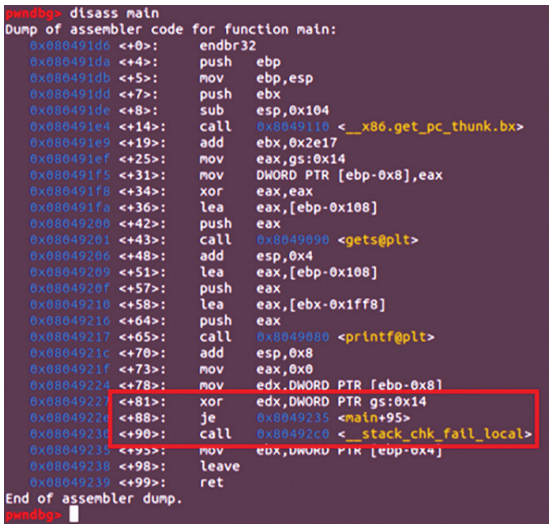


Fig. 4. Typical disassembly result w/o stack canary.

The compiler adds the stack allocation and deallocation codes as function prologue and epilogue. Our algorithm detects this code via binary analysis and randomly modifies the size of the stack frame dummy padding differently for each legacy instance. The point of our idea is that we can make this modification solely based on in-place binary patching. To build our system, our patch framework analyzes each binary architecture of the target binary, such as ARM and Intel, and applies a tailored algorithm to each CPU architecture. For example, ARM uses fixed-length instruction encoding, while Intel uses variable-length instruction encoding.

Fig. 5 shows the general function prologue and epilogue for stack frame allocation and de-allocation. In Fig. 5, the size of the local variable is 0x10. Because 0x10 is part of the 32-bit operand encoding (10 00 00 00) of the instruction, this instruction can be changed in-place without reducing or expanding the code. The essence of our idea is to randomly increase this size to make the distance between the stack buffer and the return address unpredictable. The unnecessarily increased buffer size can be considered as a dummy padding inside memory space which has no harm to program execution. However, if this value is decreased, the program may be damaged as the original buffer cannot hold the given data.

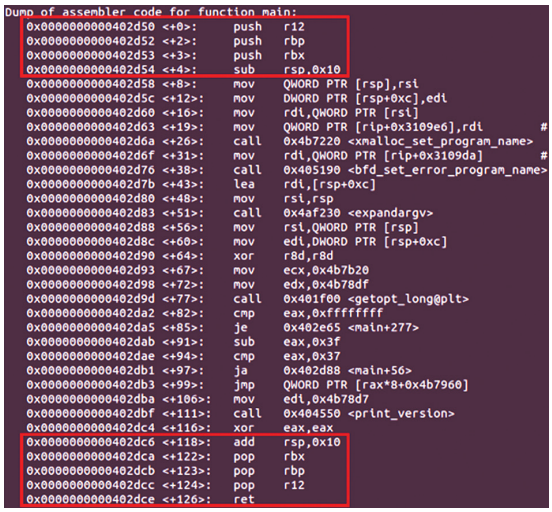


Fig. 5. Typical function's prologue and epilogue code (Binutils cxxfilt). Such code patterns can change depending on compiler options.

The patch appears simple if the stack frame allocation and de-allocation code are symmetric (e.g., add a value and subtract the same value later). However, there are cases where the stack allocation codes are more complicated. For example, when a stack frame of size 100 is allocated by subtracting 100 from to stack pointer and then de-allocated twice with the size of 50, it is difficult to determine which value is the stack frame size because the correspondence of allocation and de-allocation codes do not match based on the same operand size. Therefore, we use additional heuristics to filter out exceptional cases of stack frame allocation.

Furthermore, an additional check is performed to determine whether the changed size is appropriate for the diversification patch. The stack size must be increased with 4-byte granularity considering the CPU word alignment, which improves performance when the CPU accesses the memory and prevents alignment faults [25,28].

4.3. HEAP DIVERSIFICATION

The overall concept of applying diversification to heap memory is similar to the previous stack case (binary analysis and finding the appropriate patch call site), but the detailed method is quite different. Heap memory corruption vulnerabilities are generally related to use-after-free and type confusion. Use-after-free refers to a vulnerability that can occur when a dynamically allocated heap space is freed and reused, whereas type confusion refers to a vulnerability that occurs when the instance of an object confuses the type. The heap is a memory segment with complex data dependency

based on multiple pointer chains. The code pointer in the stack only needs to consider the return address, whereas, in the heap, it is relatively difficult to find a code pointer based on their memory location. Hence, even slight diversification of the heap memory layout can have a significant impact on exploit development.

Our binary analysis algorithm for heap memory diversification differs from the stack analysis. Generally, heap layout can be dynamically changed using standard methods such as IAT hooking in Windows and LD_PRELOAD in Linux. However, this approach is only possible for *dynamically linked binaries*. However most legacy binaries are based on static linking. This raises the question of how the heap allocation call site can be detected based on a generic algorithm. In this paper, we consider the binary has symbols to indicate the heap allocation function. Based on such a premise, our algorithm tracks down heap memory allocation sites in the legacy binary by parsing the symbol table. Then, the call site is backtracked to find the memory allocation size parameter (to modify it). However, a few problems occur while finding this size information.

The first is the data interleaving problem. In Fig. 7, the address 0x807f378 represents a heap allocation call site (malloc). To find the allocation size parameter, the instruction must be backtracked, but there are interleaved data inside the code (0x807f370). This data can be misinterpreted as instruction and hinder the analysis. To solve this problem, our algorithm additionally traces the boundary of each basic-blocks and runs backtrace only within the scope inside the block. However, even within the same basic block, there are exceptional cases.

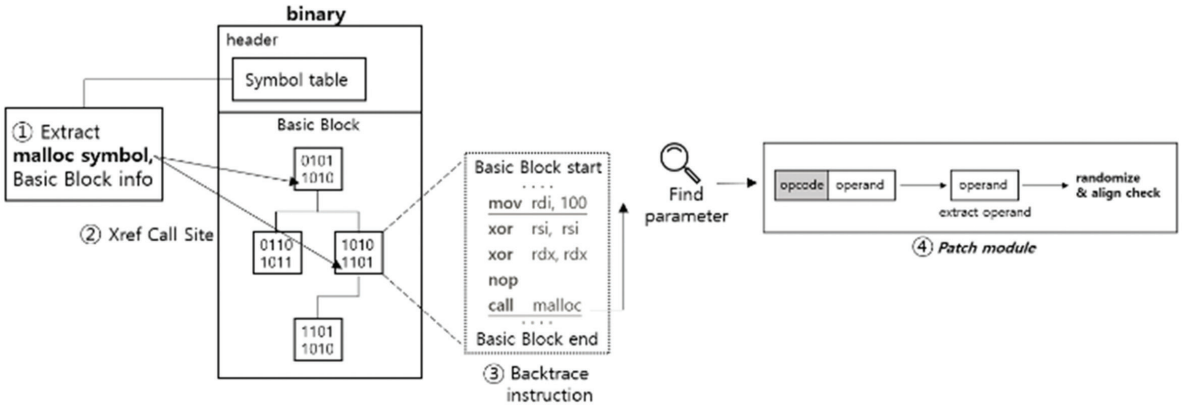


Fig.6. shows the malloc symbol and basic block information extracted from the symbol table in the header of the binary. After extracting the symbol information, the call site information is obtained from the symbol, and the parameter information is obt.

```
.text:0807F368                  BCC     loc_807F30A
.text:0807F36A                  B       loc_807F320
.text:0807F36A ; End of function _ram_flash_wait(sem_s *)
.text:0807F36A
.text:0807F36A ; ---------------------------------------------
.text:0807F36C off_807F36C      DCD _ZL18dm_operations_data
.text:0807F36C                                          ; DATA
.text:0807F36C                                          ; dm_op
.text:0807F370 dword_807F370    DCD 0xF4240             ; DATA
.text:0807F370                                          ; _ram_
.text:0807F374
.text:0807F374 ; =============== S U B R O U T I N E =========
.text:0807F374
.text:0807F374
.text:0807F374 ; _DWORD __fastcall _ram_initialize(unsigned ir
.text:0807F374 _ZL15_ram_initializej                    ; CODE
.text:0807F374                                          ; DATA
.text:0807F374                  PUSH    {R4-R6,LR}
.text:0807F376                  MOV     R5, R0
.text:0807F378                  BL      malloc
.text:0807F37C                  LDR     R6, =_ZL18dm_operations
.text:0807F37E                  MOV     R4, R0
.text:0807F380                  STR     R0, [R6]        ; dm_op
```

Fig. 7. Example case of data interleaving in ARM binary.

```
08019D46 83 68          LDR     R3, [R0,#8]
08019D48 1F 69          LDR     R7, [R3,#0x10]
08019D4A 38 46          MOV     R0, R7
08019D4C 01 F0 9C FA    BL      fat_semtake
08019D50 38 46          MOV     R0, R7
08019D52 01 F0 FB FA    BL      fat_checkmount
08019D56 06 46          MOV     R6, R0
08019D58 30 BB          CBNZ    R0, loc_8019DA8
```

Fig. 8. Exceptional case for tracking function parameter. In the example, it is ambiguous to determine which R0 is the actual parameter for the function fat_checkmount.

Fig. 8 shows two function parameters in the same basic block. Because function parameters are stored in a stack and restored immediately before they are used, it is difficult to determine which parameter is the actual size parameter based on static code analysis.

5. IMPLEMENTATION

To implement our algorithms, we use Patchkit [29], for ELF binary patch, and Capstone [27], Radare2 [30], and Keystone [31] tools for binary analysis and disassembly/reassembly. Overall, implementations are written in 979 lines of Python and 1,495 lines of additional Docker/management codes.

5.1. ARM STACK

Algorithm 1 – The ARM Stack pseudo code. Prefixes/Postfixes such as "BB" denotes the basic block of function, "dis" denotes the dissemble result of the Capstone library, "ins" denote each instruction, and "SP" denotes the stack pointer.

0: **function** PATCH(binary)

1: **for** func in binary_funcs() **do**

```
2:    startBB ← func_startBB_dis
3:    for ins in startBB do
4:        if ins_OpcodeName = "sub" then
5:            src, dst ← ins_operands
6:        else
7:            Skip
8:        end if
9:        if src = SP_reg & dst = stack_size then
10:           prologue ← ins_addr
11:       end if
12:   end for
13:   for ins in reversed(endBB) do
14:       if ins_OpcodeName = "add" then
15:           src, dst ← ins_operands
16:       else
17:           Skip
18:       end if
19:       if src = SP_reg & dst = stack_size then
20:           epilogue ← ins_addr
21:       end if
22:   end for
23:   if prologue = epilogue then
24:       function RANDOMIZE(stack_size)
25:           get (alignment info)
26:           ran_num ← random.randrange(0, 0x10)
27:           new_stack_size ← stack_size +ran_num
28:           Patch()
29:           end function
30:       end if
31: end for
32: end function
```

The overall flow of the stack randomization for ARM 32bit is shown in Algorithm 1. The disassembly result for the first basic block of each function is initially processed, and this result includes instructions such as push and *mov*. To find the address of the subtraction instruction of the stack (e.g., memory allocation), each instruction in the basic block is analyzed, and if the opcode is *sub*, the operand is remembered. If this value is constant (e.g., immediate type value), we conclude this is the stack frame size corresponds to the function prologue. Next, we scan each instruction to find the corresponding add instruction as the match for stack memory deallocation. Afterward, a new stack size is generated using the diversification algorithm, and the stack size is randomly increased respecting the CPU word alignment. Finally, the stack-frame size related instructions are patched in-place.

5.2. X86 STACK

The x86 stack randomization evaluation is performed using the Windows CE binary. Because the Windows CE binary cannot use the Patchkit library [29], our framework implemented a custom code detection for our algorithm. For patching, our tool traces instructions starting from the entry point and applies the same algorithm as ARM stack. In the Windows binary, it was more difficult to identify the function location, size, and internal structure compared to ARM. However, it was easy to apply the patch if the function frame used *leave* instruction for function epilogue as the instruction automatically calculates the required stack size for deallocation.

5.3. ARM HEAP

Algorithm 2 – The ARM Heap pseudo code. Prefixes/Postfixes such as "BB" denotes the basic block of function, "dis" denotes the dissemble result of the Capstone library, "ins" denote each instruction.

```
0: function PATCH(binary)
1: for target_addr in malloc_BB do
2:     bb ← malloc_BB[target_addr]
3:     startpos ← target_addr – len(bb) + 4
4:     for ins in dis(bb, startpos) do
5:     if ins_mnemonic is "movs" then
6:         if ins_operands is register then
7:             Reg ← ins_reg_name
8:         end if
9:         if Reg = "r0" then
10:            MRU ← (ins_addr, ins_str, bb[offset|opcode])
11:        end if
12:    end if
13:    end for
14:    if MRU ≠ None then
15:        get(alignment info)
16:        pad ← random.randrange(0xf0 - imm)
17:        Patch(MRU, pad)
18:    end if
19: end for
20: end function
```

The overall flow of the heap randomization for ARM 32bit is shown in Algorithm 2. Before performing randomization, the malloc position is first found by parsing the symbol table of the target binary. Then, all the basic blocks that contain the malloc address are loaded and saved. At this point, analysis is performed considering the ARM and THUMB modes.

The core of the algorithm is to find a pattern corresponding to the heap size allocation of malloc. When movs instruction is found, we check if the target register corresponds to *r0* because *r0* is the register used as the first function parameter. We determine that the value for *r0* is the malloc size based on heuristics and save all the size data. All values that contain *r0* inside the basic block are collected, and if there are multiple values, the closest value to the malloc call site is used as the malloc size parameter. The randomization patch is then performed. In the heap, memory alignment must follow 16-byte (128-bit) granularity because of the ARM NEON instructions.

5.4. X86 & X64 HEAP

To perform the heap randomization on x86 and x64, we use the malloc address from the symbol table and *basic block border* information of the entire functions based on binary analysis. We use basic block border information to apply our algorithm solely to a single block. Similarly to the ARM case, if the *mov* instruction operand is the immediate type and is closest to the call site, we consider the information as a size parameter. Finally, we also respect the memory alignment of 16-byte (128-bit) granularity considering the use of SIMD instructions using 128-bit MMX registers.

6. EVALUATION

In this study, experiments were performed by porting all related codes into a Docker container image for efficiency. The server environment for our evaluation is composed of 16GB RAM, and 1TB SSD, based on Ubuntu 16.04 64bit server. The Intel PX4 Autopilot Drone firmware (ARM 32-bit) was mainly used for stack diversification testing. The soundness of patched drone firmware was checked with basic system functionalities after the booting process.

6.1. STACK

1) Buffer overflow Toy example

We evaluated our system using a simple x86-based toy binary that has memory vulnerability. In the experiment, we used a general stack buffer overflow vulnerability exploitation as a test case.

When the example code is executed, buffer overflow is triggered, and a segmentation fault error occurs because the return address is broken. However, when a buffer overflow vulnerability was triggered for a binary with diversified stack sizes using our patch tool, the result of the attack was unpredictable.

```
void test(char *input, int len) {
    char buf[4];
    memcpy(buf, input, len);
}

int main(int argc, char* argv[]) {
    if(argc!=2){
    printf("usage: ./test [bof size]");
    return 0;
}
```

Fig. 9. A toy program for testing memory diversification against stack.

```
running buffer-overflow to unpatched binary...
./run.sh: line 15:    60 Segmentation fault

==========================================

now running buffer-overflow to patched varients
bof test
no bof
```

Fig. 10. Screen capture of running toy example case.

Fig. 11. Result after applying stack memory diversification against PX4-based drone. The screen on the left side indicates internal command/communication operates without problem after patching. Right side of the picture shows LED blinking of the UAV that indicates its normal operation.

We evaluated our system using PX4 Autopilot Drone firmware (ARM 32-bit) binary as a test case of ARM binary. When the stack memory layout in the PX4 firmware was randomized using our patch tool and then executed in the same manner as the firmware before the patch, all functions worked normally despite thousands of codes being modified for diversification. Fig. 11 shows *nsh* program, the shell interpreter of the drone. All the commands in *nsh* shell operated correctly before/after our patch.

3) Windows CE binary

We evaluated our system using a collection of 42 binaries, including *pmkdir*, *pmemmap*, and *pdebug*, which are tools that can be used in Windows CE provided by [32]. Wine [33] was used for compatibility to run Windows-exclusive programs on the Linux operating system. All functions worked normally when the diver-

sification-patched binaries were executed. Randomization patch was performed in several tens of places in the case of the Windows CE binary.

6.2. HEAP

1) Toy example

We use a simple toy example to evaluate our heap diversification patch. Fig. 13 shows the before/after result of the diversification patch to the example program. The randomization result showed that the heap allocation address was all changed, however all *printf()* functions in the example code were executed normally. The top portion of the allocation address was changed because of the ASLR effect (not because of our diversification), however, and the bottom offset portion (marked with a red box) was also changed by our patch tool (this part should not change in the normal case). We used gcc and clang with multiple optimization levels ranging from zero to three, and Fig. 13 is based on the optimization level zero.

```
void f1(){
    char* ptr = malloc(100);
    ptr[5] = 1;
    printf("heap buffer at %p\n", ptr);
}
void f2(){
    char* ptr = malloc(40);
    ptr[5] = 1;
    printf("heap buffer at %p\n", ptr);
}
void f3(){
    char* ptr = malloc(256);
    ptr[5] = 1;
    printf("heap buffer at %p\n", ptr);
}
void f4(){
    unsigned int size = getpid()*1000;
    size = size%1000;
    char* ptr = malloc(size);
    memset(ptr, 0xff, size);
    printf("heap buffer? at %p\n", ptr);
}

int main(){
    f1();
    f2();
    f2();
    f2();
    f3();
    f4();
    printf("end\n");
    return 0;
}
```

Fig. 12. A toy program for testing memory diversification against heap.

Fig. 13. Before/after applying heap randomization patch. The red box indicates the internal heap offset is changed across the patch (the upper part of the address is changed due to ASLR).

2) Binutils

We applied our system to the GNU Binutils binaries for performance evaluation. The Binutils version in our evaluation is based on 2.31.1 and the architecture is 64-bit Intel. After the randomization patch, all binaries are executed as originally intended. To measure the execution time delay due to the unnecessarily expanded memory layout, we repeatedly tested the execution time of various Binutils programs before/after we applied our system. The program execution time is measured based on a simple python script.

Table 1 summarizes the result (program execution time) of the comparison between binaries before and after patching. We noticed that there was tiny degradation in the execution time before and after the patch. However, the overall amount of performance degradation is negligible.

Table 1. This table is organized by comparing the binary execution time before and after randomizing to the Binutils binary using our tool. Table 1 is the result of 1000 runs each, and the following options were used to test.

Binary	Before	After	Option
addr2line	0.809s	0.732s	Check the file related to the execution file address and the line information
ar	0.877s	1.014s	Generate an archive
nm-new	0.916s	0.921s	Check the symbol information of the file
objcopy	0.751s	0.749s	Generate a new file after extracting the instructions and data of the file
objdump	1.113s	0.867s	Check all contents of the object file
ranlib	0.826s	0.825s	Register the library so that it can be used
readelf	3.516s	3.682s	Output the dependency of 'ld' including type
size	0.805s	0.848s	Output the file size
strings	1.874s	1.904s	Output the offset of the execution file
strip-new	0.879s	0.938s	Remove the symbol

3) PX4 drone

We used the PX4 Autopilot Drone firmware binary for heap diversification evaluation as well. Similarly to the stack experiment, all functions worked normally when the heap memory layout in the PX4 firmware was randomized with our patch tool and the binary was executed in the same manner as the firmware before the patch.

7. CONCLUSION

With the emergence of memory corruption vulnerabilities, defense technologies that apply diversification to software systems, such as stack canary and isolated heap, are being extensively researched. However, because the

source codes are unavailable, such defense technologies cannot be applied to very old legacy systems. To remedy such a problem, this study presents an alternative approach to apply memory layout adjustment solely based on binary analysis and patching. We mainly apply diversification against two memory components: stack and heap. As an experimental approach, we chose stack and heap as diversification targets because they are the most common memory section of any computer system. By applying our idea, it is expected that the attack time and cost will increase significantly as the attacker must adjust the attack payload differently each time when attacking an instance. We conducted experiments using real-world programs such as Intel PX4 Autopilot Drone firmware (ARM 32bit) binary, and Binutils in the Linux system. The results demonstrated our diversification did not break any original semantics of the program yet successfully changed the memory layout for the attacker.

8. ACKNOWLEDGMENT

This work was supported by the Sungshin Women's University Research Grant of 2022.

9. REFERENCES

[1] K. S. Lhee, S. J. Chapin, "Buffer overflow and format string overflow vulnerabilities", Software: practice and experience, Vol. 33, No. 5, 2003, pp. 423-460.

[2] J. C. Foster et al. "Buffer overflow attacks", Syngress, Rockland, CA, USA, 2005.

[3] S. Govindavajhala, A. W. Appel, "Using memory errors to attack a virtual machine," Proceedings of the Symposium on Security and Privacy, Berkeley, IL, USA, May 2003, pp. 154-165.

[4] A. Francillon, D. Perito, C. Castelluccia, "Defending embedded systems against control flow attacks", Proceedings of the 1st ACM workshop on Secure execution of untrusted code, Chicago, USA, November 2009, pp. 19-26.

[5] A. Gupta, S. Kerr, MS. Kirkpatrick, E. Bertino, "Marlin: A fine grained randomization approach to defend against ROP attacks", Proceedings of the International Conference on Network and System Security, 3-4 June 2013, pp. 293-306.

[6] B. A. Kuperman, C. E. Brodley, H. Ozdoganoglu, T. N. Vijaykumar, A. Jalote, "Detection and prevention of stack buffer overflow attacks", Communications of the ACM, Vol. 48, No. 11, 2005, pp. 50-56.

[7] P. O. Sullivan, "Preventing Buffer Overflows with Binary Rewriting", University of Maryland, Electrical Engineering, College Park, Md, PhD Thesis, 2010.

[8] Z. Jin, Y. Chen, T. Liu, K. Li, Z. Wang, J. Zheng, "A novel and fine-grained heap randomization allocation strategy for effectively alleviating heap buffer overflow Vulnerabilities", Proceedings of the 4th International Conference on Mathematics and Artificial Intelligence, Chegndu, China, April 2019, pp. 115-122.

[9] P. Wang, J. Zhang, S. Wang, D. Wu, "Quantitative Assessment on the Limitations of Code Randomization for Legacy Binaries", Proceedings of the IEEE European Symposium on Security and Privacy, Genoa, Italy, September 2020, pp. 1-16.

[10] M. Prasad, T. Chiueh, "A Binary Rewriting Defense Against Stack based Buffer Overflow Attacks", Proceedings of the USENIX Annual Technical Conference, San Antonio, TX, USA, 9-14 June 2003, pp. 211-224.

[11] V. Iyer, A. Kanitkar, P. Dasgupta, R. Srinivasan, "Preventing overflow attacks by memory randomization", Proceedings of the IEEE 21st International Symposium on Software Reliability Engineering, San Jose, CA, USA, November 2010, pp. 339-347.

[12] S. Xu, P. Xie, Y. Wang, "AT-ROP: Using static analysis and binary patch technology to defend against ROP attacks based on return instruction", Proceedings of the International Symposium on Theoretical Aspects of Software Engineering, Hangzhou, China, December 2020, pp. 209-216.

[13] K. Chen, Y. Lian, Y. Zhang, "Automatically generating patch in binary programs using attribute-based taint analysis", Proceedings of the International Conference on Information and Communications Security, December 2010, pp. 367-382.

[14] U. Müller, E. Hauck, T. Welz, J. Classen, M. Hollick, "Dinosaur Resurrection: PowerPC Binary Patching for Base Station Analysis", Proceedings of the Network and Distributed System Security Symposium 2021, February 2021, pp. 21.

[15] J. Pewny, B. Garmany, R. Gawlik, C. Rossow, T. Holz, "Cross-architecture bug search in binary executables", Proceedings of the IEEE Symposium on Security and Privacy, May 2015, pp.709-724.

[16] M. Huang, C. Song, “ARMPatch: A Binary Patching Framework for ARM-based IoT Devices”, Journal of Web Engineering, Vol. 20, No. 6, 2021, pp. 1829-1852.

[17] P. Sun, Q. Yan, H. Zhou, J. Li, "Osprey: A fast and accurate patch presence test framework for binaries", Computer Communications, Vol. 173, No. 9, 2021, pp. 95-106.

