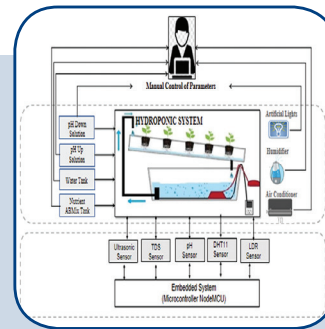
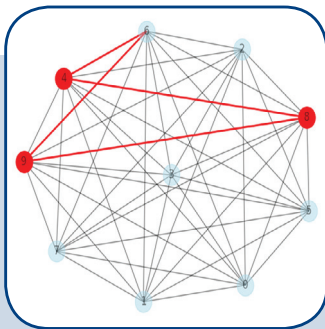
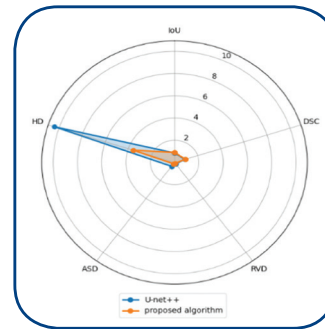
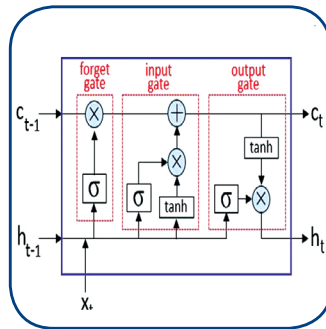
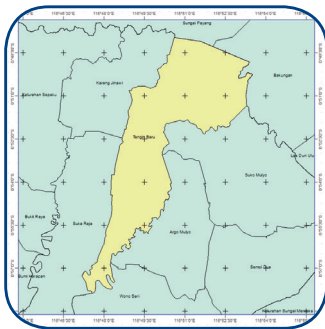