[18] Z. Xu, B. Chen, M. Chandramohan, Y. Liu, F. Song, "Spain: security patch analysis for binaries towards understanding the pain and pills", Proceedings of the 39th International Conference on Software Engineering, Buenos Aires, Argentina, May 2017, pp. 462-472.

[19] L. Zhao, Y. Zhu, J. Ming, Y. Zhang, H. Zhang, H. Yin, "Patchscope: Memory object centric patch diffing", Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, USA, November 2020, pp. 149-165.

[20] S. Lee, H. Kang, J. Jang, BB. Kang, “Savior: Thwarting stack-based memory safety violations by randomizing stack layout”, IEEE Transactions on Dependable and Secure Computing, Vol. 19, No. 4, 2021, pp. 2259-2575.

[21] M. Backes, S. Nürnberger, "Oxymoron: Making {Fine-Grained} Memory Randomization Practical by Allowing Code Sharing", Proceedings of the 23rd USENIX security symposium, Sandiego, CA, USA, August 2014, pp. 433-447.

[22] R. Nikolaev, H. Nadeem, C. Stone, B. Ravindran, “Adelie: Continuous Address Space Layout Rerandomization for Linux Drivers“, Proceedings of the 27th ACM International Conference on Architectural Support for Programming Languages and Operating System, Lausanne, Switzerland, February 2022, pp. 483-498.

[23] R. Wartell, V. Mohan, K. W. Hamlen, Z. Lin, "Binary stirring: Self-randomizing instruction addresses of legacy x86 binary code", Proceedings of the ACM conference on Computer and communications security, Raleigh, NC, USA, October 2012, pp. 157-168.

[24] A. Biswas, Z. Li, A. Tyagi, "Performance Counters and DWT Enabled Control Flow Integrity", SN Computer Science 3.1, Vol. 3, No. 48, 2022, pp. 1-19.

[25] M. Singh, “Data Structure Alignment: How data is arranged and accessed in Computer Memory?”, https://www.geeksforgeeks.org/data-structure-alignment-how-data-is-arranged-and-accessed-in-computer-memory (accessed: 2021)

[26] ARM Developer, “Cortex-R4 and Cortex-R4F Technical Reference Manual r1p3”, https://developer.arm.com/documentation/ddi0363/e/memory-protection-unit/mpu-faults/alignment-fault (accessed: 2021)

[27] Github, “Capstone Engine”, https://github.com/capstone-engine/capstone (accessed: 2021)

[28] Microsoft, “Align (C++)”, https://docs.microsoft.com/en-us/cpp/cpp/align-cpp?view=msvc-170 (accessed: 2021)

[29] Github, “patchkit”, https://github.com/lunix-bochs/patchkit (accessed: 2021)

[30] Github, Radare2: Libre Reversing Framework for Unix Geeks, https://github.com/radareorg/radare2 (accessed: 2021)

[31] Github, “Keystone Engine”, https://github.com/keystone-engine/keystone (accessed: 2021)

[32] W.J. Hengeveld, “Rapi tools”, https://itsme.home.xs4all.nl/projects/xda/tools.html (accessed: 2022)

[33] WINEHQ, “What is Wine?”, https://www.winehq.org/ (accessed: 2022)

Energy Efficient Multi-hop routing scheme using Taylor based Gravitational Search Algorithm in Wireless Sensor Networks

Original Scientific Paper

Sivasankari B
Department of Electronics and Communication Engineering,
SNS College of Technology,
Coimbatore, Tamil Nadu, India- 641035
sivasa.b.ece@snsct.org

Dharavath Champla
Department of Computer Science and Engineering,
Anna University, Chennai, Tamil Nadu,
India-600 025, champla.805@gmail.com

Pushpavalli M
Department of Electronics and Communication Engineering,
Bannari Amman Institute of Technology,
Coimbatore, Tamil Nadu, India- 641035
Sathyamangalam, India-638401
sivasa.b.ece@snsct.org, pushpavallim@bitsathy.ac.in

Ahilan A
Department of Electronics and Communication Engineering,
PSN College of Engineering and Technology,
Tirunelveli, India-627152,
listentoahil@gmail.com

Abstract *– A group of small sensors can participate in the wireless network infrastructure and make appropriate transmission and communication sensor networks. There are numerous uses for drones, including military, medical, agricultural, and atmospheric monitoring. The power sources available to nodes in WSNs are restricted. Furthermore, because of this, a diverse method of energy availability is required, primarily for communication over a vast distance, for which Multi-Hop (MH) systems are used. Obtaining the optimum routing path between nodes is still a significant problem, even when multi-hop systems reduce the cost of energy needed by every node along the way. As a result, the number of transmissions must be kept to a minimum to provide effective routing and extend the system's lifetime. To solve the energy problem in WSN, Taylor based Gravitational Search Algorithm (TBGSA) is proposed, which combines the Taylor series with a Gravitational search algorithm to discover the best hops for multi-hop routing. Initially, the sensor nodes are categorised as groups or clusters and the maximum capable node can access the cluster head the next action is switching between multiple nodes via a multi-hop manner. Initially, the best (CH) Cluster Head is chosen using the Artificial Bee Colony (ABC) algorithm, and then the data is transmitted utilizing multi-hop routing. The comparison result shows out the extension of networks longevity of the proposed method with the existing EBMRS, MOGA, and DMEERP methods. The network lifetime of the proposed method increased by 13.2%, 21.9% and 29.2% better than DMEERP, MOGA, and EBMRS respectively.*

Keywords*: ABC algorithm, Energy efficiency, Multihop routing, WSN, Taylor series*

1. INTRODUCTION

WSNs offer a wide range of uses, including disaster management, environmental control, and surveillance. They are most commonly utilized in regions where humans cannot access them. A WSN [1] is a group of several small nodes linked wirelessly for smooth transmission. These sensors' processing capacity, memory, energy, and data communication range are all constrained. Furthermore, these sensor nodes often send data via wireless radio transmission [2]. The sensors achieve their goal of autonomous event detection and data collection. A WSN's backbone is made up of low-cost sensors. A multi-hop short-distance route may be a more energy-efficient method for transmitting a message than a single-hop long-distance route [3]. The lifetime of a WSN is one of the most critical elements in determining its efficacy.

The lifespan of a network can be effectively extended by balancing the node's energy usage and enhancing energy efficacy [4]. The CH is the vital node in the clustering protocol, and it is in charge of gathering the data experienced by the node members and relaying it to the sink node. Node members only have to interact with their cluster leaders over a small distance, saving energy [5,6]. There are various options for connecting the CH to the base station, including sending data directly to the base station (BS) or using additional nodes as a next-hop. [7].

Clustering protocols strive to identify the best CH set and interchange the function of CH across all nodes to balance the node's energy consumption. Metaheuristic and heuristic algorithms are commonly utilized in optimal CH selection as efficient techniques to obtain workable solutions with variable degrees of success. The ABC algorithm is an intelligent swarm technique based on metaheuristics stimulated by bees' natural honey-gathering behavior [8,9]. In contrast to other forms of swarm intelligence, the ABC algorithm may obtain good optimization results by balancing local exploration and global development. It has a small No. of parameters and is simple to implement simultaneously.

In the hierarchical clustering technique, the secure transmission phase is the last phase of each round. The node members communicate data that has been perceived to the associated CH in this phase, and the CH subsequently transmits the acquired data to the BS [10, 11]. There is a narrow gap between node members and their CHs, and intra-cluster communication is achieved with a single hop. The distance between the CHs and the BS, on the other hand, is usually quite large. If only a single hop is used for communication, the cluster heads' unnecessary energy consumption will be exacerbated [12, 13]. Therefore, hierarchical clustering protocols often necessitate a routing algorithm to determine the optimum path amongst each CH and the BS, data can be transmitted to it. When the CH is far away from the BS, it can avoid consuming too much energy. Multi-hop routing [14] aids network routing over the range of communication, and the energy factor governs that. The latency has been reduced, but the consumption of energy is immense. Thus, the routing must conserve energy. As a result, academics are working on an energy-aware routing strategy.

The major contribution of this paper are as follows:

- The main objective of this research project is to use hybrid optimization techniques to create a WSN multi-hop routing protocol with less energy consumption.
- For CH selection, an innovative and energy-efficient ABC algorithm is used.
- An energy-efficient routing method that relies on the Taylor-based Gravitational search algorithm is proposed to identify the best path from each cluster head and base station.
- The proposed strategy can successfully lower network energy consumption, increase throughput, and lengthen network lifetime, according to simulation results.

The remaining paper is ordered in a resulting manner. The concepts of various multipath routing protocols are discussed in Section 2, and the assumptions underlying the System model are presented in Section 3. The proposed TBGSA's many phases are discussed in Section 4. Section 5 examines the suggested method's performance and compares it to existing protocols. Finally, Sect. 6 brings the paper to a conclusion.

2. RELATED WORK

A Secure Energy-Aware Routing (SEAR) technique was proposed by Singh et al. [15]. SEAR discovered malicious nodes based on trust evaluation. For this routing, a multi-factor technique was evaluated, including node trust value, hop length, and residual energy. Using trust nodes, SEAR managed the network's energy consumption and data transmission. The protocol is not designed to be scalable.

In [16] author presented an Energy-Efficient Routing Protocol (EERP) for WSN, named an A-star algorithm. This routing approach increased network lifetime via passing data packets along the shortest route. According to the next-hop sensor node's highest remaining energy, buffer occupancy, a high level of connection quality, and a minimal number of hops, the optimum path was identified. However, when a network is enormous, its performance decreases.

In [17] author proposed an energy-saving routing strategy that included clustering and sink mobility. The entire sensor field was first divided into sectors, with each sector electing a CH depending on the weight of its members. To determine the optimal scenario, node members calculated the energy consumption of several routing paths. The most significant disadvantage is that the MS location must be transmitted regularly via sensors, which may increase the network load.

In [18] author presented a multi-hop routing protocol for data routing in wireless sensor networks. The high overhead was reduced by designing a green routing protocol. The lifespan of the network can be greatly extended and its overhead decreased by using an energy-efficient protocol. Utilizing relay nodes, inter-cluster broadcasts were used in this approach to distributing accumulated cluster data. As a result, the WSN's scalability was increased, and the employment of relay nodes had a favorable influence while dissolving energy in the WSN. This strategy did not perform well with massive networks.

In [19] author presented the Balanced and Energy Efficient Multi-Hop (BEEMH) algorithm for multi-hop routing in WSN. The Dijkstra algorithm was used to create this approach. This method sparked much interest in the residual energy of nodes. Thus, the transmitter and receiver nodes were chosen among the nodes with the highest energy. The method offered a useful framework for maximizing the choice of cluster heads based on several criteria, including location and energy. Low performance was the result since the technique did not optimize the grid regions and affected the nodes' dependability of connection.

In [20] author presented an energy-aware routing scheme to conserve the energy of networks. Initial energy, total network energy, and residual energy were all considered. Nodes closer to the base station were precluded from creating a cluster and were segregated.

Throughout the data transmission process, by comparing and evaluating different metrics, the energy usage of a single-hop route was compared to a multi-hop route. Mostly the approach that conserved the least quantity of energy was chosen. The network's longevity was prolonged while the cost of communication was cut.

In [21] author presented a new multicast routing technique that reduces time while increasing packet delivery ratio and bandwidth. Nodes were established in the network environment using multi-hop pathways to efficiently forward packets without incurring an additional loss. The most critical parameters are the bandwidth and multi-hop distance between CH and cluster members. Multi-path transmission removed collisions and false detections via Clear to Send (CTS) and Request to Send (RTS) mechanisms. The pathways were stabilized to reduce crashes.

In [22] author presented an Energy Balanced Multihop Routing Scheme (EBMRS) to find the best path to the BS, while increasing energy efficiency at the inter and intra-clustering levels equally. It chooses the essential no of sensor nodes as CHs depending on PS parameters such as residual energy, n.o of hops, and BS's distance. In respect of network throughput and transmission delay, it improves WSN performance. The protocol is not designed to scale.

In [23] author introduced a CS-based system that uses Multiple Objective Genetic Algorithms (MOGA) to optimize the number of measurements, the sensing matrix, and the transmission range in order to transfer data efficiently between WSNs. The program seeks to achieve the best possible energy efficacy and precision balance. It creates a multi-hop path depending on the optimum values. The simulation findings showed that, while MOGA improves network coverage quality, it also efficiently decreases the depletion of node energy, effectively extending the network's life cycle.

In [24] author presented a multipath routing protocol that is QoS sensitive and uses a hybrid particle swarm optimization-cuckoo search technique to cluster sensor nodes. The protocol then chose multiple trustworthy paths for data transfer based on multi-hop communication using the Cluster Heads (optimal network routing). In contrast to standard protocols, it relies on fast data transit over channels that do not affect QoS (Quality of Service). In contrast to conventional QoS Centric protocols, it also extended the network lifetime by switching CHs regularly depending on residual energy and using the optimal channels for data transmission.

In [25] author presented a Multi-hop Energy Efficient Routing Protocol (DMEERP). It is divided into three pieces. Each record of the CH and members of the cluster were held and maintained by the Super Cluster Head (SCH). If the present CH fails, the activation node and weight factor are approximated to get a new CH. The path reliability ratio is calculated to route packets swiftly while minimizing packet loss. The energy model is founded on the capacity of the channel concept. DMEERP delivered a high data delivery ratio, low overhead, the ratio of the path reliability, minor delay, extended network longevity, and low consumption of energy, according to simulation results. The computational time of this method was high.

Recent techniques have shown that sink mobility in a regulated path can improve energy efficiency in WSNs, but the path limits make routing more difficult. As a result, the sink with a set speed takes less time to collect data through sensor nodes that are randomly placed. This constraint has severe implications for improving data collection and reducing energy consumption. Furthermore, these research limits influenced the development of WSN's design proficient path optimization technique.

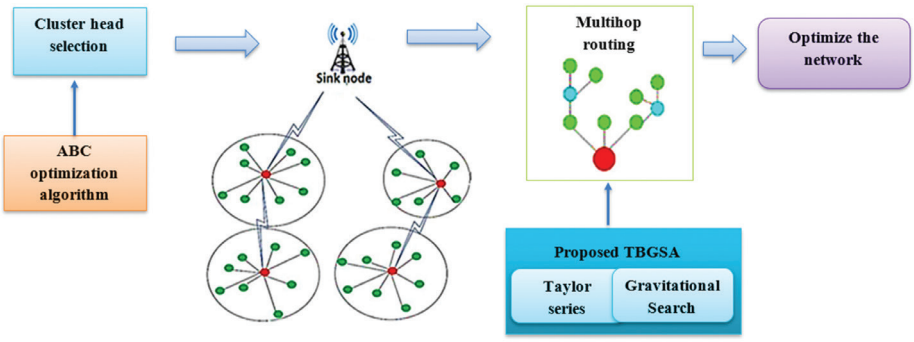


Fig. 1. Framework of TBGSA Model

3. SYSTEM MODEL

The sensor nodes are designated as $\{a_1, a_2 \ldots a_N\}$ in the network topology. In a WSN, nodes are arbitrarily distributed in an area x M m^2. M = 100 and N = 100 are used in this paper. All nodes have a limited radio range, and each has a sensor radius of d0 meters. Normal nodes, sink nodes (SN), and CH nodes are the three types of nodes found here. A clustering strategy is used in network communications, where a No. of nodes is chosen as CHs to gather data from ordinary nodes and transfer it to the SN. Fig. 1 shows an, e.g., of a WSN topology with 3 kinds of nodes. The SN is supposed to be in the middle of the area and have a limitless energy supply, whereas the sensor node's energy is inadequate in this paper. Positions of every node are considered to be static, & their

beginning energy is equal. Furthermore, every node is GPS-enabled, and its locations are known every time.

The quality of the title and the ability to derive from it keywords useful for cross-referencing and computer searches are crucial for indexing and abstracting services. A document with an inappropriate title may never be understood by the intended readership or be focused.

3.2. ENERGY MODEL

The energy required to send E_{Ty} and receive E_{Ry} of an m-bit message across a distance of x is computed using equation (1) and equation (2).

$$E_{Ty}(m,x) = E_{el} * m + \varepsilon_{amp} * m * x^{i} \quad (1)$$

$$E_{Ry}(m,x) = E_{el} * m \quad (2)$$

Where $i = \begin{cases} 2, & if\ x < x_0 \\ 4, & else \end{cases}$ and ε_{amp} is the amplified energy, E_{el} is the required energy to operate the transmitting and receiving circuits, and the value of x_0 is computed as equation (3).

$$x_0 = \sqrt{\frac{\varepsilon_f}{\varepsilon_m}} \quad (3)$$

where ε_m and ε_f stand for multi-path amplification and free space energy, respectively. Furthermore, each nominated parent node expends energy E_p to forward the m-bit long aggregated data of their 'n' child nodes to its parent node, which is worked out as following equation (4).

$$E_p = n * m * E_{el} + n * E_a * m + m * E_{Ty} \quad (4)$$

where E_a is the amount of energy spent on data aggregation by each parent node. In addition, each path uses $E\ (Pa)\ energy$, which is computed as following equation (5).

$$E(Pa) = \sum_{i=1}^{p} E_{i,n_{p(i)}}\ (m_i, x_i) \quad (5)$$

where $n_{p(i)}$ is node i's next hop on path P, and $E_{i,n\_p(i)}$ is the energy between node i and node $n_{p(i)}$.

3.3. PROBLEM STATEMENT

Many researchers in wireless sensor networks regard data transfer from one sensor node to another as a critical difficulty. Finding the shortest path improves data transmission speed and node consumption of energy, which is a crucial matter in wireless sensor networks. As a result, the main issue is to transport data utilizing the shortest path possible and a better routing system. Because the energy consumption of nodes varies depending on their functions and network placements, the routing protocol must be capable of balancing the energy consumption of the nodes. The distances between nodes and base stations are typically large in a WSN. Data transmission over a vast distance will devour a significant amount of energy. The routing protocol must reduce the amount of energy used by nodes transmitting data to the base station. How to effectively form many network nodes to decrease node energy indulgence, keep node energy consumption in check, and decrease the energy dissolution of data transmission from sensor nodes to the BS. These are all issues that must be discussed in implementing an energy-efficient routing protocol for WSNs. This research work is to use a hybrid optimization approach to design an energy-aware multi-hop routing protocol for a wireless sensor network.

4. TAYLOR SERIES BASED GRAVITATIONAL SEARCH ALGORITHM (TBGSA)

The primary goal of this research work is to use optimization in a hybrid way to design a WSN multi-hop routing protocol with less energy consumption. Two actions are performed in this work: selecting the CH and multi-hop routing. For the selection of CH, the best CH is selected using the ABC algorithm [5], and then the data is transmitted utilizing multi-hop routing. Therefore, the optimal hop for broadcasting data begins with the optimal placement of hops, which is accomplished by the suggested Taylor-based Gravitational search algorithm.

4.1. CH SELECTION USING THE ABC ALGORITHM

The ABC algorithm is a swarm-based AI algorithm that stimulates honey bees' smart searching activity. The artificial colony of the bees of the ABC algorithm has 3 types of bees: onlookers, scouts, and worker bees, each of which denotes a location in the searching pane. The honeybees fly in the searching panel depending on the 'n' dimensions if the WSN comprises n' CH sensors. The ABC uses the bee population to locate the CHs. In the area for dance, an onlooker is a bee that stays there and chooses a food source, whereas an employed bee travels to the source of food that it previously visited. The scout is a bee that is responsible for our random search. The presence of food suggests that there may be a solution to the optimization process. The quality (fitness) of the related solution is proportional to the amount of honey in the food supply.

Worker bees would be in the first part of the colony in this algorithm, whereas spectator bees are in the second quarter. The first placements for food sources are generated at random, all employed bees are assigned to a specific source of food. Every employed bee finds a new nearby source of food from the previously related source of food, and the amount of nectar from the new source of food is calculated for each round, using equation (6).

$$S_{ij} = C_{ij} + \theta(C_{ij} - C_{kj}) \quad (6)$$

where the random number is denoted as $\theta \in [1, 1]$, the candidate solution is denoted as S_i, the present solution is denoted as C_i, and C_k is a neighbor solution, and $j \in \{1, 2,..., d\}$ is an index picked at random, and the letter d denotes the solution vector's dimension. When all of the employed bees & onlooker bees have completed

the search procedure, they share data about their food sources. The onlooker bee analyses information about nectar from all employed bees. It then selects a food sourced using the roulette wheel selection approach, based on a probability connected to its amount of nectar by equation (7), which gives the best candidates a higher chance of being chosen.

$$p_i = \frac{fi_i}{\sum_{n=1}^{A} fi_n} \tag{7}$$

where fi_i is the fitness value of solution i, and A is the No. of sources of food corresponding to the employed bees.

When all food sources have been chosen, each onlooker creates a new nearby source of food is chosen, and the amount of nectar is determined. If nectar levels are higher, the bee does not memorize the old location and does not remember the new one; otherwise, it keeps it. Both employed and onlooker bees use the same food source. The employed bee is transformed into a scout. As abandoned the food supply is allotted, the employed bee of that source seems to be a scout, and any location cannot be improved by more than a specific number of cycles, known as the limit parameter. The scout, mentioned in equation (8), produces a new solution in that position at random, where the abandoned source is denoted by C_i in addition to $j \in \{1, 2,..., d\}$.

$$C_i^j = C_{min}^j + Rand(0,1)(C_{max}^j - C_{min}^j) \tag{8}$$

During the cluster setup stage, the ABC is utilized to select the CHs depending on the dependability of clusters being reduced to improve the network's longevity and period of stability. The work begins, and many clusters emerge around the base station considered a base station. The suggested work selects the nodes by taking into account their distance from the BS and their alignment with the numerous CHs. When the base detects a node with sufficient energy, it transmits the resultant data to the remaining no of nodes. The network's quality is calculated using a fitness function. The fitness function is derived using equation (9).

$$f = \sum_i \alpha(w_i, f_i) \forall f_i \in \{N_H, \lambda_o, E_{Rq}\} \tag{9}$$

where N_H denotes the n.o of hops to the sink, and λ_o denotes the total amount of traffic transmitted. E_{Rq} is the required energy.

$$\lambda_o^i = \lambda_1^i + \lambda_g^i \tag{10}$$

In the above equation (10), the number of random nodes that make up one cluster represents the size of the hired hive. The fitness function f_{im} is used to assess the nectar (network) quality using equation (11).

$$fi_m = \begin{cases} \frac{1}{1+f_m(x_m)}, & f_m(x_m) > 0 \\ 1 + |f_m(x_m)|, & f_m(x_m) < 0 \end{cases} \tag{11}$$

ABC, which follows the natural evolution process, assesses an individual's fitness. The ABC uses a search technique in which an arbitrary option guides the search depending on a parameter approach. Algorithm 1 discusses [26] how CHs are selected using the ABC algorithm. The management information system [6] includes significant elements related to ethics and information security [3-5], as well as computer security and computer ethics.

Algorithm 1. CH Selection using ABC

For each CH 'h' do

For all node 'j' do

Broadcast "hello" message

Obtain degree

e= {(p.s)⊂ S/D (p, s)≤R}

deg(p)=|e|

obtain consumed energy 0.05j<E(j)/E_0≤0.5j

calculate the distance to BS

$p = \sqrt{(x - x_{bs})^2 + (y - y_{bs})^2}$

calculate weight

wt=1/(Deg(wt)+0.05<E(wt)/E_0 ≤0.5j+p/max)-dist

end

if wt==min_wt(wt) then

state=h

broadcast" CH accepted"

else

state(wt)=on

end

end

4.2. PROPOSED TBGSA

WSN's multi-hop routing has improved communication across the network. For efficient data transmission, multi-hop routing is commonly used. However, with multi-hop routing, energy is the most significant barrier. As a result, the proposed Taylor-based gravitational search algorithm uses multi-hop routing to resolve energy issues. To reduce transmission energy, each cluster source node distributes its data to its neighbor node. The suggested TBGSA calculates the best hops for WSN routing progression using a newly created fitness function. Predicting the linear element of the equation is done by using the Taylor series to characterize previously recorded data. The Taylor series has the benefit of being a simpler and easier way to compute solutions, even in the presence of complex functions. The Taylor series provides a number of benefits, including straightforward convergence and reliable assessment of mutual functions.

The communication paths between clusters and the SN are set during this stage taking the chosen CHs into

account. Each cluster's CHs are regarded as particles in a gravitational search algorithm. Due to gravitational search, the particles in the packet are balanced at each stage of transmission. It discards those of poor quality while other particles are chosen as suitable solutions. Based on the objectives, the selected particles' weight is established. According to the proposed approach, TBGSA determines whether the route between the sensor nodes and sink nodes, which call for information transmission, decreases energy consumption and ETE delay while increasing data delivery speed and network throughput. The ideal routes improve even one of these criteria, whereas the abandoned routes are those that deteriorate even one of these measures. As a result, the suggested TBGSA's fitness function is given in equation (12).

$$f_{gr} = \min\left(\sum_{i=1}^{n}\sum_{j=1}^{n} de_{i,j} + \sum_{i=1}^{n} E_{co} - \sum_{i=1}^{n} dd - \sum_{i=1}^{n} tput\right)$$

Where, $\sum_{i=1}^{n}\sum_{j=1}^{n} de_{i,j} > 0, \sum_{i=1}^{n} E_{co} < E_{in}, \sum_{i=1}^{n} dd > 0, \sum_{i=1}^{n} tput > 0$ (12)

where f_{gr} is the fitness function of TBGSA, de_{ij} is the network's ETE delay, E_{co} is the network's total consumption of energy, E_{in} is the network's beginning energy, dd denotes the delivery rate of data at every node, and tput denotes the throughput of the network. Every population with a minimal fitness value can be chosen as a feasible information transmission solution based on the fitness function. The proposed TBGSA workflow shows in Fig. 2.

Algorithm 2. TBGSA for optimal route selection

Input: Nodes (N=N1, N2, ...Nn)

Output: Optimal route (R)

Start

Initialise the random number of nodes

{Ni==N1, N2, ...Nn}

for (i=1,2,...n iterations)

{

Clustering and cluster head using ABC

Transmit data to optimal CHs

Calculate G

$G(r)=G_0 e^{(-\frac{\beta r}{T})}$

Update the Route

if G(r)=R

{

Route identified by TBGSA

Choose optimal route

}

Else

Discard solution

}

End

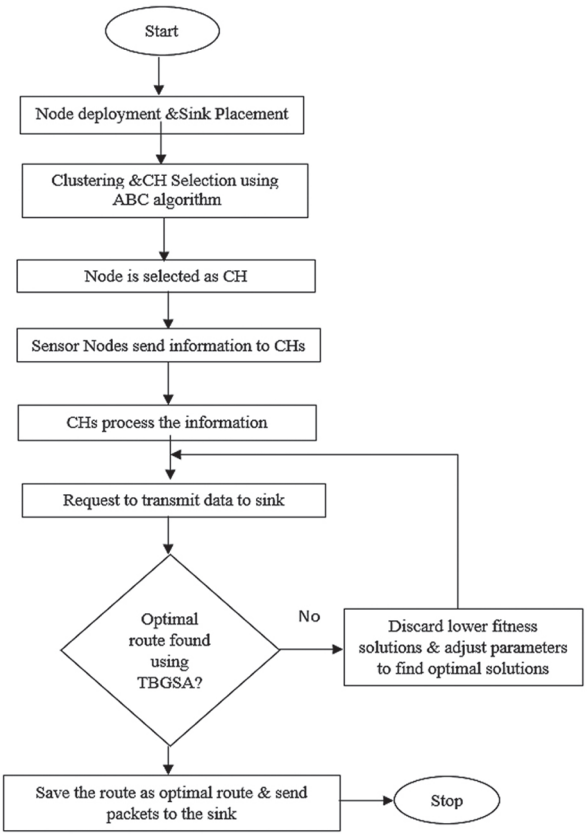


Fig. 2. Flowchart for Proposed TBGSA

5. RESULTS AND DISCUSSION

The WSN must first be configured using conventional parameters to use the proposed method. The suggested network is put into action in a real-world setting. MATLAB 2021a is used to implement this situation.

Table 1. Initial Parameters

Parameters	value
Dimension of the network	100 x 100 m
No of nodes	100
Initial node energy	0.5J
Consumption of energy for transmitting data	5×10^{8} J
Consumption of energy for receiving data	5×10^{8} J
Consumption of energy for routing packet transmission	1×10^{10} J
Consumption of energy for routing packet reception	13×10^{13} J
Consumption of energy for aggregation of data	5×10^{9} J
Maximum n.o of iterations	3500 rounds
Length of the data packet	4000 bits
N.o of transmissions at each hop	10
Packet length for routing	100 bits
Radio range	5000 m

The WSN is simulated using the primary parameters, as indicated in Table 1 and Fig. 3. The WSN is made up of 100 nodes which are dispersed throughout the n/w at random. The SN is located near the heart of the network, making it easy to reach it.

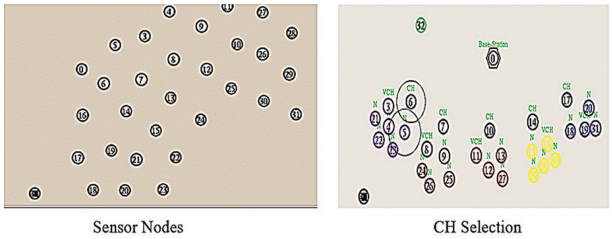


Fig. 3. Sensor Nodes and CH Selection

Equation 9 describes the fitness function, TBGSA evaluates a route's quality, and the routes that better the n/w goals are responded with their fitness value. Fig. 4 depicts the TBGSA's convergence to the optimal point. Because the fitness function used is a minimization function, the TBGSA decreases the value of the objective function at every step to converge to the ideal point, as shown in Fig. 4.

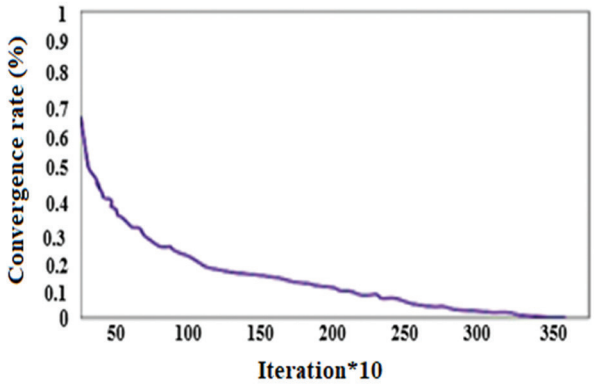


Fig. 4. TBGSA's Convergence to the Optimal Point

The effectiveness of a proposed approach is evaluated in order to raise performance in relation to the main issue. The literature, which is introduced in light of the study goals listed in the first section, offers a number of metrics for evaluating WSNs. Energy consumption, number of dead nodes, number of missed packets, End-to-End (ETE) delay, data delivery rate, & network throughput are all analyzed for the proposed technique.

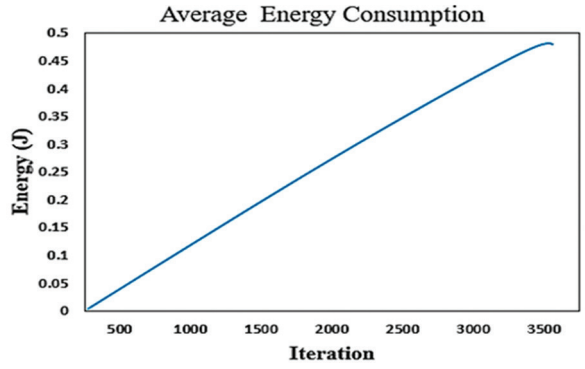


Fig. 5. Average Energy Consumption

The slope of energy used in the nodes is linear, as illustrated in Fig. 5, showing that the usage of energy used for all the nodes in the network during various iterations. As a result, specific nodes may not run out of energy before others, and at around the same time, all nodes run out of energy. As a result, all nodes progressively lose energy, showing that the network has a lengthy lifespan.

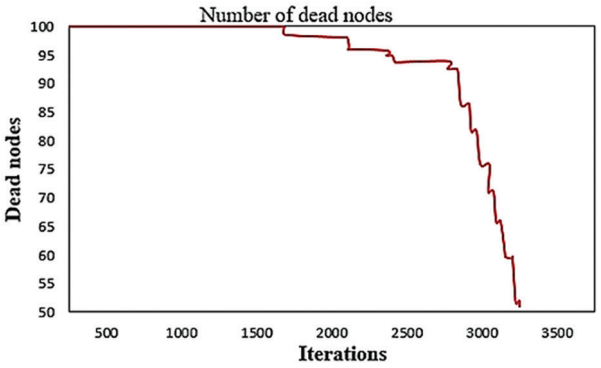


Fig. 6. Dead Nodes in the Network

Fig. 6 depicts the mortality procedure for sensor 500 nodes in the network. An alternate channel for information transmission may be established if a node fails, but the data from that area cannot be gathered. If the death of a node does not disrupt the network, it can be managed. It happened with the suggested technique in the 3020th iteration when nearly most of the nodes ran out of energy & the network was disrupted. This enhancement is because when CH selection is optimized under various characteristics, TBGSA achieves more remarkable energy preservation than other methods.

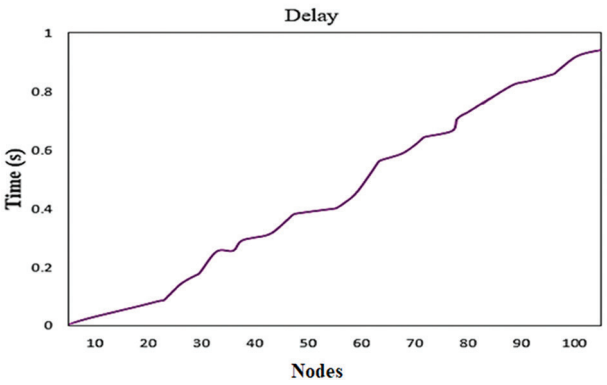


Fig. 7. ETE Delay of 100 Nodes

The network nodes' end-to-end (ETE) delay is the final metric to be assessed. The distance among the nodes is the primary cause of the ETE delay in transmission because the transmission time factor of a packet is static for the nodes. Because information is exchanged among nodes & the CH in the proposed method, the smaller mean intra-cluster distance and a shorter distance among the Cluster head and other nodes show accuracy in clustering, one of TBGSA's goals. Fig. 7 shows that the suggested technique has a 100 ms latency for hundred nodes & 3500 transmission iterations, proving a proposed method's high accuracy in clustering.

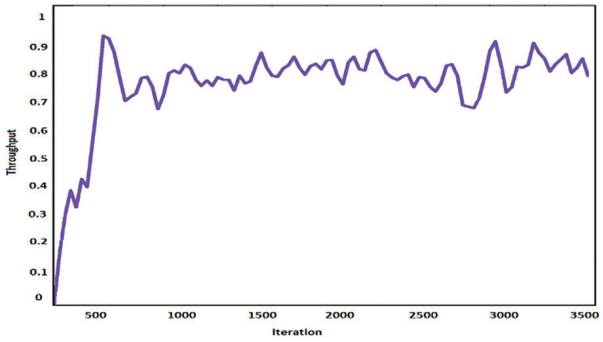


Fig. 8. Throughput of the Network

The throughput is calculated as the BW of the communication medium divided by the number of packets transferred per unit time. The network throughput increased in increasing order, reaching 100% at the end, as shown in Figure 8.

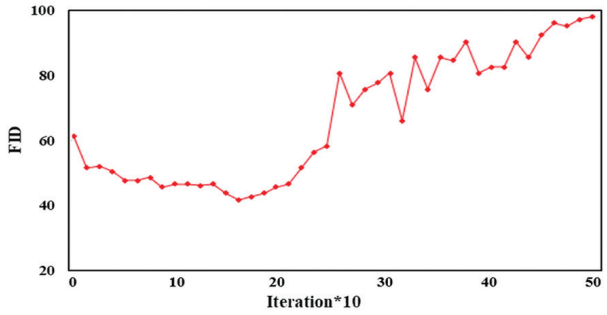


Fig. 9. Frechet Inception Distance

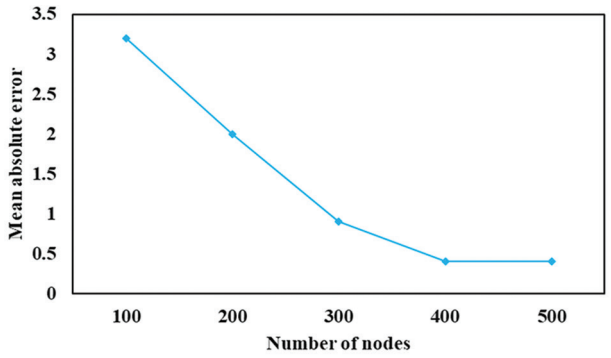


Fig. 10. Mean absolute error

Finally, we observe that the simulation results are most closely mateched with the analystical results for all distribition as the number of node increases. Acoording to the figure 9 FID improves up to 19k iterations. Figure 10 shows the results obtained for Mean absolute error. Mean absolute error (MAE) is a measure of errors between paired observations expressing the number of nodes.

5.2. COMPARISON WITH EXISTING METHODS

The effectiveness of the suggested strategies was demonstrated by comparing their performance to that of the current strategy. Performance is assessed using the following metrics: throughput, packet delivery rate, end-to-end delay, network lifetime, and energy efficiency. In a comparative study, the suggested model is compared against three existing approaches.

Table 2. Comparison of Energy Consumption

Parameters	N.o of nodes	EBMRS	MOGA	DMEERP	TBGSA (proposed)
Energy Consumption (mJ)	100	80	50	30	20
	200	110	75	55	46
	300	150	85	65	58
	400	165	120	80	76
	500	180	150	95	82

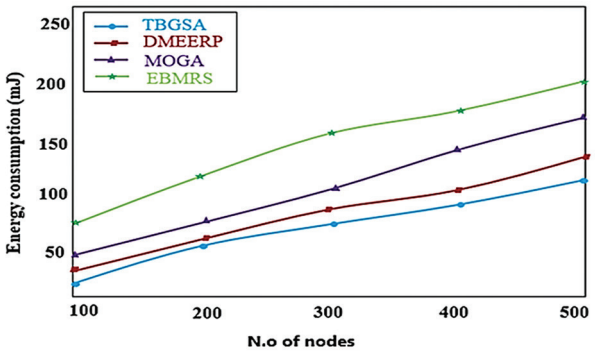


Fig. 11. Comparison of Energy Consumed

Concerning the average energy consumption, Fig. 11 compares the suggested method to existing methods. As a result, compared to all existing methods, the proposed TBGSA model has achieved the lowest energy use. The TBGSA model reduces the energy required to transfer data within the cluster. Because of the random selection of CH, EBMRS achieves a bad performance. These are some reasons why maximal energy dissipation is preferred over other models. The TBGSA model, for example, has a lower energy expenditure of 75mJ when the maximum node count is 500. In contrast, the EBMRS, MOGA, and DMEERP models are indicated in Table 2. The findings of the aforementioned experiment show that the computational complexity is significantly decreased by the suggested TBGSA approach.

Table 3. Comparison of Data Delivery Rate

Parameters	N.o of nodes	EBMRS	MOGA	DMEERP	TBGSA (proposed)
Data delivery rate (%)	100	97	97.5	98	99.3
	200	96	97	97.5	99
	300	95.5	96.8	97	98.9
	400	94	96.4	97.3	98.6
	500	93.3	96	96.8	98.8

As illustrated in Fig. 12, the suggested method has a greater delivery ratio than the existing methods. When the network size is expanded, the no of packets in TBGSA is always more than that of existing methods, as can be shown. This outcome is explained because the TB-

GSA protocol has a longer lifetime than other models, resulting in a substantial increase in data transmission to the sink node. Furthermore, the TBGSA protocol can expand network longevity and is appropriate for large-scale networks. For example, the TBGSA framework has a maximum DDR of 98.8 % when the node value is 500, but in the EBMRS, MOGA, and DMEERP approaches as indicated in Table 3, the least PDR of 93 %, 96 %, and 96.8 %, respectively.

Table 4. Comparison of ETE Delay

Parameters	N.o of nodes	EBMRS	MOGA	DMEERP	TBGSA (proposed)
End to End delay(s)	100	4	2.8	2	2
	200	5	4.5	2.4	2.2
	300	6.3	5.8	3.2	2.4
	400	7.5	7.3	3.8	2.9
	500	9.2	8.3	5.1	3.2

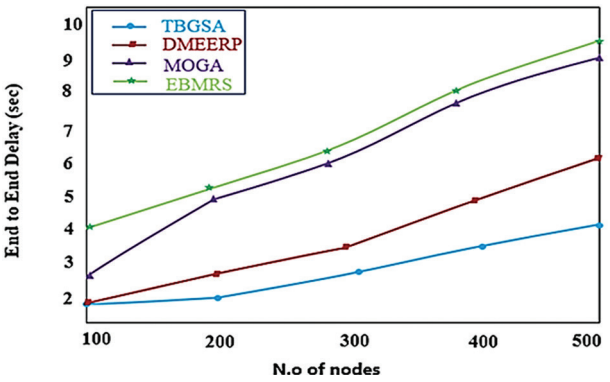


Fig. 13. End-to-End Delay

The ETE delay analysis of the TBGSA method is shown in Fig. 13 alongside a collection of prior approaches. E2E delay saves energy and ensures consistency. The E2E latency is calculated as the total time of data transmission, processing, and data delivery time. Minimum delay improves network reliability while also lowering energy consumption. For example, the TBGSA scheme achieved a minimum ETE delay of 3.2s under a higher node count of 500, the EBMRS, MOGA, and DMEERP frameworks as indicated in Table 4, achieved a maximum of 9.2s, 8.5s and 4.5s, respectively.

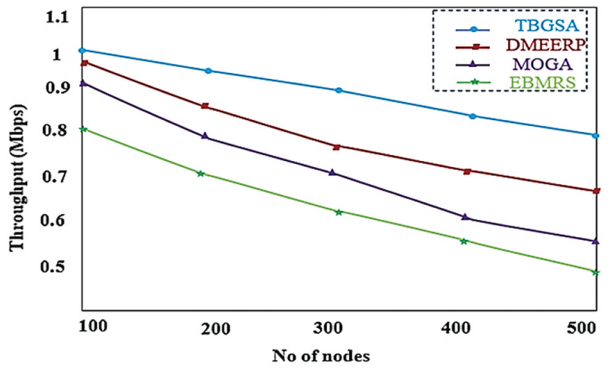


Fig. 14. Throughput

Fig. 14 compares the TBGSA model's throughput analysis to several earlier methodologies. The TBGSA method has shown qualified results with increased throughput. The TBGSA framework, for example, has reached a maximum throughput of 0.79Mbps with a node count of 500, whereas the EBMRS, MOGA, and DMEERP frameworks as indicated in Table 5, achieved low throughput of 0.65, 0.61, 0.53, 0.46, and 0.47 Mbps, respectively. As a result, data from cluster members cannot reach the base station, resulting in a sparse sensing field and limited network performance.

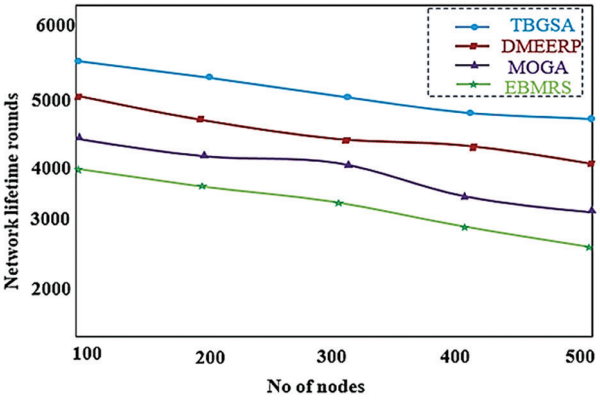


Fig. 15. Network Lifetime

Fig. 15 compares the TBGSA technique's network lifespan analysis to various existing approaches. The TBGSA model has shown superior outcomes with the highest network lifetime. The TBGSA, for example, has a longer network lifespan of 5300 rounds when the node count is 500. However, the EBMRS, MOGA, and DMEERP models have a minimum network lifetime of 3000, 4000, and 4600 rounds. The TBGSA concept aims to extend the network's longevity by boosting energy efficiency.

6. CONCLUSION

In this research present a low-energy multi-hop routing protocol that considers energy as a key factor in multi-hop routing. The technique went through two steps to achieve good multi-hop routing. The Cluster Head selection is done first, followed by the transmission of data. When selecting the CH, the best node is chosen as CH based on the ABC algorithm. Subsequently, data is communicated from one node to another utilizing various hops ideally chosen using the suggested Taylor GSA-based multiobjective fitness function. The improvement in extending the network's longevity was verified by making a comparison with the suggested technique to the findings of cutting-edge research. These results of the research appear to be extremely useful in the development of an energy-efficient multi-hop wireless sensor network, and they open the way for the expansion of commonly used applications in these networks. In the future, efforts should be made to strengthen the protocol's security features.

7. REFERENCES:

[1] M. A. Jawad, F. Khurshid, "A review of approaches to energy aware multi-hop routing for lifetime enhancement in wireless sensor networks", Proceedings of the International e-Conference on Intelligent Systems and Signal Processing, Singapore, 2022, pp. 739-757.

[2] M. Srinivas, T. Amgoth, "EE-hHHSS: Energy-efficient wireless sensor network with mobile sink strategy using hybrid Harris hawk-salp swarm optimization algorithm", International Journal of Communication Systems, Vol. 33, No. 16, 2020, p. e4569.

[3] A. Vinitha, M. S. S. Rukmini, D. Sunehra, "Energy-efficient multihop routing in WSN using the hybrid optimization algorithm", International Journal of Communication Systems, Vol. 33, No. 12, 2020, p. e4440.

[4] J. Yong, Z. Lin, W. Qian, B. Ke, W. Chen, L. Ji-fang, "Tree-based multihop routing method for energy efficiency of wireless sensor networks", Journal of Sensors, Vol. 2021, 2021, pp. 1-14.

[5] A. Rezaeipanah, P. Amiri, H. Nazari, M. Mojarad, H. Parvin, "An energy-aware hybrid approach for wireless sensor networks using re-clustering-based multi-hop routing", Wireless Personal Communications, Vol. 120, No. 4, 2021, pp. 3293-3314.

[6] R. Vinodhini, C. Gomathy, "MOMHR: a dynamic multi-hop routing protocol for WSN using heuristic based multi-objective function", Wireless Personal Communications, Vol. 111, No. 2, 2020, pp. 883-907.

[7] M. Elhoseny, R. S. Rajan, M. Hammoudeh, K. Shankar, O. Aldabbas, "Swarm intelligence–based energy efficient clustering with multihop routing protocol for sustainable wireless sensor networks", International Journal of Distributed Sensor Networks, Vol. 16, No. 9, 2020, p. 1550147720949133.

[8] T. N. Tran, T. Van Nguyen, V. N. Q. Bao, B. An, "An energy efficiency cluster-based multihop routing protocol in wireless sensor networks", Proceedings of the International Conference on Advanced Technologies for Communications, Ho Chi Minh City, Vietnam, 2018, pp. 349-353.

[9] M. B. Yashoda, V. Shivashetty, "Bi-CRS: bio-inspired cluster-based routing scheme for d2d communication in IoT", Proceedings of International Conference on Recent Trends in Computing, Singapore, 2022, pp. 187-199.

[10] A. A. Kamble, B. M. Patil, "Systematic analysis and review of path optimization techniques in WSN with mobile sink", Computer Science Review, Vol. 41, 2021, p. 100412.

[11] R. K. Yadav, R. P. Mahapatra, "Energy aware optimized clustering for hierarchical routing in wireless sensor network", Computer Science Review, Vol. 41, 2021, p. 100417.

[12] M. Adnan, L. Yang, T. Ahmad, Y. Tao, "An unequally clustered multi-hop routing protocol based on fuzzy logic for wireless sensor networks", IEEE Access, Vol. 9, 2021, pp. 38531-38545.

[13] I. Abasıkeleş-Turgut, "Multihop routing with static and distributed clustering in WSNs", Wireless Networks, Vol. 27, No. 6, 2021, pp. 3797-3809.

[14] B.M. Sahoo, T. Amgoth, H.M. Pandey, "Particle swarm optimization-based energy efficient clustering and sink mobility in heterogeneous wireless sensor network", Ad Hoc Networks, Vol. 106, 2020, p. 102237.

[15] O. Singh, V. Rishiwal, L. Kumar, P. Yadav, "Secure energy aware routingin wireless sensor networks", Proceedings of the 4th International Conference on Internet of Things: Smart Innovation and Usages, Ghaziabad, India, 2019, pp. 1-6.

[16] A. Ghaffari, "An energy efficient routing protocol for wireless sensor networks using A-star algorithm", Journal of Applied Research and Technology, Vol. 12, No. 4, 2014, pp. 815-822.

[17] J. Wang Y. Gao, W. Liu, A. K. Sangaiah, H. J. Kim, "Energy efficient routing algorithm with mobile sink support for wireless sensor networks", Sensors, Vol. 19, No. 7, 2019, p. 1494.

[18] K. Cengiz, T. Dag, "Energy aware multi-hop routing protocol for WSNs", IEEE Access, Vol. 6, 2017, pp. 2622-2633.

[19] A. E. Fawzy, M. Shokair, W. Saad, "Balanced and energy-efficient multi-hop techniques for routing in wireless sensor networks", IET Networks, Vol. 7, No. 1, 2018, pp. 33-43.

[20] Y. Liu, Q. Wu, T. Zhao, Y. Tie, F. Bai, M. Jin, "An improved energy-efficient routing protocol for wireless sensor networks", Sensors, Vol. 19 No. 20, 2019, p. 4579.

[21] P. Arivubrakan, S. Sundari, "A protocol-based routing technique for enhancing the quality of service in multi hop wireless networks", International Journal of Innovative Technology and Exploring Engineering, Vol. 8, No. 8, 2019, pp. 1182-1186.

[22] V. K. Arora, V. Sharma, "A novel energy-efficient balanced multi-hop routing scheme (EBMRS) for wireless sensor networks", Peer-to-Peer Networking and Applications, Vol. 14, No. 2, 2021, pp. 807-820.

[23] M. Al Mazaideh, J. Levendovszky, "A multi-hop routing algorithm for WSNs based on compressive sensing and multiple objective genetic algorithms", Journal of Communications and Networks, Vol. 99, 2021, pp. 1-10.

[24] C. Mohanadevi, S. Selvakumar, "A qos-aware, hybrid particle swarm optimization-cuckoo search clustering based multipath routing in wireless sensor networks", Wireless Personal Communications, Vol. 127, 2022, pp. 1985-2001.

[25] V. Nivedhitha, A. G. Saminathan, P. Thirumurugan, "DMEERP: A dynamic multi-hop energy efficient routing protocol for WSN", Microprocessors and Microsystems, Vol. 79, 2020, p. 103291.

[26] S. R, Deepa, "Cluster optimization in wireless sensor network based on optimized Artificial Bee Colony algorithm", IET Networks, Vol. 10, No. 6, 2021, pp. 295-303.

Review of Path Selection Algorithms with Link Quality and Critical Switch Aware for Heterogeneous Traffic in SDN

Review Paper

Muhammad Nura Yusuf
Faculty of Computing, Univerisiti Teknologi Malaysia, Johor 81310, Malaysia and
Department of Mathematical Science, Abubakar Tafawa Balewa University, Bauchi PMB 0284, Nigeria
ymnura@atbu.edu.ng

Kamalrulnizam bin Abu Bakar
Faculty of Computing, Univerisiti Teknologi Malaysia, Johor 81310, Malaysia
knizam@utm.my

Babangida Isyaku
Sule Lamido University, PMB 047 Kafin Hausa, Jigawa State, Nigeria
Babangida.isyaku@slu.edu.ng

Ajibade Lukuman Saheed
Faculty of Computing, Univerisiti Teknologi Malaysia, Johor 81310, Malaysia
saheed2066@graduate.utm.my

Abstract *– Software Defined Networking (SDN) introduced network management flexibility that eludes traditional network architecture. Nevertheless, the pervasive demand for various cloud computing services with different levels of Quality of Service requirements in our contemporary world made network service provisioning challenging. One of these challenges is path selection (PS) for routing heterogeneous traffic with end-to-end quality of service support specific to each traffic class. The challenge had gotten the research community's attention to the extent that many PSAs were proposed. However, a gap still exists that calls for further study. This paper reviews the existing PSA and the Baseline Shortest Path Algorithms (BSPA) upon which many relevant PSA(s) are built to help identify these gaps. The paper categorizes the PSAs into four, based on their path selection criteria, (1) PSAs that use static or dynamic link quality to guide PSD, (2) PSAs that consider the criticality of switch in terms of an update operation, FlowTable limitation or port capacity to guide PSD, (3) PSAs that consider flow variabilities to guide PSD and (4) The PSAs that use ML optimization in their PSD. We then reviewed and compared the techniques' design in each category against the identified SDN PSA design objectives, solution approach, BSPA, and validation approaches. Finally, the paper recommends directions for further research.*

Keywords*: Software Defined Networking (SDN), Path Selection Algorithms (PSA), Routing, Quality of Service (QoS), Traffic Management, Flow table Management*

1. INTRODUCTION

The proliferation of the Internet of Things (IoT) applications has significantly increased the number and heterogeneity of traffic in modern networks[1]. A prior study reveals that active internet devices will rise from 26.66 billion in 2019 to 41 billion in 2027[3]. The signs of these become apparent during the COVID-19 pandemic [2]. The traffic arrival rate will increase the traffic volume on the internet and the rate of new path setup requests. Furthermore, the traffic flow is heterogeneous because they show non-uniform arrival rate, duration, and size. The heterogeneity affects their quality of service (QoS) requirements and demands of network resources. They behave differently en route to their destination.

Large flows like Elephant Flows (EF) are very few, about 1 -10 % of total network traffic. However, they are Long-Lived (LLF) and tend to consume network buffers. Their behaviour consequently imposes congestion and delays to most Mice Flows(MF)[4]. For instance, applications like Hadoop demand an enormous amount of throughput to perform an all-to-all transfer of petabytes during the shuffle phase. This demand is similar to a virtual machine migration that consumes high bandwidth. On the other hand, MFs are delay-sensitive and require high-priority queues. These technologies exhibit characteristic flows that require a highly dynamic network technology to meet their requirements[5].

However, techniques like Equal Cost Multiple Paths (ECMP) [6] do not distinguish among flows during routing. It amalgamates flows, irrespective of their requirements, on the same path. Using ECMP can lead to switching buffer overflows and inefficient band-

width utilization. Meanwhile, critical traffic must arrive at their destination free of any delays. In contrast, bandwidth-intensive traffics must be assured a high throughput. For these reasons, efficient resource management and optimum path selection specific to each traffic are crucial to the operation of modern Data Centre Networks (DCN)[7], Wide Area Networks (WAN)[8], Enterprise Networks [9], and Internet Exchange Point Networks (IXP)[10]. This paper studies networks management applications' ability to adapt to the needs of these technologies.

Routing involves designing network policies and configuring devices to send a flow from source to destination. However, a traditional network [11] is unsuitable for this task. The unsuitability is because the devices in traditional are distributed and do not have a global knowledge of the entire network. As such, the network operator must be on-site to adhere to vendor specifications during configuration [12–14]. Thus, network management is not flexible and precise [15]. Meanwhile, the emergence of Software Defined Networking (SDN) [16] provided a new paradigm with better flexibility to efficiently manage a network in a way that overcomes most of the limitations of traditional architecture. This paper explores how SDN architecture handles path selection for traffic routing.

The SDN's flexibility comes from separating the Control Plane (CP) from the Data Plane (DP). With the separation, all network control functions are programmed centrally as opposed to the traditional network. The centralization frees the DP to focus on forwarding packets only. The controller abstracts the DP from all network applications at Application Plane (AP) and communicates only through Northbound (NBi) and Southbound (SBi) Interfaces. The controller extracts the network statistics from the DP in real-time to feed various network applications. Examples of these applications are routing [1], security [17], congestion [18], prioritized services [19], QoS [20], load balancing [21], energy [22], and many others. The applications run their algorithms to implement new rules simultaneously throughout the network seamlessly. So, the controller can dynamically make changes to a network in response to events like the arrival of a new flow, traffic bursts, or topology changes [23]. Other reasons could be intrusion detection, a node, or link failure, which may occur every 30min in large networks [2]. In any of these situations, a route computation process to select another path is triggered to update the DP with new rules. The update operation is done proactively using protection or reactively using restoration approaches [24]. In both cases, the key challenge for the controller and the underlying Path Selection Algorithms (PSA) is to swiftly complete the DP's operations swiftly with the least convergence time and transient congestion.

The frequency of rules update is 1.5 to 5s in the restoration approach[1]. The time spent to complete one cycle of update operation consists of (1) the rule computation delay, (2) the transmission and propagation delay in distributing the rule to all the switches, and (3) the delay in installing the rule. Depending on the number of switches along the path, the process can increase the network convergence time. It is challenging for the controller to complete this operation within the carrier-grade network requirement of 50ms[24] or 10ms when dealing with critical data [25]. As reported by [26,27], the operation might be even more complex within a few milliseconds requirement of a vehicle communication system and IoT applications. Another factor to consider is the differences in controllers' processing abilities. Maestaro[28], and NOX[29] can handle 600,000 and 30,000 flow requests/s, respectively, but Ryu can only support 6000 flow requests/s [30].

Although the aggressive use of wildcards[31–36] in the protection approach may help reduce the communication overhead and yield faster forwarding performance. However, TCAM's limitation to supporting only 2000 rule entries makes this approach less attractive [3]. Thus, the restoration approach is receiving significant attention because of its adaptability to contemporary network dynamics [1]. Although at the detriment of communication overhead and the challenge of strict adherence to the conflicting QoS requirements. Given the importance of the field, various techniques have been proposed lately. However, flexibility in customization to adapt to traffic variabilities in modern-day networks is still limited. This limitation calls for exploring and reviewing more path selection criteria to provide researchers with valuable references in state-of-the-art to do more work in the field. This paper critically reviews the proposed techniques within Ten (10) years. Other studies have conducted a similar study.

For instance, Khan et al. in [5] surveyed QoS provision techniques in Service-Oriented Architecture SOA on SDN. Segment, static, and dynamic link cost routing based on Fog-Enabled IoT Platform, have been surveyed in [23,37,[38] respectively. On the contrary, this paper focused on SDN-PSA in various use cases in WAN, DCN, and WSN. Karakus and Duresi [39] partially survey routing for multimedia flows. In contrast, this paper focused on SDN PSA, classified based on path selection criteria. Tomovic et al. [40] and Guck [41] compared the QoS routing in large-scale SDN with a focus on bandwidth and delay. Guck's study focused on unicast communication and considered only the algorithms that find a single path. This paper compares general communication techniques (multicast) and multipath algorithms. Likewise, low latency transmission strategies in SDN have been surveyed in [25]. In contrast, this paper covers other QoS metrics, such as Delay, Loss, and Bandwidth during path selection. Load balancing has been covered in [15], [42–44]. The papers discussed approaches such as Controller Placement Problem (CPP) [45] and Switch Migration (SM) [46]. On the other hand, this paper focuses on PSA techniques at csCP and dmCP for load balancing, along with QoS

and practical resource usage. Waziral et al. survey issues related to topology discovery in [47]. SDN's energy issues were surveyed in [49–51]. On the contrary, this paper covers other PS goals, like Load Balancing and Resource Utilization, using various methods other than ML. This paper identifies the baseline algorithm used for PS and classifies the PS based on Selection Criteria (PSC). Then compares them for use cases, design goals, selection constraints, solution approach, and validation techniques. Table 1 provides a comparison summary to highlight these differences.

The key contributions of this research are as follows:

- The paper identified the baseline shortest path (sp) algorithms used for path selection in SDN.
- The paper Identifies different PSA problem types and design objectives in SDN
- Provide a classification of PSA based on Link Quality, Switch role, and Flow characteristics
- Finally, the study identifies and discusses potential future research directions

Figure 1 provides the paper's organizational chart. Section 2.gives an overview of Path Selection Algorithms in SDN. The overview covers how a PSA works in SDN, highlighting some baseline algorithms upon which many PSAs are built. The section also provides design objectives and different PSA problem types. It concludes by introducing the PSA classifications. Section 3. provided the critical review in four (4) sub-sections according to Link Quality, Switch role, Flow characteristics, and Machine learning approaches. Section 4 provides future research directions. Finally, section 5. concludes the paper.

Table 1: Comparison of Related Papers: **RP**: Routing Path, **SA**: Solution Approaches, **PSC**: Path Selection Criteria, **PSCr**: Path Selection Constraints, **ML**: Machine Learning, **IET**: Implementation and Evaluation Tools, **LB**: Load Balance, **RU**: Resource Utilization, **FT**: Fault tolerance

Ref.	Comm Tech		Routing Path		Use Case			Design Goal (DG)				PS Techniques Classification				IET	Missing aspect in contrast to this paper
	Uni cast	Multi cast	Single	Multi	Wired	Wire less	IoT	LB	QoS	RU	FT	SA & Problem Formulation		PSCr	PSC		
												ML	Others				
[5]	NA		X	X	SOA			X	✓	✓	X	X	✓	X	X	✓	Comm Tech, FT, ML, PSCr, PSC,
[15]	NA		✓	✓	Both		X	✓	X	X	X	X	✓	X	X	X	Comm Tech, QoS, RU, FT, ML, PSCr, PSC, IEA
[23]	✓	X	✓		L2&3VPN, DCN			X	✓	✓	✓	X	Optimization	X	✓	✓	Multi cast, Multi path, LB, ML, PSCr
[25]	X	X	Not Stated		Not Identified			✓	✓	✓	X	✓	✓	X	X	✓	Comm Tech, Use Case, PSCr, PSC,
[37]	✓	X	Single		Wired			X	✓	X	X	X	Optimization	✓	✓	✓	Comm Tech, Multi-Path, IoT, LB, RU, FT,
[38]	✓	X	Both		Fog, IoT			✓	✓	X	X	X	✓	X	X	✓	Multi cast, RU, FT
[39]	NA		Not Stated		WAN			X	✓	X	X	X	Optimization	X	X	X	Comm Tech, PSCr, PSC Path, IEA
[40]	X	X	Not Stated		WAN			X	✓	X	X	X	✓	✓	✓	X	Comm Tech, Path, IEA
[41]	✓	X	✓	X	✓	X	X	X	✓	X	X	X	✓	✓	✓	X	IEA, LB, RU
[42]	X	X	X	X	Both		X	✓	X	X	X	X	✓	X	X	X	Comm Tech, QoS, RU, ML, PSCr, PSC,
[43]	X	X	X	X	Both		X	✓	X	X	X	X	✓	X	X	X	Comm Tech, QoS, RU, ML, PSCr, PSC, IEA
[44]	Not Specified		X	X	cSC, dDC, & 5G			✓	✓	X	X	✓	✓	X	X	X	Comm Tech, RP, RU, FT, PSCr, PSC, IEA
[48]	X	X	X	X	✓	✓	X	X	✓	✓	X	✓	X	X	X	✓	Comm Tech, Path, Env, SA,
[49]	NA		X	X	Energy, DCN			X	✓	✓	X	X	Optimization	X	X	✓	Comm Tech, RP, LB, QoS, ML, PSCr, PSC
[50]	NA		Not Stated		Energy			✓	X	✓	X	X	Optimization	X	✓	X	Comm Tech, RP, QoS, FT, ML, PSCr&IET
[51]	NA		Sleep Schedule		Wired		X	X	✓	✓	X	X	Optimization	X	X	X	Comm Tech, LB, FT, PSCr, PSC, IET
This Paper	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

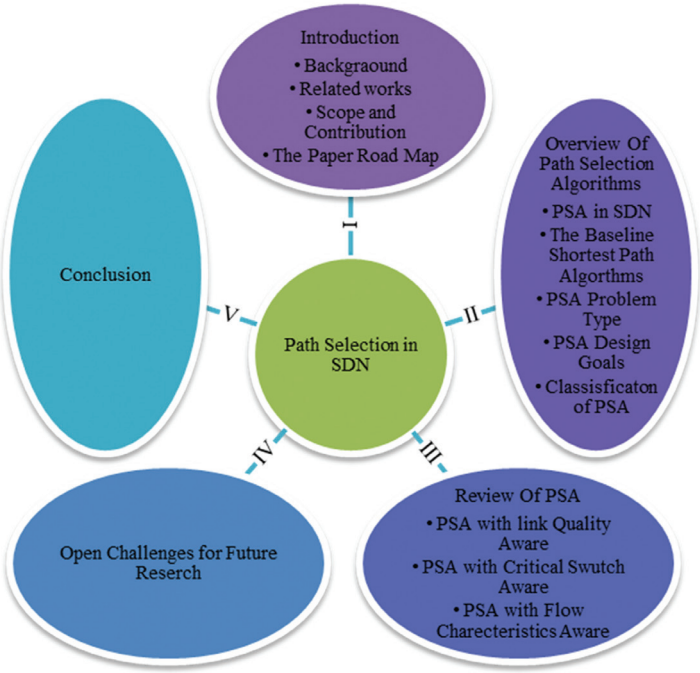


Fig. 1. Organization of the Survey

2. OVERVIEW OF SDN PATH SELECTION ALGORITHMS

2.1. PATH SELECTION ALGORITHM (PSA) IN SDN

SDN controller used an OpenFlow [52] to communicate with the DP. The standard uses a Link Layer Discovery Protocol (LLDP) to discover the DP topology. Upon discovery, the controller uses a PSA to compute a rule that guides flow to their destinations. It then instructs the switches to install the rules in their respective flow tables. In most cases, a PSA is invoked whenever a new flow with no corresponding entry in the flow table arrives. Another reason could arise from topology changes due to link or node failure. In both situations, the PSA must converge the network with the new rule to avoid disruptions. The controller uses a thread monitoring mechanism to track these changes by collecting statistics in a fixed cycle processing. The mechanism periodically issues a request to switches to get the information. In the Ryu framework [53], the issuance of information requests for all the registered switches is repeated infinitely after a set time interval. The network rules must be regularly updated every 1.5 to 5s due to traffic variabilities [1].

2.2. BASELINE SHORTEST PATH (SP) ALGORITHMS USED FOR PS IN SDN

Most of the PSA for SDN often use mechanisms of the Shortest Path problem as a foundation. Some of these are based on algorithms like Dijkstra[54], A* [55], or Bellman-Ford[56]. Others are based on k Shortest Paths (kSP), like Yen's algorithm [57]. In contrast, others use Restricted Shortest Path (kRSP) or Constrained Shortest Path (kCSP) like A* Prune [58]. These algorithms are proven to be NP-Complete [59]. So some exact algorithms for handling the problem in a brute-force manner or approximation like Lagrangian Relaxation Based Aggregated Cost (LARAC) are used.

Dijkstra [54] is a centralized algorithm that calculates the Shortest Path (SP) to multiple destinations in a graph with non-negative link weight. It keeps a queue with a list of partial paths starting at the source and going through intermediate nodes before the destination. The least weighted path from the queue is chosen at each iteration. So n paths are created by extending the partial paths to n external links of the node where the path ends. Only routes with a lower weight than the queue's current route that led to the same destination are enqueued. Depending on the weight values, the Dijkstra loosens up by removing a path with a higher weight value. The algorithm uses a Breadth-First Search (BFS) to visit a node.

On the other hand, the **A* algorithm** [55] is an enhancement of [54] by introducing a guess function at each node to estimate the path's cost from the reference node to the terminal node. The priority queue's outlier paths with minimum projected cost are expanded initially. The nearer the estimated cost is to the actual cost, the faster the algorithm converges. However, the overhead brought forth by the computation of the guess function constitutes the trade-off to consider. Furthermore, in contrast to [54], a **Bellman-Ford** (BFA) [56] computes the Shortest Path (SP) tree in a network with negative edge weight. Unlike Dijkstra, BFA operates in a distributed manner. The algorithm keeps track of the best path connected to V nodes. It performs |V|-1 number of iterations to update each node's current best path. This way, all SPs will ultimately be found because the path to any node is only|V|-1. It instantly

halted if an iteration yielded no update because successive iterations would not result in any change.

The earliest algorithm to address a k-shortest path problem is **Yen's** [57]. The algorithm runs in two (2) phases. In the first phase, it works based on traditional SPA to find the initial path. Then, the other (k-1) shortest paths are found regarding the initial path's intermediate switches. The algorithm is similar to [55] and always stops once the terminal switch is visited k times. Additionally, **the A* prune** [58] identifies k-SP from source to destination in a communication network. The technique is an exact algorithm similar to [55] in using a guess function for each metric. Paths are considered by adjusting the cost estimation and pruning according to the cost projection exceeding the corresponding end-to-end bound. The process continues to iterate until k-CSPs are found or the candidate path container is empty.

To deal with NP-Complete problems such as kRSP or kCSP, a Lagrangian Relaxation Based Aggregated Cost (LARAC) [62] is used to relax the CSP with a combined delay and link cost to SP problem with a modified cost function. The method allows dropping some constraints of the first problem and introducing them in the optimization goal. However, the technique has a duality gap deficiency because it does not guarantee the best path return.

2.3. THE DESIGN GOAL/OBJECTIVES OF PSA

In many cases, PSA algorithms' design goal is threefold: QoS satisfaction, Resource Utilization, and Load Balancing.

QoS satisfaction deals with the ability of a network to consistently adhere to the performance expected by an application in terms of delay, jitter, bandwidth, throughput, and loss. QoS routing is one of the most critical components of a network management framework [41]. Ensuring it in a path is challenging due to network dynamics. Users' conflicting interests compound the difficulty, as exemplified by the high demands of pervasive applications like VoIP, video conferencing, telemedicine, and online game. Many routing strategies are designed to adapt to these demands and select a route with optimized QoS requirements [5]. Other PSAs are designed with resource utilization objectives. These PSA allocate network resources such as CPU, Memory, and bandwidth to traffic based on availability, priority, and requirement [61–63].

Lastly, other PSAs are designed to distribute network loads for task processing, packet transmission, and storage to network components based on their residual capacity [42]. The default settings of control algorithms are often similar across all CP[25]. In most cases, they are based on Shortest Path Algorithms (SPA) such as [54],56]. For example, the default setting of the Beacon [64] is [54]. The situation conditioned the CP to take the same PSD, irrespective of network conditions. However, some paths are more appealing than others, thus becoming critical as all the switches often select a node connected to them as the next hop. Thus, many flows are sent to the same path simultaneously. However, suppose the network experiences a traffic burst caused by some hot events. In that case, the component already running at full will be under added strain. The unbalanced distribution of the traffic along paths can lead to some paths becoming congested. Consequently, the network begins to experience delay and packet loss, resulting in an ultimate failure if the situation persists. Some PSAs are designed with load-balancing objectives to mitigate this situation at th DP. Another thing to note is that load imbalance affects the CP in response time delay as it affects the DP[25]. A dmCP can be overwhelmed due to an imbalanced distribution of flow request processing tasks. At the CP, the problem is being addressed through Controller Placement Problem (CPP) solutions [65–67] and Switch Migration (SM) [68–72]. Therefore, the load balance problem is addressed through the PSA at the DP and CPP or SM at CP. This paper focused on the PSA for load balancing at the DP.

2.4. PSA PROBLEM TYPE

A PSA is designed to best-effort traffic or for traffic with stringent QoS requirements. Similarly, PSA can be designed with fixed (non-adaptive) or dynamic (adaptive) link costs for the PSD. The diagram in Figure 2 shows the distribution of PSA problem types.

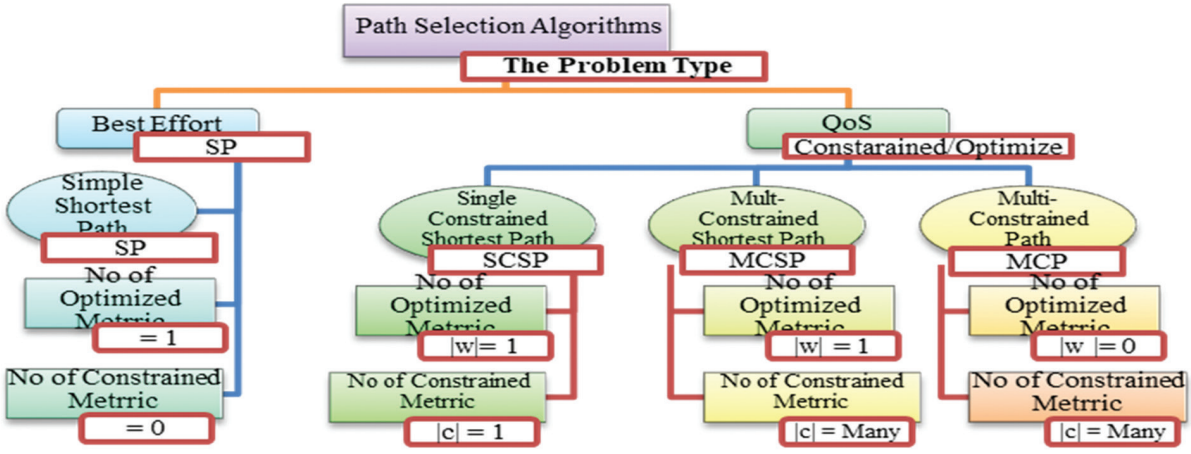


Fig. 2. PSA Problem Type

Best effort PSA does not aspire to guarantee that traffic will be delivered to its destination or meet any QoS requirements. The shortest path based on a minimum number of hop count with no other constrain is often used as the PS metrics. In contrast, PSA with QoS makes Path Selection Decisions (PSD) with predefined objectives. It aims to optimize some performance metric(s) while keeping others below a prescribed threshold. The metric to optimize is referred to as cost or weight. While the ones to be kept below a certain threshold are called constraints. Depending on requirements, the PSA can be designed to address different problems. For example, a PSA problem can be defined as a Single-Constrained Shortest Path (SCSP), a Multi-Constrained Shortest Path (MCSP), or a Multi-Constrained Path (MCP) Algorithm.

Consider a PSA for a network modelled as a graph $G=(S, E)$, with S, and E as a set of switches and communication links. The total number of switches and links are $|S|$ and $|E|$. If $w \in R_+^{|E|}$, is a vector that denotes the weight of the links between two adjacent switches $s \in S$. Let $c \in R_+^{M}$, denotes another vector to holds M elements corresponding to the threshold of the constrains metrics. Also, let $M \in R_+^{Mx|E|}$, be a matrix that represents the values of the constraints for individual link $e_{ij} \in E$, between source i and destination j. Also, let P_{ij}, be the complete path between the source s_i and destination s_j, in the set of all paths P in G. With these factors, SCSP, MCSP, or MCP can be formulated mathematically to optimize any QoS metric (Min or Max) represented by the weight vector w as:

$$SelectP = min\, w^T X, \qquad \forall\, p_{ij} \in P \tag{1}$$

$$\mathcal{M}X \leq c \tag{2}$$

$$X = \begin{cases} 1, & if \quad e_{ij} \in p_{ij} \\ 0, & if \quad e_{ij} \notin p_{ij} \end{cases} \tag{3}$$

SCSP finds a route with minimum end-to-end QoS metric while keeping another metric below a prescribed bound. It corresponds to a situation where c in equation (2) = 0. In contrast, an MCSP is defined to optimize many end-to-end QoS metrics constrained by individual bounds. It corresponds to a situation where c in equation (2) >1. Lastly, an MCP is a selection problem defined without an optimization metric. The route to be selected while keeping some QoS metrics below a prescribed threshold. It corresponds to a situation where c in equation (2) > 1.

2.5. CLASSIFICATION OF PSA SOLUTIONS IN SDN

As shown in Figure 3, Path Selection Algorithms PSA considers link quality, switch role (critical node) flow characteristics, or a combination of these during path selection decision-making. This paper classifies the existing PSA based on these criteria.

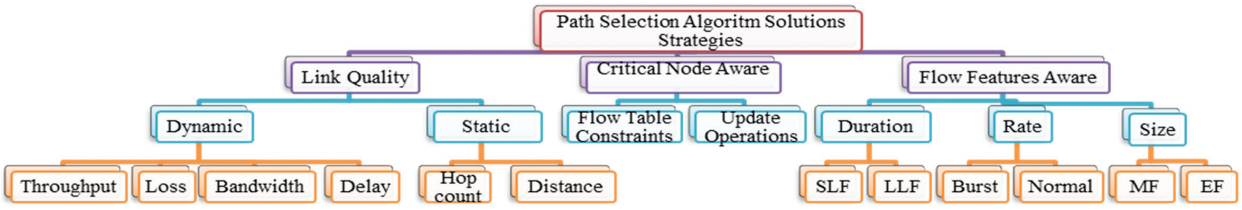


Fig. 3. Classification oF PSA in SDN

2.5.1. Link Quality

Link quality is estimated in teams of metrics such as delay, bandwidth, loss, throughput, or jitter. Depending on the problem, these metrics can be modelled as an objective function of an optimization problem while keeping other metrics as constraint(s). These metrics are either static or dynamic. Dynamic PSA uses metrics such as bandwidth. It is dynamic because the metric's value may change each time traffic is routed. In contrast, PSA, with a metric like a hop count, is static because it remains fixed unless there is a total change in topology [58]. The metrics are also classified as additive, multiplicative or non-additive. Additive metrics are delay, hops, and jitter. They imply that a path's end-to-end metric can be found accurately by summating the weight of different links. The weight of a path with non-additive metrics, like bandwidth, can only be established by the value of that constraint at the blockage link.

2.5.2. Critical Node(Switch)

A node/switch in a communication network is critical if many switches select it as the next hop because it falls along the paths of other switches. Consequently, because of its position, the communication frequency of such a switch with the CP for rule installation is higher than regular switches due to the number of SP passing through it. Such as, a switch tends to generate high communication overheads. However, the regular switches are responsive and dependent on the efficiency of the critical switch. Tools in graph theory such as degree [74], betweenness [75], information [76], closeness centralities, and PageRank [77] can be used to measure switch importance in a network.

Consider a network $G=(S, E)$, where S is a set of switches and E communication links. The network identifies a critical switch in terms of parameters such as switch update operations, switch flow table residual capacity or switch port's residual capacity[73]. The number

of rules operations finds the switch update operation on each switch. The value reflects the total traffic load between different sources to destination switch pairs. For instance, a switch $s_\Omega \in S$ serving many intermediary switches through, $|P_{sp}|$ number of shortest paths between s_i to s_j switch pairs is considered more critical than regular switch $s_\omega \in S$. Switch s_Ω will be heavily loaded with a high number of rules in its flow table in comparison to s_ω. See Table (2) for a comparison of different measurement metrics.

Table 2. Comparison Table of Node Importance Evaluation Metrics

	Metric		Description
Node and edge Centrality	Eigenvectors **EV**		**EV** measures the importance of a node in a network. It is based on the principle that connections to nodes with a high degree contribute more to its score than connections to nodes having a low score.
	PageRank **PR**		The score computed by **PR** is higher for nodes that are highly connected with nodes that are highly connected themselves. **PR** score is iterated until convergence. **PR** is a variant of the EV centrality measure.
	Hyperlink-Induced Topic Search: **HITS**		• **HITS** calculates two scores: **Hub** and **Authority** • The more a node has outgoing links, the higher the Hub score. While the more a node has incoming links, the higher its **Authority** score. • Initially, every node is considered a **Hub** • **Authority** scores are fixed to a constant. The scores are updated and converge after a few iterations.
	Degree Centrality **DC**		A **DC**-based measure of individual centrality corresponds to how well-connected the individual is within their local environment.
	Closeness Centrality **CC**		**CC** measures centrality on a global scale based on how close a node is to all the other nodes. The idea is that a switch is central if it can interact with all others quickly
	Betweenness Centrality **BC**	Geodesic	**BC** of a node measures the extent of the role of the intermediary node in interacting with others.
		Path/flow	
		Random path	
	Information Centrality **IC**		**IC** metric is based on the concept of efficient propagation of information over the network. IC of a node is the relative drop in the network efficiency caused by the removal of the node from the network. It combines the idea of CC and BC measures.

2.5.3. Flow Characteristics

A flow has been defined differently based on the research context [3]. Most definitions convey that a flow is a set of packets sharing common identification properties while passing an observation point during a specific period [78]. These properties, in most cases, include five tuples like protocols such as (TCP, UDP, ICMP), source-destination IP, and port numbers.

Network traffic is made up of different types of flows [52]. Examples are EF and MF. EF are typically very few, about 1 -10 % of total network traffic. However, they are Long-Lived flows(LLF). As such, they tend to rapidly consume network buffers which consequently impose congestion and queuing delays to the minor majority of MF [4]. Each of these traffic types has specific QoS demand and behaves differently en route to destinations. Therefore, considering flow features during PSD is necessary because of their potential impact on network performance.

Flow statistics such as packet counts, bytes count, duration, per-flow packet size distribution (PSD), rate, or burst [81,82] are used to classify flows. These metrics can be used in isolation or combination to classify the flows using a threshold or threshold-less approach.

- Flow Size fS: is the number of bytes transmitted in a flow. It can be quantified using byte or packet count. With fS, flows are classified using a threshold Th such that if the $fS>Th$, then the flow is an EF and an MF otherwise.
- Flow Duration fD: This is the elapsed time between the first and the last packet of a flow. In a threshold-based, any flow with $fD<Th$ are tortoises, while flows with $fD \geq Th$ are dragonflies.
- Flow Arrival Rate fR is found by dividing the size fS of the transmitted data by the total flow duration fD.$fR=fS/fD$. If $fR>Th$, it is a cheetah and snail otherwise.
- Burst: Traffic burst investigates the extent of traffic and connection dominance in a network. Burst traffic involves packets with a short inter-arrival period. If packets inter-arrival time $iAT>Th$ are called porcupines and stingrays otherwise [78].

However, there is no unanimous settlement among the proposed flow classification approaches on what flow feature to adopt. However, the majority of the approaches used flow size [79], [84–87]. To set up the classification threshold, many other techniques used duration [88], rate [89–91], or burst as well. So, many PSA in SDN designed their solutions while considering these factors. These algorithms are reviewed in section 3.3 and summarised in Table 6

3. REVIEW OF PATH SELECTION ALGORITHMS IN SDN

3.1. PSA WITH LINK QUALITY AWARE

To choose a path for traffic in an SDN-Edge computing network, Hu et al.[92] considered loss and latency to design a Path Selection Method (PSM). The edge node at the SDN boundary is configured to assign network resources, such as bandwidth, according to flow requirements. Flows are directed via a path with lower packet loss. At the same time, delay-sensitive flows are routed via a path with minimal delay. Likewise, Alnajim and Salehi [93] proposed an incremental scheduling QoS-aware path selection technique to quickly redirect real-time applications with time-bound flows. The technique avoids bottleneck links causing scheduling impasses by selecting a path with enough residual bandwidth from the list of candidates' paths. Before the optimum PS, the technique incorporates an offline pre-routing where initial K paths are formed using Yen's [57]. The run time of this stage is high. The authors controlled it by coupling a Fibonacci Heap with Dijkstra [55] and [57]. However, in [92], [93], the switch load was not considered when choosing the path. Hence, QoS parameters like latency and the PDR should also be considered to ensure the quality of the link choice. In a similar technique in [31], switch port capacity in terms of data Transmission (Tx) and Receive Rate (Rx) is monitored. The data is pulled to determine the maximum capacity or load the port can accommodate. The statistics are fed to an application on a Floodlight [96]. The controller defines rules that send a flow via a port with Least Loaded Path LLP. In the validation, the authors used Mininet and Iperf to generate traffic. However, the constant gathering of statistics may significantly increase the controller overhead, thus increasing flow setup latency.

On the other hand, a QoS-driven and SDN-assisted Multipath Selection Scheme (QSMPS) was proposed in [94] to address the adaptability problem of traditional MPTCP. The method checks and assesses the network status using a scalable SDN- technique to gather statistics. Based on the data, an optimal number of sub-flows are found by QSMPS, which are distributed along the routes with short delays. The authors validate QSMPS in topology on Mininet using a Ryu framework. However, a best-case scenario based on residual bandwidth might not always ensure that distinct flows' requirements are met. Besides, the proposed scheme did not classify traffic according to its uniqueness. Because other flows can choose a short setup latency and a link quality that decreases the number of times links change due to topology changes. To compute the best paths while taking the QoS criteria for each flow into account, Saha et al.[95] introduced a greedy heuristic based on Yen's [57]. Multiple metrics are jointly considered in formulating the problem to get the best paths. The multi-constraint QoS-aware route is solved using Integer Linear Programming (ILP). The authors validate it on POX [96] using a Mininet with D-ITG [97] to generate IoT- traffic. However, in large networks with frequent topology changes, the ILP-based method may slow the convergence of the routing rules [88]. Perner and Carle [98] study the effects of various optimization on network link utilization and latency. A path selection technique was designed and bounded with some constraints to meet its requirements. One of the constraints is TCAM's size limitation. The constraint is modelled such that the number of outgoing flows does not exceed the maximum number of forwarding table entries. However, the switch update operation is not considered. Similarly, just like [87], the technique may suffer from high routing convergence time due to the ILP model.

Khalili et al. [99] designed a mechanism that determines a controller's flow setup latency by crafting and sending Special Ping Packets (SPP) between source-destination pairs. The Round-Trip Time RTT of the packets is measured. The time is attached to the individual path to get the total path setup latency. The technique is designed for scCP. Accordingly, Ravuri et al. [100] consider using dmCP to improve [99] overhead, scalability, and Single Point Failure SPOF experience in scCP. The paper emulates the DP on Mininet and implements the dmCP with floodlight[101], ONOS [102], and Kandoo [103]. However, both schemes may experience a high flow table operation, affecting path setup and switching time.

Intersection-based routing using SDN and Fog computing is proposed in [104] to address communication coverage holes in VANET. The controller's global network knowledge is leveraged to collect street score information from fog nodes and feed it into Dijkstra [54] to build the routing path. In a similar approach, [105] uses SDN to propose a dynamic routing to cope with the problem of frequent changes in a Flying Ad Hoc Network (FANET) of a drone (UAV) network. The technique Hybridizes a traditional OLSR with an SDN controller to perform topology discovery, statistics gathering, and route computation. OLSR and SDN alternate network control depending on the network status. However, [106] note that the conventional routing commonly used in DCN, like OSPF, incurs significant overhead with high convergence time. For this reason, controller-side Regular Topology Routing (cRetor) is proposed. cRetor differs from other topology–aware routing in compatibility with various other topologies. Including topology description language in the scheme ends the need for LLDP to run first. The action frees and relieves the scarce bandwidth and processing loads on the controller, respectively. The authors claim that the route calculation time of the technique is fast, which makes the convergence time shorter. The overheads and the failover perfor-

mance are admirable. Alidadi et al. [109] proposed a low-complexity SDN-MPLS algorithm encompassing a quid pro quo between load balancing, hop count, and power consumption in mobile –SDN with restricted bandwidth during PS.

Table 3. Comparison of PSA with Link Quality Aware

Ref	Path Selection Criteria						Use Case					Design Objective			Solution Approach & BLA		Validation & Implementation Tools
	Parameters			No of Constraints			Custom/ Others	DCN	WAN	IoT	LB	QoS	RU	AI	MM / Others	
	Link QoS	Critical Switch	Flow Features	MCP	SCSP	MC SP										
[92]	✓	X	X	X	X	✓				✓		✓	✓	X	✓	Simulation
[93]	✓	X	X	X	✓	✓	Random	X	X	X	X	✓	X	X	Fibonacci, Yen,Dijkstra,	
[31]	✓	X	X	X	✓	X	Fat Tree	X	X	X	X	✓	X	X	Dijkstra[54]	Floodlight Mininet, Iperf
[94]	✓	X	X	X	✓	X	✓	X	X	X	X	✓	X	X	✓	Ryu,[53] NetworkX
[95]	✓	X	✓	✓	X	X	X	X	X	✓	X	✓	X	X	ILP, Yen	POX, Mininet [107]
[98]	✓	✓	X	X	X	✓	Critical	X	X	X	X	✓	✓	X	ILP	ITZ
[99]	✓	X	X	X	✓	X	csCP	X	X	X	X	✓	X	X	Polling method	Floodlight, Mininet [107]
[100]	✓	X	X	X	✓	X	dmCP	X	X	X	X	✓	X	X		Floodlight, ONOS, Kando
[104]	✓	X	X	X	✓	X	VANET	X	X	✓	X	✓	X	X	Dijkstra	SUMO,NS2[108]
[105]	✓	X	X	X	X	✓	FANET	X	X	✓	X	✓	X	X	Hybridized	
[106]	✓	X	X	X	X	✓	DCell	✓	X	X	X	✓	X	X	A*[58]	Floodlight, Mininet
[109]	✓	X	X	X	X	✓	Mobile	X	✓	X	✓	✓	✓	X	------	MIRA

3.1.1 Multipath PSA With Link Quality Aware

Megyesi et al. [110] proposed a mechanism for measuring Available Bandwidth (ABW) during path selection. The authors substitute the distance metrics of [54] with the link ABW to decide which path to select. The method is modelled as a max-flow problem to adapt to different use cases requiring the choice of the best available path among multipath. A Ford-Fulkerson algorithm is employed for the max-flow problem. Another technique is proposed by Dutra et al. [111]. Where, for each of the ingress traffic in DCN, the technique provides it with the required end-to-end bandwidth and effective use of the switches along the selected path. The action leads to a reduction in path use cost and execution time.

Similarly, in the work of Celenlioglu and Mantar [112], a pre-established multi-paths (PMP) between each source-destination switch is used to visualize an underlying DP topology to design a PSA with resource management through Admission Control (AC), Load Balancing(LB), and Path Resizing(PR). In another approach, [113] and [114] proposed GridFTP for the parallel transfer of a large amount of scientific data along multiple paths based on the Dijkstra[54]. Equally, using a multipath approach, Tariq and Bassiouni [115] extended their initial design of a QoS Aware Multipath (QAMO) [116] for a traditional optical network to SDN. The extension supports adaptive QoS differentiation with a priority factor for burst traffic and link state. It uses a Dijkstra [54] to find K paths between the source-destination pair.

Furthermore, [117] proposed a multipath forwarding approach for next-generation networks using SDN. The technique addresses the need for a conducive inter-networking environment for future data-centric applications. The technique exploits the edge diversity of transit ISPs available across many IXPs to propose cross-layer coordination with SDN flexibility. The authors use [118] to discover the initial K-shortest paths. Before invoking a route reconciliation and update strategy to re-evaluate the initial choice at the Control eXchange Authority (CXA) Controller implemented with Ryu[53] SDN framework. Likewise, in a similar effort, A Dynamic Multipath Scheduling Protocol (DMSP) for identifying and isolating congestion-susceptible links in DCN using SDN is demonstrated in [119]. DMSP split the flow traffic among multipath to reduce the congestion. However, the splitting is done identically among the available paths. To address the problem of the unequal split of DMSP [119]. Farrugia et al. [122] proposed a Globally Optimised Multipath Routing (GOMR) algorithm that splits a flow traffic equally among multiple paths using a stochastic mechanism. GOMR leverages the global knowledge of network topology and its traffic statistics to formulate the problem as linear programming LP to optimize per packet multipath routing proposed.

Table 4. Comparison Table for PSA with Multipath

Ref	Path Selection Criteria						Use Case				Design Objective			Solution Approach & BLA		Validation & Implementation Tools
	Parameters			No of Constraints			Custom / Others	DCN	WAN	IoT	LB	QoS	RU	AI	MM / Other	
	Link QoS	Critical Switch	Flow Features	MCP	SCSP	MCSP										
[110]	✓	X	X	X	✓	X	X	✓	X	X	X	✓	X	X	Dijkstra, Ford Fulkerson	Floodlight, Mininet, DITG
[111]	✓	X	X	✓	X	X	X	X	X	X	X	X	✓	X	X	
[112]	✓	X	X	✓	X	X	Intra Domain	X	X	X	✓	✓	✓	X	✓ MM	Floodlight, Mininet Ditg
[113] [114]	✓	X	X	X	✓	X	Virtual & Real Env	X	✓	X	X	✓	✓	X	Dijkstra	OvS
[115]	✓	X	X	X	✓	X	DCN	✓	X	✓	X	✓	X	X	Dijkstra	C++
[117]	✓	X	X	X	X	✓	Next-Gen ISP, IXPs	X	✓	X	X	✓	X	X	EPPStein KSP[118]	Ryu, Mininet
[119]	✓	X	X	X	X	✓	Fat Tree	✓	X	X	X	✓	✓	X	EPPStein KSP[118]	Ryu, Mininet
[120]	✓	X	X	✓	X	X	Butterfly, G´EANT	✓	✓	X	✓	✓	X	X	LP	Ns-3, GLPK, LEMON

3.2. PSA WITH CRITICAL SWITCH AWARE

Yan et al. [24] employed a TCAM-aware flow rerouting approach to address the fault tolerance problem. The authors formulate the problem as optimization with an objective function that finds a set of backup paths with TCAM and bandwidth as constraints. Solutions are sought using Forward Local Rerouting (FLR) and Backward Local Rerouting (BLR) heuristics. The two heuristics are used alternatively depending on the network state. However, the coexistence of the multiple heuristics at the controller may introduce an extra computational complexity. SwitchReduce[37] presents another intermediate switch state and controller participation technique. The technique is founded on wildcard identical action flows, RouteHeaders (first hop-based routing), and division of labour principles. SwitchReduce ensures the number of rules in the flow table is decided by the corresponding actions taken on flows going through it and does not increase linearly according to the flows. However, the technique is not adaptive to topology changes and has not been validated on large DCN.

In a similar effort, Perner and Carle [98] study the effects of various optimization objective functions on network performance metrics, such as link utilization and latency. Path selection technique bounded with some constraints is designed to meet the network requirements. One of the constraints considered is the TCAM size limitation. The constraint is modelled such that the number of outgoing flows does not exceed the maximum number of forwarding table entries. Astaneh et al. [121,122] proposed another work to restore SDN failures by rerouting disrupted flow with several switch update operations. The technique finds the switches with a small number of flow table entries to reroute the flows during the path restoration process. It is designed as a local restoration plan problem formulated as an ILP to trade-off between path cost and the switch update operation. Dijkstra [54] is leveraged to select a path with a minimum hop count. However, the technique may suffer from high routing convergence time due to the ILP. In another technique, Malik et al. [123], [124] proposed an alternative way to reduce switch update operation and preserve the limited size of the flow table during rerouting at the time of failure. They proposed an optimum path selection while looking at shared links across paths. The shared links have a higher tendency to bring down the consumption of flow table spaces. The technique picks a route with the most shared links from the paths list. it is validated with POX[96] on the Mininet. However, when the switch utilization rate rises, there is a greater likelihood of load imbalance and possible overflow.

Incidentally, Yu et al. [125] presented a path selection method based on node significance and flow prediction. The authors use Deep Neural Networks (DNNs) Q-Leaning [126] to balance the network load. The scheme comprises three algorithms running as the intelligent centre on the controller. The identification of the critical nodes is made using H-index. Flow forecast is achieved with DNN, and path selection is based on Q-leaning according to node importance. However, despite the benefits, the TCAM space constraint might limit the potential use of DNN[127] and Q-Leaning [126] that so much relies on historical data. In another work, Gotani et al. [128], [129] proposed a technique to reduce the effect of switch processing latency during path setup. Since the time to add flow entries is different for different switches. The paper designed a scheme of three methods to select an optimized path while minimizing the total switching time. The authors made path selection decisions based on path-switching delay in multiple paths with varying switch processing times. In this manner, path screening took place, and the one with the set of switches that requires the least processing time was selected. In a large-scale network, the solution might not deliver the best performance. Also, in the work of Isyaku et al. [1], [73] link quality and switch update operation is considered for path selection. However, the authors ignore the heterogeneity of the traffic traversing the network at the point of the path selection decision.

To minimize the impact of high demands of flow rule updates in space-hungry TCAM, the work in [130] considers the diversity of instruction types and switch behaviour to propose the ***RuleTailor*** algorithm. RuleTailor is an efficient, measurement-based optimization framework for SDN flow routing rules updates. Different from the consideration of switch behaviour in RuleTailor, the techniques proposed in [131,132] adopt the concept of aggregation of the routing rules. The approach aims to trim the number of possible rules in the table. [131] classify paths into two, one for popular flow and the other for non-popular flow. On the other hand, Jia and Wang [132] translate the Destination Address and Source-Port on Demand (DATSPToD) so that aggregate routing rules are taken to minimize entries in the flow table. The technique is meant for Large-scale SDN with scattered address space allocation. Similarly, DATSPToD is founded on address modification and port rewriting to address the problem of inefficient routing due to interleaved allocation of a non-contiguous IP address. Similarly, using Mixed Integer Non-Linear Programming (MINLP), Guo et al. [133] introduce path cardinality constraints on a PSA to prune the number of rules in TCAM. Dijkstra [54] is applied to find initial paths. Then passes them through a route optimization function before invoking an H-Permissible to prune and select a path based on the cardinality constraints. However, the inclusion of the optimization function adds up to the complexity by order of $(LP(N^2+L, L))/2$ [77]. Maaloul et al. [134] proposed a technique for SDN-based CGN to optimize energy. The controller dynamically turns on or off a network component while considering their residual space. The authors formulate the problem as Binary Integer Linear Programming (BILP) with TCAM and link usage as constraints. The aim is to minimize the power consumption of links and switches. A First-Fit Heuristics (FFH) is designed to solve the problem based on traffic demands. Dijkstra [54] is used as the baseline algorithm to identify, sort, and select paths according to First-Fit Most-Power (FFMP), First-Fit Least-Power (FFLP), and First-Fit Random (FFR).

Table 5. Comparison Table of PSA with Critical Switch Aware

Ref	Path Selection Criteria						Use Case				Design Objective			Solution Approach & BLA		Validation & Implementation Tools
	Parameters			No of Constraints			Custom	DCN	WAN	IoT	LB	QoS	RU	AI	MM	
	Link QoS	Critical Switch	Flow Features	MCP	SCSP	MCSP										
[24]	X	TCAM Size	X	✓	X	X	Failure									GLPK, Internet2
[36]	X	✓	X	X	X	X	ITZ	✓	X	X	X	X	✓	X		NoX OVS, Mininet
[98]	✓	TCAM Size	X	X	X	✓	Critical system	X	X	X	X	✓	✓	X	✓ ILP	ITZ
[121], [122]	X	Update Operation	X	X	X	✓	ITZ	X	X	X	X	X	✓	X	ILP, Dijkstra	ERnet USnet
[123], [124]	✓	Size & Update	X	X	✓	X	✓	X	X	X	X	✓	X	X	✓	POX Mininet NetworkX
[125]	X	Node Role	✓	X	X	✓	Random	X	X	X	✓	✓	X	DNN Q-L	X	---
[128], [129]	X	Switch CPU	X	✓	X	X	Disaster	X	X	X	X	✓	X	X	✓	NA
[1], [73]	✓	✓	X	X	X	✓	Custom	X	X	X	X	✓	✓	X	✓	Ryu, Mininet
[130]	X	Switch Role	✓	X	✓	X		X	X	X	✓	X	✓	X	✓	Ryu OVS, Iperf
[131]	X	TCAM compression	X	X	X	X	Flow rule update	X	X	X	X	X	✓	X	Aggregation Markov	Floodlight OVC MATLAB
[132]	X	TCAM	X	X	X	X		X	✓	X	X	X	✓	X	✓	NA
[134]	✓	TCAM power cost	X	X	✓	X	Energy CGN	X	X	X	✓	X	✓	X	BILP, Heuristic Dijkstra	SNDlib, CPLEX MATLAB
[133]	✓	TCAM	X	X	✓	X	ITZ	X	X	X	X	X	✓	X	MINLP Dijkstra,ARA	C++

3.3 PSA WITH FLOW CHARACTERISTICS AWARE

Several PSAs have been proposed to support the streaming of video flows [135–139]. Civanlar et al. [135] formulate the problem as Linear Programming (LP) to minimise weighted route length and packet loss. The technique finds the best path to accommodate video and the shortest path for the best effort. Harold et al. [136] designed the Routing Module (RM) of their proposed Video scheme based on the A* Prune [58]. The technique returns a list of paths that satisfy bandwidth, jitter, and delay constraints. It includes three modules for policy, admission control, and path reservation, with adequate resources for traffic control. However, the policies cannot guarantee the reservation of paths for all requests made. Thus, delay and packet loss might be experienced. The authors attribute the limitation to the scalability associated with a single controller. In a similar effort, [137] proposed an adaptive technique to reroute video with QoS support using LARAC [140]. The authors implement the algorithm on a Floodlight [101]. In a dif-

ferent approach, [138] improves the QoS provisioning of video services from the server side. The authors proposed a framework for load-balancing over a single-operator network to improve the QoS of video streaming. The framework monitors the load of the servers to track packet loss and delay variation. It redirects streaming requests using LARAC to a video server with a lighter load. In another work, a technique of PS for multi-media flows is proposed by Chooprateep et al.[139]. The authors designed a Video (VPSA) that finds a path based on Yen's algorithm using the controller's data of the link's past and present bandwidth utilization. Depending on whether a suitable path exists, video flows are either refused or allowed upon arrival. However, the coefficient used in the problem model is not adaptive to the characteristics of the traffics. Moreover, the algorithm considers only a single dynamic QoS metric while ignoring static metrics such as link latency and video holding.

In Egilmez et al. [20], an OpenFlow controller application, OpenQoS, implemented on a Floodlight [101], is proposed to handle multimedia traffic separately with end-to-end QoS assurance. OpenQoS is designed based on LARAC [140] approach. The algorithm uses the packet's header field in the MPLS structure of every incoming traffic to classify and separate multimedia traffic from data traffic. Network statistics are collected via feature_request messages every 1ses to enable the calculation of the QoS of each available route. With the separation and QoS knowledge of all routes, each traffic class is handled differently. The multimedia flows are placed on the route with the required QoS resources. At the same time, the data traffic is handled with the best effort forwarding. Saha [95] took the QoS metric specific to each flow into account to propose a greedy heuristic based on Yen's k-SP algorithm. The algorithm selects the ideal path for each flow. Multiple metrics are considered in formulating the problem as Integer Linear Programming (ILP). The model is validated using POX [96] controller and Mininet with D-ITG[97] to generate an IoT-based use case traffics. However, in large networks with frequent topology changes, using the ILP-based method may result in slow network convergence [88]. Kotani and Okabe [141], employed a packet filtering technique to separate the most critical flows from others at the level of packet-in messages. The filtering protects CP from a high packet rate and reduces the load on DP switches. An experiment reveals that with the mechanism, switches could significantly moderate the CPU loads, thereby preserving the space constraint of TCAM. However, the rate restriction mechanism in the technique presents some cases of packet lost and slight overhead. HiQoS [142] is also proposed by Jinyao et al. as a Multipath QoS solution. The PS scheme includes a module for the differential handling of flows with QoS requirements and a module that finds multiple paths based on a modified Dijkstra[54]. The controller uses the IP address of the source switch to separate several types of services and supplies diverse bandwidth assurances to each class. Bandwidth guarantee to specific traffic is achieved through the queuing mechanisms provided by the Openflow protocol [52]. HiQoS is bench marked with LiQoS and MiQoS.

The work in [143] proposed a framework for service differentiation support in SDN. It ensures the necessary QoS level for all multimedia applications. The approach leverages the controller's monitoring ability to get traffic statistics and network status every 3s. However, the controller might be overwhelmed with high overhead at this monitoring rate. The authors try to minimize that by restricting the statistics query to ingress switches only. In the work of Assefa and Ozkasap [144], a Machine Learning Framework for traffic aware energy efficient routing is proposed. The goal of MER-SDN is twofold, energy usage optimization and network performance. In a similar effort, Deng and Wang [145] applied Simulated Annealing (SA) Optimization to design a PSA to meet the specific QoS requirements of SDN-based IoT applications. AQRA classifies traffic into high, medium, and low priority using Class Identifier (QCI) obtained from the application profile set by the service providers. The initial path is determined using Dijkstra but updated later for each flow and placed in the switch flow table with rules designed by the SA-based routing module according to the application profile.

Table 6. Comparison of PSA with Flow Features Aware

Ref	Path Selection Criteria						Use Case / Aim				Design Objective			Solution Approach & BLA		Validation & Implementation Tools
	Parameters			No of Constraints												
	Link Quality	Critical Switch	Flow Features	MCP	CSP	MCSP	Custom	DCN	WAN	IoT	LB	QoS	RU	AI	MM	
[135]	X	X	✓	X	X	X	X	X	X	X	X	✓	X	X	LP	NOX
[136]	X	X	✓	X	✓	X	X	X	X	X	X	✓	X	X	A* Prune	FlowMonitorNS3
[137]	✓	X	✓	X	✓	X	✓	X	X	X	X	✓	X	X	LARAC	Floodlight Mininet
[138]	X	X	✓	X	X	X	X	X	X	X	✓	✓	X	X	LARAC	ODL
[20]	✓	X	✓	X	✓	X	X	X	X	X	X	✓	X	X	LARAC	Floodlight VLC MP
[95]	X	X	✓		X	✓	X	X	X	X	X	✓	X	X	Yen	X
[139]	✓	X	✓	X	✓	X	✓	X	X	X	X	✓	X	X	Yen	ITZ
[141]	X	✓	✓	X	X	X	X	X	X	X	✓	X	X	X	---	OVS
[142]	✓	X	✓	X	✓	X	✓	X	X	X	X	✓	X	X	Dijkstra	Floodlight, Mininet
[143]	✓	X	✓	X	✓	X	X	X	X	X	X	✓	✓	X	Dijkstra	POX
[144]	✓	X	✓	NA	NA	NA	X	NA	NA	NA	NA	NA	X	✓	NA	POX, SNDlib
[145]	✓	X	✓	X	X	✓	Campus	X	X	✓	✓	✓	X	SA	Dijkstra	Ryu, D-ITG

3.4. PSA WITH MACHINE LEARNING TECHNIQUES

Forecasting and classifying flow traffic is crucial to efficient resource utilization and QoS provisioning during PS for routing traffic in a modern network. The parameters included in the QoS specifications are usually captured in the SLA between service providers and subscribers. These parameters, described in section 2.4.2, are monitored and acquired from switches, ports, and flows using the OpenFlow built-in data collection module. Other metrics, such as residual bandwidth, link utilization, delay, and jitter, require extra effort to be acquired. There is a compelling need to efficiently measure these metrics and map each traffic with appropriate network resources to meet users' needs. Thus, additional intelligence is necessary to execute these tasks as desired. Thus, researchers have leveraged different Machine Learning (ML) techniques to synthesize network statistics controllers for traffic classification, routing, resource management, and load balancing. Traffic classification can be based on application or flow behaviours. The reason for using the former parameter is based on the need to separate Delay-Sensitive (DS) applications from Non-Delay-Sensitive (nDS). The DS application always required speedy detection and redistribution on the network to avoid SLA violations. However, with the wild upsurge of applications on the internet, it would be unrealistic to identify all the applications, especially in a large-scale network. The latter parameter help in separating EF from MF because the long-lived features of EF hurt the MF significantly. (See section 1.1 for detail). Different AI techniques can help detect, classify, and schedule each flows class as appropriate. Refer to Table 7 for a comparison summary of these techniques.

Cui and Xu [146] propose a PSA with load balancing in SDN based on multiple path features fed into Artificial Neural Network (ANN) model. ANN integrates the information and selects a path with a minimum aggregate load. The choice of the ANN to process the collected network statistics is due to its support for an infinite number of input vectors with undefined distribution in contrast to logistic regression and other probability methods. The technique goes through a Forward Propagation Learning (FPL) phase where the ABW, PLR, TL, and HC with a pre-set weight are supplied as input neurons. A Weight Adjustment (WA) phase is to adjust this weight until the fittest neural node is returned. This way, the controller finds a Least Loaded Path (LLP) to route traffic. Assefa and Ozkasap [144] propose a Machine Learning Framework for energy efficient routing and QoS optimization.

In a similar effort, Deng and Wang [145] applied Simulated Annealing (SA) to design a PSA to meet the specific QoS requirements of SDN-based IoT applications. AQRA incorporates a traffic classification module to categorize applications into high, medium, and low priority. The classification is according to QoS Class Identifier (QCI) obtained from the application profile set by the service providers. An initial path is determined using Dijkstra but updated later by SA for each flow. The SA-based module designs the routing rules according to the application profile. In another work [147], Energy Optimize Routing with Congestion Control for SDN WBAN is developed using Spider Monkey Optimization techniques. The network's weight/cost of available paths is modelled with residual energy level, link reliability, path loss, and queue length. Therefore, an optimum path among the paths is selected using the SMO algorithm. In a different approach, Naïve Bayes is used by El-Garoui et al. [148] to solve a routing problem. The solution optimizes Communication Overhead (CO) and Transmission Latency (TL) between pervasive nodes in SDN-VANET. The MLT influences CO reduction between the controller and RSU by predicting vehicle location as per RSU.

3.4.1. Genetic Algorithm (GA) Approach

Yu and Ke[149] acknowledge that video streaming is a pervasive killer application in the modern internet that require a highly efficient routing method to meet users' QoS demand. For this reason, they exploit the Genetic algorithm (GA) to develop a routing algorithm (GA-SDN) that can enhance video traffic over SDN. GA-SDN model the network as a connected graph with candidate paths from the source to the destination represented as (s_i *list,* s_j). For any ingress traffic, the algorithm identifies a video in two ways; (1) ToS/DSCP bits of the packet and (2) Port number. If any packets whose information matches any video stream protocol, GA-SDN will not forward it according to the default SP. Instead, the algorithm will check the link utilization to determine whether the available bandwidth can provide the required QoS support. The technique is benchmarked against BF Algorithm[56]. Similarly,[150] deploys a secured GA-Based module in an SDN controller to perform a route calculation task that selects a path with optimized energy-consumed nodes in an IoT environment. Block-Chain technology is used to maintain a list of malicious activities of nodes in the DP, which the GA-Based routing module consider when taking a PS decision.

In contrast to GA-SDN, Li et al. [151] use Non-dominated Sorting Genetic Algorithm (NSGA II) to model a multi-objective optimization PS decision in SDN. It contains Monitoring (NMM), Awareness(NAM), and Reconfiguration (NRM) Modules. NRM receives instruction from NMM to reroute traffic to a better path when link utilization is high, or AVB is less than the flow requirements. The authors claimed that using NSGA II influences the reduction of forwarding latency and packet loss ECMP [6]. The work in [152] is another example of a routing problem solved with NSGA II. The authors applied the algorithm to propose a secure routing with untrusted DP switches.

3.4.2. Reinforcement Learning Approach

Reinforcement Learning (RL)[153] leverage the monitoring module of OpenFlow to gather comprehensive

network statistic to build state representation space for the reward and action tuples to use in SDN. RL variants like Q-Leaning, Q-routing, and SARSA-Leaning have been applied in SDN to optimize different PS problems [154]–[159].

Conversely, [154] propose Q-FDBA to address video streaming problems related to Quality of experience (QoE) fairness. A Q-Leaning is used as the cluster decision algorithm to maximize the QoE. Other work by [155] combined State-Action-Reward-State-Action (SARSA)- with variable ε-Greedy function to solve SDN PS problems concerning congestion, packet queue waiting time, and transmission speed. In another approach, Huong et al. [157] combine RL with Deep Neural Network to develop a scheme called RLLP to address a load balance problem during PS. The reward function that guides the RLLP for load-balancing decisions considers the link delay, standard deviation, utilization, and discount rates. Whereas [158] uses RL components to propose a framework for traffic-aware energy-efficient routing HyMER. Similarly, in Shi et al. [159], the SARSA version of RL is also applied to develop a delay-aware PSA for SDN-supported power distribution application in an IoT environment (SDRS). SDRA uses an RL agent to adapt to the fluctuating network state and make a PS decision that improves system performance in terms of delay. Furthermore, an energy-efficient routing problem in large-scale SDN-IoT is also handled [160]. The authors propose two-level control mechanisms involving Multi-hop clustering MHC-RPL and a Q-Routing version of RL. However, in addition to its high convergence time, the requirement of a Q-learning algorithm to maintain a Q-table for storing *state*, *action*, and *reward* space information greatly limited its applicability to solving routing problems in SDN. Thus, to address the shortcomings of RL as experienced in Q_FDBA, Yu et al. [161] introduced a Deep Deterministic Policy Gradient (DDPG) mechanism to replace Q-table use with a neural network. The authors proposed Deep Reinforcement Learning (DROM) to optimize the routing procedure. DROM is benchmarked with OSPF concerning convergence time, delay, and throughput metrics.

Table 7. Comparison Table of PSA with Machine Learning

Ref	Path Selection Criteria						Use Case / Aim				Design Objective				Solution Approach & BLA		Validation & Implementation Tools
	Parameters			No of Constraints													
	Link Quality	Critical Switch	Flow Features	MCP	CSP	MCSP	IoT	DCN	WAN	Others	LB	QoS	FT	RU	AI	MM / Other	
[146]	✓	X	X	✓	X	X	X	✓	X	✓	✓	X	X	X	ANN		
[144]	✓	X	✓	X	X	X	X	✓	X	Energy	X	X	X	✓	✓	X	POX, Mininet SNDlib
[145]	✓	X	✓	X	X	✓	✓	X	X	Campus Network	✓	✓	X	X	SA	Dijkstra	Ryu, Mininet-WiFi, D-ITG
[147]	✓	X	X	✓	X	X	✓	X	X	WBAN	X	✓	X	✓	Spider Monkey		MATLAB
[148]	✓	X	X	X	✓	X	✓	X	X	VANET	X	✓	X	X	Naïve Bayes	X	Ryu, Mininet SUMO
[149]	✓	X	✓	✓	X	X	X	X	X	✓	X	✓	X	X	GA	X	Ns2, [108]myEvalSVC
[150]	X	✓	X	X	✓	X	✓	X	X	Security &Energy		✓	X	X	GA	X	MetaMask Ganache
[151]	✓	X	X	X	X	X	X	X	X	✓	✓	✓	X	✓	NSGA-II	X	Ryu, Iperf, Mininet
[152]	X	RL, Untruth	X	X	✓	X	X	X	X	Security	X	✓	X		NSGA-II	LP	Matlab
[154]	✓	X	✓	✓	X	X	X	X	X	✓	X	✓	X	X	Q-Leaning	X	ODL Mininet
[155]	✓	X	X	✓	X	X	X	X	X	✓	✓	✓	X	✓	SARSA & ε-Greedy	X	Mininet [107]
[157]	✓	X	X	✓	X	X	X	X	X	✓	✓	✓	X	✓	RL	X	Ryu, Mininet, DITG
[158]	✓	X	X	✓	X	X	X	✓	X	Energy	✓	X	X	✓	RL, Leaning	X	POX, MininetSNDlib
[159]	✓	X	X	X	✓	X	✓	X	X	PLC	X	✓	X	X	SARSA	X	--
[160]	✓	X	X	X	✓	X	✓	X	X	Energy	X	✓	X	✓	RL Q-Routing	X	Cooja, Contiki, RPL
[161]	✓	X	✓	X	✓	X	X	X	X	✓	X	✓	X	X	DRL	X	TensorFlow Keras,,OMNeT
[162], [163]	✓	✓	✓	X	✓	X	X	✓	X	X	X	✓	X	✓	DRL	X	OMNeT
[164]	✓	X	X	X	✓	X	X	✓	X	Backbone Network	X	✓	X	X	DRL, SA	X	OMNeT++
[165]	✓	X	X	X	✓	X	X	✓	X	Sprint	X	✓	X	X	DRL	X	OMNeT++
[166]	✓	X	X	X	X	✓	X	X	X	NSF NetARPANet	X	✓	X	X	DRL	Yen	Ryu, Mininet
[167]	✓	X	X	✓	X	X	✓	X	X	Security	X	✓	X	✓	DRL	X	Tensorflow
[168]	✓	✓	✓	X	X	✓	✓	X	X	FASNET		✓	X	X	ACO	X	TinyOS, MintRoute
[169]	✓	X	X	✓	X	X	X	✓	X	Energy	✓	✓	X	✓	ACO	LP	Floodlight, Mininet, Iperf3

The works in [162][163] also did a similar thing with DRL to address the problem of amalgamating EF with MF on the same path. The authors consider multiple network resources such as switch cache, link bandwidth to map, and schedule flow according to QoS requirements. In another effort, Maheswari et al. [164] also involved DRL to optimize routing procedures in SDN. The techniques aim to optimize network delay, network operation, and maintenance costs. The Traffic Matrix (TM), link weight, and network delay are represented by the state, action, and reward tuples. Likewise, in the work of Xu et al. [165], a DRL technique is integrated into SDN PSA to optimize performance concerning delay, hop count, and throughput. Chen et al. [166] formulated a traffic engineering problem in SDN. They developed an RL-Routing, based on DRL to find an optimized solution. Network delay and throughput are modelled in the state representation space of DRL. The reward function taps these metrics from this space to build an action space. The action space comprises a list of all paths and their associated cost (*reward*). Furthermore, the vulnerability of PSA in SDN to dynamic change of flow control rules at the time of malicious activities motivates Gou [167] to propose a DRL-based QoS- Aware security routing algorithm (DQSP) for IoT applications. DQSP is modelled to be immune to Gray Hole Attacks (GHA) and DDoS attacks. DQSP is evaluated in terms of PDR, E2E Delay, and probability of path attack (PA).

However, the *state* space of RL is overpopulated with many metrics, whose extraction and calculation from the MM might overwhelm the controller.

3.4.3. Ant Colony Optimization Approach

Ant-Colony Optimization (ACO) technique is exploited to develop a Traffic Differentiated Routing (TDR) [168]. The authors formulate a transmission reliability and prediction model as an LP. The model considers link availability and node forwarding ability to develop a TDR to guarantee the QoS of Flying Ad-hoc Sensor Networks (FASNETs). The also model seeks to arrive at a spanning tree that minimizes the average delay and improves data integrity for reliability-sensitive applications. The NP-hard nature of the problem compels the authors to seek a solution from the ACO. Torkzadeh et al. [169] incorporate an energy optimization constraint for a load-balancing routing problem in SDN. The authors propose a two-phase solution to solve the problem. A minimum graph ACO is employed in the first phase to prune the network topology. All inconsequential DP switches are discharged during routing, leaving only an energy-minimized sub-graph. In the second phase, a QoS-weighted PS technique is developed to route the traffic along paths with a balanced load based on a dynamic threshold value.

4. OPEN CHALLENGES FOR FUTURE RESEARCH

4.1. PSA WITH LINK QUALITY AWARE

4.1.1. Network State Information Problem

All PSA considering link quality parameters, whether static or dynamic (See section 2.4 for details), depend on Network State Information (NSI) for the PSD. The controller collects the NSI from the DP at time intervals and feeds it into the PSA to make the decision. PSA assumes this information to be accurate and adequate to make the right decision. The NSI is collected through sampling or polling techniques. However, if the former technique is used, the NSI might be inadequate or inaccurate at the time of PSD. However, it is a big challenge to maintain a high level of statistic collection accuracy in practice due to the periodic manner of the collection. Depending on the collection interval, the PSA might be called upon to take PSD during this interval. In this circumstance, the PSA must use the old information available. However, this might mislead PSA to make false positive or false negative decisions. Selecting the sampling period depends on the topology size and density if an overhead reduction is critical in the network. Therefore, it will be interesting to undertake further study to explore how the NSI collection period by the controller can dynamically adapt to topology and traffic changes. The study should provide a balance between adequate NSI for accurate PSD and message collection overhead on the controller. The idea of modelling link cost as a probabilistic metric is an exciting possibility in addressing inaccurate NSI.

Furthermore, it is essential to note that PSA with multiple constraints might not necessarily be an NP-complete problem. Instead, it further depends on other factors, such as topology size. Therefore, solution searching using an exact algorithm approach should be able to distinguish the scenario for which the complexity of the problem is polynomial to refine the solution searching approach.

4.1.2. Service Level Agreement (SLA) Evaluation Inconsistency

Subjectivity in evaluating satisfaction and compliance level of the QoS requirement captured in SLA by customers and service providers affects the trust and future relationship between them. The situation might sometimes lead to substantial financial loss for both parties [5]. This scenario is possible due to the lack of a QoS satisfaction measurement framework that can meticulously verify, audit, and validate the guarantee level pledged to each as defined in the SLA. Therefore, it is an exciting research concept to imagine a quantitative and unified technique to objectively provide a detailed evaluation of such agreement. Incorporating AI techniques might be helpful.

4.2 PSA WITH CRITICAL SWITCH AWARE

Several initiatives for different use cases have been proposed in recent years to solve a path selection problem for SDN. However, only a few have examined how SDN performs in dynamic, large-scale telecommunications networks, where heavy traffic flows are constantly generated. Similarly, the limited switch memory influences SDN performance in large-scale networks due to increased update operations and security-related risks. It is proven that in such an environment, many flows carry many packets in a short amount of time. Switch memory cannot hold the necessary amount of flow entries. This limitation remains one of the research problems that need additional studies. i.e., managing massive flows with many packets in the power-hungry tiny switch flow table.

4.2.1. Rule Update Operation on Switch flow table

Flow-table TCAM uses exact and wildcard matching rules for update operations [1]. Therefore, an effective PSA should understand the suitability of each possibility while formulating an optimization objective. Unlike exact rules, which must provide individual flow rules for each entry in a switch flow table. On the other hand, the wildcard rules allow multiple flows to be composed as one. This way, all identical entries can be recycled among various flows, thereby minimizing the number of entries and the overhead of frequent flow setup requests. PSA should also consider that TCAM operates slowly during an update operation. Therefore, packets may experience delays, especially in large networks. Hence, failure recovery must strictly comply with the CGN latency requirements.

4.2.2. Reactive and Proactive Rule Installation Hybridization

The number of critical switches in a network can be reduced by hybridizing reactive and proactive flow rule installation approaches. Employing a proactive approach for time-constraint applications is a preference. In contrast, best-effort or applications without deadline violation can embrace the reactive approach. For PSA to effectively utilize the limited TCAM space and decrease overhead and packet delay, efficient flow rule allocation should incorporate both reactive and proactive approaches. Therefore, designing a PSA while taking traffic variations along with these approaches will be an intriguing research topic. Re-routing rules should be executed in under 25ms to satisfy the strict QoS specifications of real-time applications.

4.3. PSA WITH FLOW CHARACTERISTICS AWARE

Traffic management has become essential to computer network design requirements [170]. The concept impacts various computer network areas of concern. E.g., Security issues such as intrusion and violation detection, glitch discovery such as TCP incast [171], and anomaly tracking such as DDoS [172]. Other areas are Traffic Engineering (TE), Quality of Service, Resource Management, Energy, PS Optimization and flow rerouting, Service Level Agreement (SLA) Appraisal, and Auditing. Most of the PSA with traffic awareness involves detecting and classifying a flow according to priority, size, or duration. The following issues are areas of concern that call for further investigation.

4.3.1. Flow Feature Selection Dilemma

Flow statistics such as size, duration, rate, or burst are used to compare a flow against a pre-defined threshold to classify flows accordingly. The threshold value can be fixed or adaptive. Selecting the flow feature for flow classification depends on the design objectives of the problem. There is no unanimous settlement on

4.3.2. Threshold Value Determination Challenge

The selection of a threshold value by most of the existing flow-aware PSA lacks a specific and systematic justification. Both static and adaptive threshold values run into this problem. Most of the papers reviewed in section 3.3 cited previously published works, for which, at the end of the citation chain, there is no justification for the chosen threshold. That trend is observed across numerous works. As a result, rather than being systematic, threshold selection in existing studies appears ad hoc. A study in [79] suggests that the preconfigured fixed threshold parameter for EF detection incurred high detection error rates because of its ability to adapt dynamically in real-time to constant traffic variability in a contemporary network. The tendency to report false positive and false negative errors is significant. The techniques do not adequately consider the dynamics of network traffic. Instead of being static, network traffic is dynamic. It may alter over time in response to variables such as the time of day, configuration changes, failures, or adjustments to the topology and instrumentation. Every time the threshold needs to be adjusted, the classification must be repeated to reflect the dynamic nature of the network. Manual threshold adjustment is nonetheless impossible due to the nondeterministic and frequent changes in network traffic conditions.

4.3.3. Flow Identification Overhead and Accuracy challenge

Flow detection mechanisms incorporated in PS schemes incurred some overhead in the stage of statistics gathering. Statistics-gathering techniques can be through sampling, polling, triggering, or hash functions. Each of these techniques has specific strengths and weaknesses regarding application areas and scenarios.

Sampling is one of the most adopted methods to acquire and profile networks for traffic analysis. Popular sampling method such as sFlow [35] has been incor-

porated by several (EFDM). It is scalable and adaptable to traffic heterogeneity in DCN [79], as several switches can be monitored efficiently by the sFlow protocol. In contrast, polling techniques deal with every flow entry in the flow table. A prior study [79] reveals that 64kb or 88bytes messages will return to the controller for each polling request. This data might not seem much, but if the average DCN size with 100 edge switches is factored in, that can reach up to 10mf/s on average. Summing up that for 88bytes, the total data to return will be up to 65GB. This data will dominate the limited bandwidth of the northbound interface. Therefore, employing this method to gather network statistics for EF identification in DCN will cause significant overhead and bandwidth mismanagement. Ref [4] suggests that EF are few in a DCN; thus, it is inefficient to accumulate data on each flow to detect EF in the network.

On the contrary, the triggering approach sets up a sniffer agent or applications at the end host [84]. The technique detects and classifies flows before transmission directly and precisely. Once the dimensions of a flow (e.g., socket buffer, flow size) surpass a set-up threshold, the EFDM decides that the flow is indeed an EF. The method reduces the overhead. However, it is impractical in DCN due to the requirement of changing the operation of each end host. For these reasons, it is interesting to conduct further studies in that direction.

4.4. PSA WITH MACHINE LEARNING TECHNIQUES

One of the significant challenges of employing ML techniques to solve real-life problems is the data set availability for model training. Privacy and confidentiality issues associated with computer networks make sharing this data difficult and scarce. The situation is worse in SDN because the technology is still emerging. For this reason, future research should be directed toward building and expanding the existing Opensource data set, such as SDN ITZ.

Secondly, PSA sought through ML should be adaptable to factors like communication mediums and applicable to technologies like Low-power and lossy network (LLNs) use in WSN. The forms of communication such as unicast, multicast, or broadcast applications (use case) like fog computing, DCN, and 5G. Likewise, traffic heterogeneity should be considered along with prediction patterns to guide the adaptation. Thus, NSI must be obtained regularly for accurate prediction and PS policy formulation. Thirdly, the ML model should incorporate safe mechanisms to preserve NSI integrity. This mechanism inclusion is necessary to avoid inaccurate findings or inconsistent decisions because the data collection could be impeded or even altered.

4.4.1. Training Dataset Scarcity Problem

One of the significant challenges of employing ML to solve real-life problems is the data set availability for model training. Privacy and confidentiality issues associated with computer networks make sharing this data difficult and scarce. The situation is worse in SDN because the technology is still emerging. For this reason, future research should be directed toward building and expanding the existing Opensource data set, such as SDN ITZ.

4.4.2. Intelligent Flow Table Management

In large networks, many flows arrive regularly, necessitating the installation of relevant rules in a flow table to occupy substantial storage space. Most SDN TCAM-related solutions in the literature are only evaluated on small networks [173]. What works for these networks cannot be compared to the number of devices in an extensive network such as WAN, DCN, or (IoT). ML techniques are handy for managing devices like switches and controllers. Many ML approaches, however, concentrate on flow classification and flow monitoring. Most research focuses on selecting the optimum traffic flow to be installed in advance rather than forecasting traffic flow for real-time applications and best-effort traffic. OpenFlow provides built-in data collection that stores flow-statistic like packet counts. This statistical data shows how frequent traffic flows. It will be interesting to develop a plan that takes advantage of this built-in data collection to reduce TCAM space usage. So the flow matching rate can speed up. Perhaps by applying fuzzy theory in the choice of recurrently used flow rules to be placed in the flow table.

5. CONCLUSION

SDN provides flexibility in managing the complexity and demand of our modern network, which was unable to be provided by traditional architecture. In this paper, we picked the network management task of path selection for routing traffic and reviewed the existing algorithms under four categories. (1) To guide their PSD, the PSAs with static link quality under different traffic conditions or dynamic link quality. (2) The PSAs that evaluate the criticality of a switch in terms of an update operation, flow table, and port capacity to guide PSD. (3) The PSAs that consider the traffic flow heterogeneities in terms of size, duration, burst, and priority to guide PSD. (4) The PSAs that use ML for PSD decisions. For each category, the papers were reviewed considering their path selection criteria, use case, design objectives, solution approach, baseline algorithms, and validation approaches. A comparison summary table is given at the end of each category. Based on the review, some persistent challenges related to each category are identified and recommended for further study. For instance, inaccurate and inadequate NSI in PSA with dynamic link quality and SLA evaluation are challenges that need further study. Secondly, rule update operations overhead is a challenge in PSA with critical switch awareness. The paper suggests hybridizing reactive and proactive rule installation approaches to optimize PSA convergence. Thirdly, the paper identifies flow feature selection and threshold

value challenges during flow classification. Furthermore, the paper identifies flow identification overhead and accuracy as issues requiring further research efforts associated with PSAs considering traffic dynamics. Lastly, the paper identifies training dataset scarcity as a problem faced by PSA employing ML.

6. REFERENCES

[1] B. Isyaku, K. A. Bakar, M. S. M. Zahid, E. H. Alkhammash, F. Saeed, F. A. Ghaleb, "Route path selection optimization scheme based link quality estimation and critical switch awareness for software-defined networks", Applied Sciences, Vol. 11, No. 19, 2021,

[2] Z. R. Alashhab, M. Anbar, M. M. Singh, Y. B. Leau, Z. A. Al-Sai, S. A. Alhayja'a, "Impact of Coronavirus Pandemic Crisis on Technologies and Cloud Computing Applications", Journal of Electronic Science and Technology, Vol. 19, No. 1, 2021, pp. 25-40.

[3] B. Isyaku, M. S. Mohd Zahid, M. Bte Kamat, K. Abu Bakar, F. A. Ghaleb, "Software Defined Networking Flow Table Management of OpenFlow Switches Performance and Security Challenges: A Survey", Futur. Internet, Vol. 12, No. 9, 2020, p. 147.

[4] C. Y. Lin, C. Chen, J. W. Chang, Y. H. Chu, "Elephant flow detection in datacenters using OpenFlow-based Hierarchical Statistics Pulling", Proceedings of the IEEE Global Communications Conference, Austin, TX, USA, 8-12 December 2014, pp. 2264-2269.

[5] S. Khan, F. K. Hussain, O. K. Hussain, "Guaranteeing end-to-end QoS provisioning in SOA based SDN architecture: A survey and Open Issues", Future Generation Computer Systems, Vol. 119, 2021, pp. 176-187,

[6] C. Hopps, "Analysis of an Equal-Cost Multi-Path Algorithm", Network Working Group C. Hopps Request for Comments: 2992 NextHop Technologies Category: Informational, 2000, pp. 130-139.

[7] M. Al-Fares, A. Loukissas, A. Vahdat, "A scalable, commodity data centre network architecture", Proceedings of the ACM SIGCOMM Conference on Data Communication, 2008, p. 63

[8] H. Balakrishnan, M. Stemm, S. Seshan, R. H. Katz, "Analyzing stability in wide-area network performance", Performance Evaluation Review, Vol. 25, No. 1, 1997, pp. 2-12.

[9] M. Casado et al. "Rethinking enterprise network control", IEEE/ACM Transactions on Networking, Vol. 17, No. 4, 2009, pp. 1270-1283.

[10] A. Gupta et al. "An industrial-scale software defined internet exchange point", Proceedings of the 13th Usenix Conference on Networked Systems Design and Implementation, 2016, pp. 1-14 .

[11] I. Z. Bholebawa, R. K. Jha, U. D. Dalal, "Performance Analysis of Proposed Network Architecture: OpenFlow vs Traditional Network", Wireless Personal Communications, Vol. 86, No. 2, 2016, pp. 943-958,

[12] J. M. Jaffe, F. H. Moss, "A Responsive Distributed Routing Algorithm for Computer Networks", IEEE Transactions on Communications, Vol. 30, No. 7, 1982, pp. 1758-1762.

[13] R. Gallager, "A Minimum Delay Routing Algorithm Using Distributed Computation", IEEE Transactions on Communications, Vol. 25, No. 1, 1977, pp. 73-85.

[14] J. M. Mcquillan, I. Richer, E. C. Rosen, "The New Routing Algorithm for the ARPANET", IEEE Transactions on Communications, Vol. 28, No. 5, 1980, pp. 711-719.

[15] A. A. Neghabi, N. J. Navimipour, M. Hosseinzadeh, A. Rezaee, "Load Balancing Mechanisms in the Software Defined Networks: A Systematic and Comprehensive Review of the Literature", IEEE Access, Vol. 6, 2018, pp. 14159-14178.

[16] D. Kreutz, F. M. V. Ramos, P. E. Verissimo, C. E. Rothenberg, S. Azodolmolky, S. Uhlig, "Software-defined networking: A comprehensive survey", Proceedings of IEEE, Vol. 103, No. 1, 2015, pp. 14-76.

[17] M. Ejaz Ahmed, H. Kim, "DDoS attack mitigation in the internet of things using software-defined networking", Proceedingd of the IEEE Third International Conference on Big Data Computing Service and Applications, Redwood City, CA, USA, 2017, pp. 271-276.

[18] Q. Gao, W. Tong, S. Kausar, L. Huang, C. Shen, S. Zheng, "Congestion-aware multicast plug-in for an SDN network operating system", Computer Networks, Vol. 125, 2017, pp. 53-63.

[19] P. Yi, T. Hu, Y. Hu, J. Lan, Z. Zhang, Z. Li, "SQHCP: Secure-aware and QoS-guaranteed heterogeneous

controller placement for software-defined networking", Computer Networks, Vol. 185, No. 2020, p. 107740.

[20] H. E. Egilmez, S. T. Dane, K. T. Bagci, A. M. Tekalp, "OpenQoS: An OpenFlow controller design for multimedia delivery with end-to-end Quality of Service over Software-Defined Networks", Proceedings of the Asia Pacific Signal and Information Processing Association Annual Summit and Conference, Hollywood, CA, USA, 3-6 December 2012.

[21] H. Zhong, Y. Fang, J. Cui, "Reprint of 'LBBSRT: An efficient SDN load balancing scheme based on server response time", Future Generation Computer Systems, Vol. 80, 2018, pp. 409-416.

[22] M. W. Hussain, B. Pradhan, X. Z. Gao, K. H. K. Reddy, D. S. Roy, "Clonal selection algorithm for energy minimization in software-defined networks", Applied Soft Computing, Vol. 96, 2020, p. 106617.

[23] Z. N. Abdullah, I. Ahmad, I. Hussain, "Segment routing in software-defined networks: A survey", IEEE Communications Surveys & Tutorials, Vol. 21, No. 1, 2019, pp. 464-486.

[24] P. Murali Mohan, T. Truong-Huu, M. Gurusamy, "Fault tolerance in TCAM-limited software-defined networks", Computer Networks, Vol. 116, 2017, pp. 47-62.

[25] B. Yan, Q. Liu, J. Shen, D. Liang, B. Zhao, L. Ouyang, "A survey of low-latency transmission strategies in software-defined", Computer Science Review, Vol. 40, 2021, p. 100386.

[26] J. Mei, K. Zheng, L. Zhao, Y. Teng, X. Wang, "A Latency and Reliability Guaranteed Resource Allocation Scheme for LTE V2V Communication Systems", IEEE Transactions on Wireless Communications, Vol. 17, No. 6, 2018, pp. 3850-3860.

[27] V. Seedha Devi, T. Ravi, S. B. Priya, "Cluster Based Data Aggregation Scheme for Latency and Packet Loss Reduction in WSN", Computer Communications, Vol. 149, No. 2019, pp. 36-43.

[28] Z. Cai, A. Cox, E. T. S. Ng, "Maestro: A System for Scalable OpenFlow Control", Cs.Rice.Edu, 2011 p. 10.

[29] N. Gude, J. Pettit, B. Pfaff, N. Mckeown, S. Shenker, "NOX : Towards an Operating System for Networks", ACM SIGCOMM Computer Communication Review, Vol. 38, No. 3, 2008, pp. 105-110.

[30] L. Zhu, M. M. Karim, K. Sharif, F. Li, X. Du, M. Guizani, "SDN Controllers: Benchmarking & Performance Evaluation", arXiv:1902.04491, 2019.

[31] M. F. Rangkuty, R. Muslim, T. Ahmad, M. H. A. Al-Hooti, "Path selection in software-defined network data plane using a least loaded path", Proceedings of the International Conference on Advanced Computer Science and Information Systems, Depok, Indonesia, 17-18 October 2020, pp. 135-140.

[32] E. T. B. Hong, C. Y. Wey, "An optimized flow management mechanism in OpenFlow network", Proceedings of the International Conference on Information Networking, Da Nang, Vietnam, 11-13 January 2017, pp. 143-147.

[33] A. R. Curtis, J. C. Mogul, J. Tourrilhes, P. Yalagandula, P. Sharma, S. Banerjee, "DevoFlow: Scaling flow management for high-performance networks", Computer Communications Review, Vol. 41, No. 4, 2011, pp. 254-265.

[34] M. Yu, J. Rexford, M. J. Freedman, J. Wang, "Scalable flow-based networking with DIFANE", ACM SIGCOMM Computer Communications Review, Vol. 40, No. 4, 2010, pp. 351-362.

[35] P. Phaal, "Traffic Monitoring using sFlow", www.sflow.org/sFlowOverview.pdf. (accessed: 2022)

[36] A. S. Iyer, V. Mann, N. R. Samineni, "SwitchReduce: Reducing switch state and controller involvement in OpenFlow networks", Proceedings of the IFIP Networking Conference, Brooklyn, NY, USA, 22-24 May 2013, pp. 1-9.

[37] E. Akin, T. Korkmaz, "Comparison of Routing Algorithms with Static and Dynamic Link Cost in Software Defined Networking (SDN)", IEEE Access, Vol. 7, 2019, pp. 148629-148644.

[38] F. Y. Okay, S. Ozdemir, "Routing in Fog-Enabled IoT Platforms: A Survey and an SDN-Based Solution", IEEE Internet of Things Journal, Vol. 5, No. 6, 2018, pp. 4871-4889.

[39] M. Karakus, A. Durresi, "Quality of Service (QoS) in Software Defined Networking (SDN): A survey", Journal of Network and Computer Applications, Vol. 80, no, 2016, pp. 200-218.

[40] S. Tomovic, I. Radusinovic, N. Prasad, "Performance comparison of QoS routing algorithms applicable to large-scale SDN networks", Proceedings of the International Conference on Computer as a Tool, Salamanca, Spain, 8-11 September 2015.

[41] J. W. Guck, A. Van Bemten, M. Reisslein, W. Kellerer, "Unicast QoS Routing Algorithms for SDN: A Comprehensive Survey and Performance Evaluation", IEEE Communications Surveys & Tutorials, Vol. 20, No. 1, 2018, pp. 388-418.

[42] M. R. Belgaum, S. Musa, M. M. Alam, M. M. Su'Ud, "A Systematic Review of Load Balancing Techniques in Software-Defined Networking", IEEE Access, Vol. 8, 2020, pp. 98612-98636.

[43] M. Hamdan et al. "A comprehensive survey of load balancing techniques in a software-defined network", Journal of Network and Computer Applications, Vol. 174, No. 2020, 2020, p. 102856.

[44] F. Chahlaoui, H. Dahmouni, "A Taxonomy of Load Balancing Mechanisms in Centralized and Distributed SDN Architectures", SN Computer Science, Vol. 1, No. 5, 2020, pp. 1-16.

[45] B. Heller, R. Sherwood, N. Mckeown, "The controller placement problem", Computer Communications Review, Vol. 42, No. 4, 2012, pp. 473-478.

[46] J. Cui, Q. Lu, H. Zhong, M. Tian, L. Liu, "A Load-Balancing Mechanism for Distributed SDN Control Plane Using Response Time", IEEE Transactions on Network and Service Management, Vol. 15, No. 4, 2018, pp. 1197-1206.

[47] R. Wazirali, R. Ahmad, S. Alhiyari, "Sdn-OpenFlow topology discovery: An overview of performance issues", Applied Sciences, Vol. 11, No. 15, 2021.

[48] R. Amin, E. Rojas, A. Aqdus, S. Ramzan, D. Casillas-Perez, J. M. Arco, "A Survey on Machine Learning Techniques for Routing Optimization in SDN", IEEE Access, Vol. 9, 2021, pp. 104582-104611.

[49] M. F. Tuysuz, Z. K. Ankarali, D. Gözüpek, "A survey on energy efficiency in software-defined networks", Computer Networks, Vol. 113, 2017, pp. 188-204.

[50] B. G. Assefa and Ö. Özkasap, "A survey of energy efficiency in SDN: Software-based methods and optimization models", Journal of Network and Computer Applications, Vol. 137, No. 2018, 2019, pp. 127-143.

[51] F. Dabaghi, Z. Movahedi, R. Langar, "A survey on green routing protocols using sleep-scheduling in wired networks", Journal of Network and Computer Applications, Vol. 77, No. 2016, 2017, pp. 106-122.

[52] N. McKeown et al., "OpenFlow", ACM SIGCOMM Computer Communications Review, Vol. 38, No. 2, 2008, pp. 69-74.

[53] Course Hero, "Lab 4-1-Ryu Introduction" (accessed: 2022)

[54] E. W. Dijkstra, " A note on two problems in connexion with graphs", Numerische Mathematik, Vol. 271, 1959, pp. 269-271.

[55] P. E. Hart, N. J. Nilsson, B. Raphael, "Formal Basis for the Heuristic Determination eijj ", Systems Science and Cybernetics, Vol. 4, No. 2, 1968, pp. 100-107.

[56] R. Bellman, "On a Rooting Problem", 1956.

[57] J. Y. Yen, " Finding the K Shortest Loopless Paths in a Network", Management Science, Vol. 17, No. 11, 1971, pp. 712-716.

[58] G. Liu and K. G. Ramakrishnan, "A *Prune: An algorithm for finding K shortest paths subject to multiple constraints", Proceedings of the Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society, Anchorage, AK, USA 22-26 April 2001, pp. 743-749

[59] R. K. Ahuja, "Network Flows", Indian Institute of Technology, 1993.

[60] M. R. Garey, D. S. Johnson, "Computers, Complexity, and Intractability", Computers and Intractability: A Guide to the Theory of NP-Completeness, 1979, p. 115.

[61] M. I. Hamed, B. M. ElHalawany, M. M. Fouda, A. S. T. Eldien, "A novel approach for resource utilization and management in SDN", Proceedings of the 13th International Computer Engineering Conference, Cairo, Egypt, 27-28 December 2018, pp. 337-342, ,

[62] N. Thazin, K. M. Nwe, Y. Ishibashi, "Resource Allocation Scheme for SDN-Based Cloud Data Center Network", Procedings of ICCA, 2019 pp. 15-22.

[63] M. Jeon, N. Kim, Y. Jang, B. D. Lee, "An efficient network resource management in SDN for cloud services", Symmetry., Vol. 12, No. 9, 2020.

[64] D. Erickson, "The Beacon OpenFlow controller", Proceedings of the second ACM SIGCOMM workshop on Hot topics in software defined networking, August 2013, pp. 13-18.

[65] M. Khorramizadeh and V. Ahmadi, "Capacity and load-aware software-defined network controller placement in heterogeneous environments", Computer Communications, Vol. 129, 2018, pp. 226-247.

[66] M. T. I. Ul Huque, W. Si, G. Jourjon, V. Gramoli, "Large-Scale Dynamic Controller Placement", IEEE Transactions on Network and Service Management, Vol. 14, No. 1, 2017, pp. 63-76.

[67] W. H. F. Aly, "Generic Controller Adaptive Load Balancing (GCALB) for SDN Networks", Journal of Computer Networks and Communications, Vol. 2019, 2019, pp. 1-9.

[68] K. S. Sahoo et al., "ESMLB: Efficient Switch Migration-Based Load Balancing for Multicontroller SDN in IoT", IEEE Internet of Things Journal, Vol. 7, No. 7, 2020, pp. 5852-5860.

[69] W. H. F. Aly, "Generic Controller Adaptive Load Balancing (GCALB) for SDN Networks", Journal of Computer Networks and Communications, Vol. 2019, 2019, pp. 1-9.

[70] A. Mahjoubi, O. Zeynalpour, B. Eslami, N. Yazdani, "LBFT: Load Balancing and Fault Tolerance in distributed controllers", Proceedings of the International Symposium on Networks, Computers and Communications, June 2019, pp. 1-6.

[71] H. Chen, G. Cheng, Z. Wang, "A game-theoretic approach to elastic control in software-defined networking", China Communications, Vol. 13, No. 5, 2016, pp. 103-109.

[72] C. Wang, B. Hu, S. Chen, D. Li, B. Liu, "A Switch Migration-Based Decision-Making Scheme for Balancing Load in SDN", IEEE Access, Vol. 5, 2017, pp. 4537-4544.

[73] B. Isyaku, K. bin A. Bakar, F. A. Ghaleb, M. S. M. Zahid, "Path Selection With Critical Switch-Aware For Software Defined Networking", Proceedings of the IEEE Symposium on Wireless Technology & Applications, Shah Alam, Malaysia, 17 August 2021, pp. 22-26.

[74] L. C. Freeman, "Centrality in social networks conceptual clarification", Social Networks, Vol. 1, No. 3, 1978, pp. 215-239.

[75] L. C. Freeman, "A Set of Measures of Centrality Based on Betweenness", Sociometry, Vol. 40, No. 1, 1977, p. 35.

[76] V. Latora, M. Marchiori, "A measure of centrality based on network efficiency", New Journal of Physics, Vol. 9, No. 2008, 2007,

[77] B. Fan, Z. Li, N. Shu, Y. Li, "Identification of Key Nodes Based on PageRank Algorithm", Proceedings of the IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference, Chongqing, China, 12-14 March 2021, pp. 2398-2402.

[78] A. Pekar, A. Duque-Torres, W. K. G. Seah, O. M. Caicedo Rendon, "Towards threshold-agnostic heavy-hitter classification", International Journal of Network Management, No. 2021, pp. 1-22,

[79] C. Bi, X. Luo, T. Ye, Y. Jin, "On precision and scalability of elephant flow detection in a data centre with SDN", Proceedings of the IEEE Globecom Workshops, Atlanta, GA, USA, 9-13 December 2013, pp. 1227-1232.

[80] M. Afaq, S. U. Rehman, W. C. Song, "A Framework for Classification and Visualization of Elephant Flow in SDN-Based Networks", Procedia Computer Science, 2015, Vol. 65, No. Iccmit, 2021, pp. 672-681.

[81] S. Sarvotham, R. Riedi, R. Baraniuk, "Connection-level analysis and modelling of network traffic", Proceedings of the First ACM SIGCOMM Workshop on Internet Measurement, 2001, p. 99.

[82] S. Shakkottai, N. Brownlee, K. Claffy, "A study of burstiness in TCP flow", Lecture Notes in Computer Science, Vol. 3431, 2005, pp. 13-26, ,

[83] K. C. Lan, J. Heidemann, "A measurement study of correlations of Internet flow characteristics", Computer Networks, Vol. 50, No. 1, 2006, pp. 46-62.

[84] A. R. Curtis, W. Kim, P. Yalagandula, "Mahout: Low-overhead datacenter traffic management using end-host-based elephant detection", Proceedings IEEE INFOCOM, Shanghai, China, 10-15 April 2011, pp. 1629-1637.

[85] X. Li and C. Qian, "Low-complexity multi-resource packet scheduling for network function virtualization", Proceedings IEEE INFOCOM, Hong Kong, China, 26 April - 1 May 2015, pp.1400-1408.

[86] F. Carpio, A. Engelmann, A. Jukan, "DiffFlow: Differentiating short and long flows for load balancing in data centre networks", Proceedings of the IEEE Global Communications Conference, December 2016.

[87] F. Estrada-Solano, O. M. Caicedo, N. L. S. Da Fonseca, "NELLY: Flow Detection Using Incremental Learning at the Server Side of SDN-Based Data Centers", IEEE Transactions on Industrial Informatics, Vol. 16, No. 2, 2020, pp. 1362-1372.

[88] S. C. Madanapalli, M. Lyu, H. Kumar, H. H. Gharakheili, V. Sivaraman, "Real-time detection, isolation, and monitoring of elephant flows using commodity SDN system", Proceedings of the IEEE/IFIP Network Operations and Management Symposium, Taipei, Taiwan, 23-27 April 2018, pp. 1-5.

[89] R. Liu, H. Gu, X. Yu, X. Nian, "Distributed flow scheduling in energy-aware data centre networks", IEEE Communications Letters, Vol. 17, No. 4, 2013, pp.801-804.

[90] M. Chiesa, G. Kindler, M. Schapira, "Traffic Engineering with equal-cost-multipath: An algorithmic perspective", IEEE/ACM Transactions on Networking, Vol. 25, No. 2, 2017, pp. 779-792.

[91] S. C. Chao, K. C. J. Lin, M. S. Chen, "Flow Classification for Software-Defined Data Centers Using Stream Mining", IEEE Transactions on Services Computing, Vol. 12, No. 1, 2019, pp. 105-116.

[92] C. L. Hu, C. Y. Hsu, S. E. Khuukhenbaatar, Y. Dashdorj, Y. Dong, "Path Selection with Joint Latency and Packet Loss for Edge Computing in SDN", Proceedings of the 20th Asia-Pacific Network Operations and Management Symposium, Matsue, Japan, 18-20 September 2019.

[93] A. Alnajim and S. Salehi, "Incremental Path-Selection and Scheduling for Time-Sensitive Networks", Proceedings of the IEEE Global Communications Conference, Waikoloa, HI, USA, 9-13 December 2019.

[94] K. Gao, C. Xu, J. Qin, S. Yang, L. Zhong, G. M. Muntean, "QoS-driven Path Selection for MPTCP: A Scalable SDN-assisted Approach", Proceedings of the IEEE Wireless Communications and Networking Conference, Marrakesh, Morocco, 15-18 April 2019,

[95] N. Saha, S. Member, S. Bera, S. Member, "Sway : Traffic-Aware QoS Routing", IEEE Transactions on Emerging Topics in Computing, 2018. (in press)

[96] L. R. Prete, A. A. Shinoda, C. M. Schweitzer, R. L. S. De Oliveira, "Simulation in an SDN network scenario using the POX Controller", Proceedings of the IEEE Colombian Conference on Communications and Computing, Bogota, Colombia, 4-6 June 2014,

[97] A. Botta, A. Dainotti, A. Pescapé, "A tool for the generation of realistic network workload for emerging networking scenarios", Computer Networks, Vol. 56, No. 15, 2012, pp. 3531-3547.

[98] C. Perner and G. Carle, "Comparison of Optimization Goals for Resilient Routing", Proceedings of the IEEE International Conference on Communications Workshops, May 2019, pp. 1-6.

[99] R. Khalili, Z. Despotovic, A. Hecker, "Flow Setup Latency in SDN Networks", IEEE Journal on Selected Areas in Communications, Vol. 36, No. 12, 2018, pp. 2631-2639.

[100] H. K. Ravuri, M. T. Vega, T. Wauters, B. Da, A. Clemm, F. De Turck, "An Experimental Evaluation of Flow Setup Latency in Distributed Software Defined Networks", Proceedings of the IEEE Conference on Network Softwarization, June 2021, pp. 432-437.

[101] L. V. Morales, A. F. Murillo, S. J. Rueda, "Extending the floodlight controller", Proceedings of the IEEE 14th International Symposium on Network Computing and Applications, Cambridge, MA, USA, 28-30 September, pp. 126-133.

[102] P. Berde et al., " ONOS: towards an open, distributed SDN OS", Proceedings of the third workshop on Hot topics in software defined networking, August2014, pp. 1-6.

[103] S. Hassas Yeganeh, Y. Ganjali, " Kandoo: a framework for efficient and scalable offloading of control applications", Proceedings of the first workshop on Hot topics in software defined networks, August 2012, p. 19.

[104] N. Noorani and S. A. H. Seno, "Routing in VANETs based on intersection using SDN and fog computing", Proceedings of the 8th International Conference on Computer and Knowledge Engineering, Mashhad, Iran, 25-26 October 2018, pp. 339-344.

[105] A. T. Albu-Salih, S. A. H. Seno, S. J. Mohammed, "Dynamic routing method over hybrid SDN for flying ad hoc networks", Baghdad Science Journal, Vol. 15, No. 3, 2018, pp. 361-368.

[106] Z. Jia, Y. Sun, Q. Liu, S. Dai, C. Liu, "CRetor: An SDN-Based Routing Scheme for Data Centers with Regular Topologies", IEEE Access, Vol. 8, 2020, pp. 116866-116880.

[107] ONF, "Open Networking Foundation", www.opennetworking.org/mininet/ (accessed: 2022)

[108] K. Fall, K. Varadhan, The network simulator (ns-2), www. isi. edu/nsnam/ns (accessed 2022)

[109] A. Alidadi, S. Arab, T. Askari, "A novel optimized routing algorithm for QoS traffic engineering in SDN-based mobile networks", ICT Express, Vol. 8, No. 1, 2022, pp. 130-134.

[110] P. Megyesi, A. Botta, G. Aceto, A. Pescapè, S. Molnár, "Available bandwidth measurement in software-defined networks", Proceedings of the 31st Annual ACM Symposium on Applied Computing, April 2016, pp. 651-657,

[111] D. L. C. Dutra, M. Bagaa, T. Taleb, K. Samdanis, "Ensuring End-to-End QoS Based on Multi-Paths Routing Using SDN Technology", Proceedings of the IEEE Global Communications Conference, Singapore, 4-8 December 2017, pp. 1-6, ,

[112] M. R. Celenlioglu, H. A. Mantar, "An SDN based intra-domain routing and resource management model", Proceedings of the IEEE International Conference on Cloud Engineering, Tempe, AZ, USA, 9-13 March 2015, pp. 347-352.

[113] C. Huang, C. Nakasan, K. Ichikawa, H. Iida, "A multipath controller for accelerating GridFTP transfer over SDN", Proceedings of the IEEE 11th International Conference on e-Science, Munich, Germany, 31 August - 04 September 2015, pp. 439-447.

[114] C. Huang, C. Nakasan, K. Ichikawa, H. Iida, "An SDN-Based Multipath GridFTP for High-Speed Data Transfer", Proceedings of the IEEE 36th International Conference on Distributed Computing Systems, Nara, Japan, 27-30 June 2016, pp. 763-764.

[115] S. Tariq, M. Bassiouni, "QAMO-SDN: QoS aware Multipath TCP for software-defined optical networks", Proceedings of the 12th Annual IEEE Consumer Communications and Networking Conference, Las Vegas, NV, USA, 9-12 January 2015, pp. 485-491.

[116] S. Tariq, M. Bassiouni, "Performance evaluation of MPTCP over optical burst switching in data centres", Proceedings of the International Telecommunications Symposium, Sao Paulo, Brazil, 17-20 August 2014, pp. 0-4.

[117] A. Basit, S. Qaisar, S. H. Rasool, M. Ali, "SDN Orchestration for Next Generation Inter-Networking: A Multipath Forwarding Approach", IEEE Access, Vol. 5, 2017, pp. 13077-13089.

[118] D. Eppstein, "Finding the k Shortest Paths", SIAM Journal on Computing, Vol. 28, No. 2, 1998, pp. 652-673.

[119] S. A. Hussain, S. Akbar, I. Raza, "A dynamic multipath scheduling protocol (DMSP) for full performance isolation of links in software-defined networking (SDN)", Proceedings of the 2nd Workshop on Recent Trends in Telecommunications Research, Palmerston North, New Zealand, 10 February 2017,

[120] N. Farrugia, V. Buttigieg, J. A. Briffa, "A globally optimised multipath routing algorithm using SDN", Proceedings of the 21st Conference on Innovation in Clouds, Internet and Networks and Workshops, Paris, France, 19-22 February 2018, pp. 1-8.

[121] S. A. Astaneh, S. Shah Heydari, "Optimization of SDN Flow Operations in Multi-Failure Restoration Scenarios", IEEE Transactions on Network and Service Management, Vol. 13, No. 3, 2016, pp. 421-432.

[122] S. Astaneh, S. S. Heydari, "Multi-failure restoration with minimal flow operations in software-defined networks", Proceedings of the 11th International Conference on the Design of Reliable Communication Networks, Kansas City, MO, USA, 24-27 March 2015, pp. 263-266.

[123] A. Malik, B. Aziz, M. Bader-El-Den, "Finding most reliable paths for software-defined networks", 13th International Wireless Communications and Mobile Computing Conference, June 2017, pp. 1309-1314,

[124] A. Malik, R. de Fréin, B. Aziz, "Rapid restoration techniques for software-defined networks", Applied Sciences, Vol. 10, No. 10, 2020.

[125] C. Yu, Z. Zhao, Y. Zhou, H. Zhang, "Intelligent Optimizing Scheme for Load Balancing in Software

Defined Networks", Proceedings of the IEEE 85th Vehicular Technology Conference, Sydney, NSW, Australia, 4-7 June 2017, pp. 5-9.

[126] C. J. C. H. Watkins and P. Dayan, "Technical Note: Q-Leaning", Reinforcement Learning, Vol. 292, Boston, Springer, 1992, pp. 55-68.

[127] R. Miikkulainen et al. "Evolving deep neural networks", Artificial Intelligence in the Age of Neural Networks and Brain Computing, 2019, pp. 293-312.

[128] K. Gotani et al. "Design of an SDN Control Method Considering the Path Switching Time under Disaster Situations", Proceedings of the 5th International Conference on Information and Communication Technologies for Disaster Management, December 2018, pp. 1-4.

[129] K. Gotani, H. Takahira, M. Hata, L. Guillen, S. Izumi, T. Abe, "OpenFlow Based Information Flow Control Considering Route Switching Cost", Proceedings of the IEEE 43rd Annual Computer Software and Applications Conference, July 2019, Vol. 2, pp. 527-530.

[130] B. Zhao, J. Zhao, X. Wang, T. Wolf, "RuleTailor: Optimizing Flow Table Updates in OpenFlow Switches with Rule Transformations", IEEE Transactions on Network and Service Management, Vol. 16, No. 4, 2019, pp. 1581-1594.

[131] C. Wang, H. Y. Youn, "Entry aggregation and early match using hidden Markov model of flow table in SDN", Sensors, Vol. 19, No. 10, 2019.

[132] W. K. Jia, X. Wang, "Flow aggregation for large-scale SDNs with scattered address space allocation", Journal of Network and Computer Applications, Vol. 169, 2020, p. 102787.

[133] Y. Guo, H. Luo, Z. Wang, X. Yin, J. Wu, "Routing optimization with path cardinality constraints in a hybrid SDN", Computer Communications, Vol. 165, 2021, pp. 112-121.

[134] R. Maaloul, R. Taktak, L. Chaari, B. Cousin, "Energy-Aware Routing in Carrier-Grade Ethernet Using SDN Approach", IEEE Transactions on Green Communications and Networking, Vol. 2, No. 3, 2018, pp. 844-858.

[135] S. Civanlar, M. Parlakisik, A. M. Tekalp, B. Gorkemli, B. Kaytaz, E. Onem, "A QoS-enabled OpenFlow environment for scalable video streaming", Proceedings of the IEEE Globecom Workshops, Miami, FL, USA, 6-10 December 2010, pp. 351-356.

[136] H. Owens and A. Durresi, "Video over Software-Defined Networking (VSDN)", Computer Networks, Vol. 92, 2015, pp. 341-356.

[137] T. F. Yu, K. Wang, Y. H. Hsu, "Adaptive routing for video streaming with QoS support over SDN networks", Proceedings of the International Conference on Information Networking, Cambodia, 12-14 January 2015, pp. 318-323.

[138] S. Yilmaz, A. M. Tekalp, B. D. Unluturk, "Video streaming over software-defined networks with server load balancing", Proceedings of the International Conference on Computing, Networking and Communications, Garden Grove, CA, USA, 16-19 February 2015, pp. 722-726.

[139] A. Chooprateep, Y. Somchit, "Video Path Selection for Traffic Engineering in SDN", Proceedings of the 11th International Conference on Information Technology and Electrical Engineering, Pattaya, Thailand, 10-11 October 2019.

[140] A. Jüttner, B. Szviatovszki, I. Mécs, Z. Rajkó, "Lagrange relaxation based method for the QoS routing problem", Proceedings of the Conference on Computer Communications. Twentieth Annual Joint Conference of the IEEE Computer and Communications Society, Anchorage, AK, USA, 22-26 April 2001, pp. 859-868.

[141] D. Kotani, Y. Okabe, "A Packet-In message filtering mechanism for protection of control plane in OpenFlow switches", Proceedings of the ACM/IEEE Symposium on Architectures for Networking and Communications Systems, Marina del Rey, CA, USA, 20-21 October 2014, pp. 695-707.

[142] J. Yan, H. Zhang, Q. Shuai, B. Liu, X. Guo, "HiQoS: An SDN-based multipath QoS solution", China Communications, Vol. 12, No. 5, 2015, pp. 123-133,

[143] S. Tomovic, N. Prasad, I. Radusinovic, "SDN control framework for QoS provisioning", Proceedings of the 22nd Telecommunications Forum Telfor, Belgrade, Serbia, 25-27 November 2014, pp. 111-114,

[144] B. G. Assefa, O. Ozkasap, "MER-SDN: Machine learning framework for traffic-aware energy-efficient routing in SDN", Proceedings of the IEEE 16th International Conference on Dependable, Auto-

nomic and Secure Computing, 16th International Conference on Pervasive Intelligence and Computing, 4th International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress, 2018, pp. 966-973.

[145] G. C. Deng, K. Wang, "An Application-aware QoS Routing Algorithm for SDN-based IoT Networking", Proceedings of the IEEE Symposium on Computers and Communications, Natal, Brazil, 25-28 June 2018, pp. 186-191.

[146] C. X. Cui, Y. B. Xu, "Research on load balance method in SDN", International Journal of Grid and Distributed Computing, Vol. 9, No. 1, 2016, pp. 25-36.

[147] O. Ahmed, F. Ren, A. Hawbani, Y. Al-Sharabi, "Energy Optimized Congestion Control-Based Temperature Aware Routing Algorithm for Software Defined Wireless Body Area Networks", IEEE Access, Vol. 8, 2020, pp. 41085-41099.

[148] L. El-Garoui, S. Pierre, S. Chamberland, "A new sdn-based routing protocol for improving delay in smart city environments", Smart Cities, Vol. 3, No. 3, 2020, pp. 1004-1021.

[149] Y.-S. Yu and C.-H. Ke, "Genetic algorithm-based routing method for enhanced video delivery over software-defined networks", International Journal of Communication Systems, Vol. 31, No. 1, 2018, p. 3391.

[150] S. Abbas, N. Javaid, A. Almogren, S. M. Gulfam, A. Ahmed, A. Radwan, "Securing Genetic Algorithm Enabled SDN Routing for Blockchain-Based Internet of Things", IEEE Access, Vol. 9, 2021, pp. 139739-139754,

[151] D. Li, X. Wang, Y. Jin, H. Liu, "Research on QoS routing method based on NSGAII in SDN", Journal of Physics: Conference Series, Vol. 1656, No. 1, 2020,

[152] A. Yazdinejad, R. M. Parizi, A. Dehghantanha, G. Srivastava, S. Mohan, A. M. Rababah, "Cost optimization of secure routing with untrusted devices in software-defined networking", Journal of Parallel and Distributed Computing, Vol. 143, 2020, pp. 36-46.

[153] E. F. Morales, J. H. Zaragoza, "An introduction to reinforcement learning", Decis. Theory Model. Appl. Artif. Intell. Concepts Solut., 2011, pp. 63-80.

[154] J. Jiang, L. Hu, P. Hao, R. Sun, J. Hu, H. Li, "Q-FDBA: Improving QoE fairness for video streaming", Multimedia Tools and Applications, Vol. 77, No. 9, 2018, pp. 10787-10806,

[155] Z. Yuan, P. Zhou, S. Wang, X. Zhang, "Research on Routing Optimization of SDN Network Using Reinforcement Learning Method", Proceedings of the 2nd International Conference on Safety Produce Informatization, Chongqing, China, 28-30 November 2019, pp. 442-445.

[156] C. Fang, C. Cheng, Z. Tang, C. Li, "Research on Routing Algorithm Based on Reinforcement Learning in SDN", Journal of Physics: Conference Series, Vol. 1284, No. 1, 2019,

[157] T. T. Huong, N. D. D. Khoa, N. X. Dung, N. H. Thanh, "A global multipath load-balanced routing algorithm based on Reinforcement Learning in SDN", Proceedings of the International Conference on Information and Communication Technology Convergence, Jeju, Korea, 16-18 October 2019, pp. 1336-1341.

[158] B. G. Assefa and O. Ozkasap, "HyMER: A Hybrid Machine Learning Framework for Energy Efficient Routing in SDN", Proceedings of the IEEE 16th International Conference on Dependable, Autonomic and Secure Computing, 16th International Conference on Pervasive Intelligence and Computing, 4th International Conference on Big Data Intelligence and Computing and Cyber Science and Technology Congress, August 2018, pp. 974-980.

[159] Z. Shi, J. Zhu, H. Wei, "SARSA-based delay-aware route selection for SDN-enabled wireless-PLC power distribution IoT", Alexandria Engineering Journal, Vol. 61, No. 8, 2022, pp. 5795-5803.

[160] A. Ouhab, T. Abreu, H. Slimani, A. Mellouk, "Energy-efficient clustering and routing algorithm for large-scale SDN-based IoT monitoring", Proceedings of the IEEE International Conference on Communications, Dublin, Ireland, 7-11 June 2020.

[161] C. Yu, J. Lan, Z. Guo, Y. Hu, "DROM: Optimizing the Routing in Software-Defined Networks with Deep Reinforcement Learning", IEEE Access, Vol. 6, 2018, pp. 64533-64539,

[162] W. X. Liu, "Intelligent Routing based on Deep Reinforcement Learning in Software-Defined Data-Center Networks", Proceedings of the IEEE Symposium on Computers and Communications, Barcelona, Spain, 29 June - 3 July 2019.

[163] W. Liu, J. Cai, Q. C. Chen, Y. Wang, "DRL-R: Deep reinforcement learning approach for intelligent routing in software-defined data-center networks", Journal of Network and Computer Applications, 2020, p. 102865.

[164] D. N. Maheswari, C. B. Sujitha, "Routing Optimization in Sdn Using Deep Reinforcement Learning", Journal of Engineering, Computing and Architecture, Vol. 10, No. 5, 2020, pp. 40-47.

[165] C. Xu, W. Zhuang, H. Zhang, "A Deep-reinforcement Learning Approach for SDN Routing Optimization", Proceedings of the 4th International Conference on Computer Science and Application Engineering, October 2020.

[166] Y. R. Chen, A. Rezapour, W. G. Tzeng, S. C. Tsai, "RL-Routing: An SDN Routing Algorithm Based on Deep Reinforcement Learning", IEEE Transactions on Network Science and Engineering, Vol. 7, No. 4, 2020, pp. 3185-3199.

[167] X. Guo, H. Lin, Z. Li, M. Peng, "Deep-Reinforcement-Learning-Based QoS-Aware Secure Routing for SDN-IoT", IEEE Internet of Things Journal, Vol. 7, No. 7, 2020, pp. 6242-6251.

[168] W. Qi, Q. Song, X. Kong, L. Guo, "A traffic-differentiated routing algorithm in Flying Ad Hoc Sensor Networks with SDN cluster controllers", Journal of the Franklin Institute, Vol. 356, No. 2, 2019, pp. 766-790.

[169] S. Torkzadeh, H. Soltanizadeh, A. A. Orouji, "Energy-aware routing considering load balancing for SDN: a minimum graph-based Ant Colony Optimization", Cluster Computing, Vol. 24, No. 3, 2021, pp. 2293-2312.

[170] M. Hayes, B. Ng, A. Pekar, W. K. G. Seah, "Scalable Architecture for SDN Traffic Classification", IEEE Systems Journal, Vol. 12, No. 4, 2018, pp. 3203-3214.

[171] A. M. Abdelmoniem, B. Bensaou, A. J. Abu, "Mitigating incast-TCP congestion in data centres with SDN", Annals of Telecommunications, Vol. 73, No. 3-4, 2018, pp. 263-277.

[172] A. Gutnikov, "DDoS attacks in Q4 2021", www.securelist.com/category/ddos-reports/ (accessed: 2022)

[173] B. Isyaku, K. Bin Abu Bakar, M. N. Yusuf, M. S. Mohd Zahid, "Software Defined Networking Failure Recovery with Flow Table Aware and flows classification", Proceedings of the 11th IEEE Symposium on Computer Applications & Industrial Electronics, April 2021, pp. 337-342.

INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

Published by Faculty of Electrical Engineering, Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia.

About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

Copyright

Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

TITLE OF ARTICLE (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

Author/Authorized Agent **Date**

CONTACT
International Journal of Electrical and Computer Engineering Systems (IJECES)
Faculty of Electrical Engineering, Computer Science and Information Technology Osijek
Josip Juraj Strossmayer University of Osijek
Kneza Trpimira 2b
31000 Osijek, Croatia
Phone: +38531224600,
Fax: +38531224605,
e-mail: ijeces@ferit.hr