# International Journal of Electrical and Computer Engineering Systems



# 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 15, Number 8, 2024 | Pages 631 - 718

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

## EDITOR-IN-CHIEF

**Tomislav Matić**  
J.J. Strossmayer University of Osijek,  
Croatia

**Goran Martinović**  
J.J. Strossmayer University of Osijek,  
Croatia

## EXECUTIVE EDITOR

**Mario Vranješ**  
J.J. Strossmayer University of Osijek, Croatia

## ASSOCIATE EDITORS

**Krešimir Fekete**  
J.J. Strossmayer University of Osijek, Croatia

**Damir Filko**  
J.J. Strossmayer University of Osijek, Croatia

**Davor Vinko**  
J.J. Strossmayer University of Osijek, Croatia

## EDITORIAL BOARD

**Marinko Barukčić**  
J.J. Strossmayer University of Osijek, Croatia

**Tin Benšić**  
J.J. Strossmayer University of Osijek, Croatia

**Matjaz Colnarič**  
University of Maribor, Slovenia

**Aura Conci**  
Fluminense Federal University, Brazil

**Bojan Čukić**  
University of North Carolina at Charlotte, USA

**Radu Dobrin**  
Mälardalen University, Sweden

**Irena Galić**  
J.J. Strossmayer University of Osijek, Croatia

**Ratko Grbić**  
J.J. Strossmayer University of Osijek, Croatia

**Krešimir Grgić**  
J.J. Strossmayer University of Osijek, Croatia

**Marijan Herceg**  
J.J. Strossmayer University of Osijek, Croatia

**Darko Huljenić**  
Ericsson Nikola Tesla, Croatia

**Željko Hocenski**  
J.J. Strossmayer University of Osijek, Croatia

**Gordan Ježić**  
University of Zagreb, Croatia

**Ivan Kaštelan**  
University of Novi Sad, Serbia

**Ivan Maršić**  
Rutgers, The State University of New Jersey, USA

**Kruno Miličević**  
J.J. Strossmayer University of Osijek, Croatia

**Gaurav Morghare**  
Oriental Institute of Science and Technology,  
Bhopal, India

**Srete Nikolovski**  
J.J. Strossmayer University of Osijek, Croatia

**Davor Pavuna**  
Swiss Federal Institute of Technology Lausanne,  
Switzerland

**Marjan Popov**  
Delft University, Nizozemska

**Sasikumar Punnekkat**  
Mälardalen University, Sweden

**Chiara Ravasio**  
University of Bergamo, Italija

**Snježana Rimac-Drlje**  
J.J. Strossmayer University of Osijek, Croatia

**Krešimir Romić**  
J.J. Strossmayer University of Osijek, Croatia

**Gregor Rozinaj**  
Slovak University of Technology, Slovakia

**Imre Rudas**  
Budapest Tech, Hungary

**Dragan Samardžija**  
Nokia Bell Labs, USA

**Cristina Seceleanu**  
Mälardalen University, Sweden

**Wei Siang Hoh**  
Universiti Malaysia Pahang, Malaysia

**Marinko Stojkov**  
University of Slavonski Brod, Croatia

**Kannadhasan Suriyan**  
Cheran College of Engineering, India

**Zdenko Šimić**  
The Paul Scherrer Institute, Switzerland

**Nikola Teslić**  
University of Novi Sad, Serbia

**Jami Venkata Suman**  
GMR Institute of Technology, India

**Domen Verber**  
University of Maribor, Slovenia

**Denis Vranješ**  
J.J. Strossmayer University of Osijek, Croatia

**Bruno Zorić**  
J.J. Strossmayer University of Osijek, Croatia

**Drago Žagar**  
J.J. Strossmayer University of Osijek, Croatia

**Matej Žnidarec**  
J.J. Strossmayer University of Osijek, Croatia

## Proofreader

**Ivanka Ferčec**  
J.J. Strossmayer University of Osijek, Croatia

## Editing and technical assistance

**Davor Vrandečić**  
J.J. Strossmayer University of Osijek, Croatia

**Stephen Ward**  
J.J. Strossmayer University of Osijek, Croatia

**Dražen Bajer**  
J.J. Strossmayer University of Osijek, Croatia

## 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>

## Copyright

Authors of the International Journal of Electrical  
and Computer Engineering Systems must transfer  
copyright to the publisher in written form.

# TABLE OF CONTENTS

<b>Empowering Mental Health: CNN and LSTM Fusion for Timely Depression Detection in Women</b> .....	631
<i>Original Scientific Paper</i>	
Divya Pakkattil   Ravindran Sri Devi	
<b>Sentiment Mining in E-Commerce: The Transformer-based Deep Learning Model</b> .....	641
<i>Original Scientific Paper</i>	
Tahani Alsaedi   Asif Nawaz   Abdulrahman Alahmadi   Muhammad Rizwan Rashid Rana   Ammar Raza	
<b>End-to-End Fully Automated Lung Cancer Volume Estimation System</b> .....	651
<i>Original Scientific Paper</i>	
Pushkar Sathe   Alka Mahajan   Deepak Patkar   Mitusha Verma	
<b>Automated Surgical Wound Classification for Intelligent Emergency Care Applications</b> .....	663
<i>Original Scientific Paper</i>	
Min Woo Park   Mee Young Sung	
<b>A Framework for 5G Network Slicing Optimization using 2-Edge-Connected Subgraphs for Path Protection</b> .....	675
<i>Original Scientific Paper</i>	
Igor Begić   Adrian Satja Kurdija   Željko Ilić	
<b>Automated vs. Semi-Automated Hydroponics: Quantifying Automation Effects on Plant Growth</b> .....	687
<i>Original Scientific Paper</i>	
Swati Jain   Mandeep Kaur	
<b>Improved Parameter Estimation of Three-Phase Squirrel-Cage Induction Motors Using the Nelder-Mead Simplex Algorithm</b> .....	695
<i>Original Scientific Paper</i>	
Son T. Nguyen   Linh V. Trieu   Tu M. Pham   Anh Hoang	
<b>CRITIC-TOPSIS Method: Design of hybrid renewable energy systems based on multi-criteria decision-making</b> .....	705
<i>Original Scientific Paper</i>	
Ade Gafar Abdullah   Yusril Nurhikam   Dadang Lukman Hakim   Nanin Trianawati Sugito   Diky Zakaria	
<b>About this Journal</b>	
<b>IJECES Copyright Transfer Form</b>	





# Empowering Mental Health: CNN and LSTM Fusion for Timely Depression Detection in Women

Original Scientific Paper

**Divya Pakkattil\***

Department of Computer Science  
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India  
muralidivya.132@gmail.com

**Ravindran Sri Devi**

Department of Computer Science  
Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India  
devisri.0878@gmail.com

\*Corresponding author

**Abstract** – Depression is a mental illness that manifests as persistent melancholy, a loss of interest in routine activities, trouble focusing, poor memory, and a lack of energy. It is a widespread mental health condition that can affect people of any age or gender. Depression is more common in women than in men. In order to identify early indicators of depression in women, this study uses a deep learning-based model utilizing convolutional neural networks and Long Short-Term Memory. With the help of a dataset of left and right hemispheric electroencephalogram data, the suggested model was trained and assessed. The suggested method entails preprocessing the electroencephalogram data, which is feature-extracted using a convolutional neural network, and sequence modeling using a Long Short-Term Memory network. With the help of Electroencephalogram data from women with and without depression, the model was trained and assessed. The results show that the suggested method successfully identified depression in women using Electroencephalogram data with excellent accuracy, sensitivity, and specificity. When it came to identifying female depression, the model had an accuracy of 99.02% on the left hemisphere and a right hemisphere accuracy of 98.06%. The study demonstrates that employing advanced deep learning techniques on electroencephalogram data enables accurate and sensitive identification of depression in women. This highlights the potential for early intervention in mental health disorders, particularly in populations with a higher depression prevalence.

---

**Keywords:** Early depression detection, EEG signals, mild depression, Long Short-Term Memory, Feature selection, Convolutional Neural Networks

---

Received: January 10, 2024; Received in revised form: May 10, 2024; Accepted: June 3, 2024

## 1. INTRODUCTION

Psychoanalytic ideas of depression [1] first appeared around the beginning of the 20th century, emphasizing the importance of unconscious conflicts and early experiences. Subsequently, with studies into the role of neurotransmitters [2], genetics, and brain architecture in depression, biological explanations developed. Depression is now understood to be a complicated condition with a variety of causes and treatments. Depression can be treated with psychotherapy [3], medication [4], and lifestyle modifications (such as regular exercise and a nutritious diet). The significance of tackling societal factors that affect mental health, such as poverty,

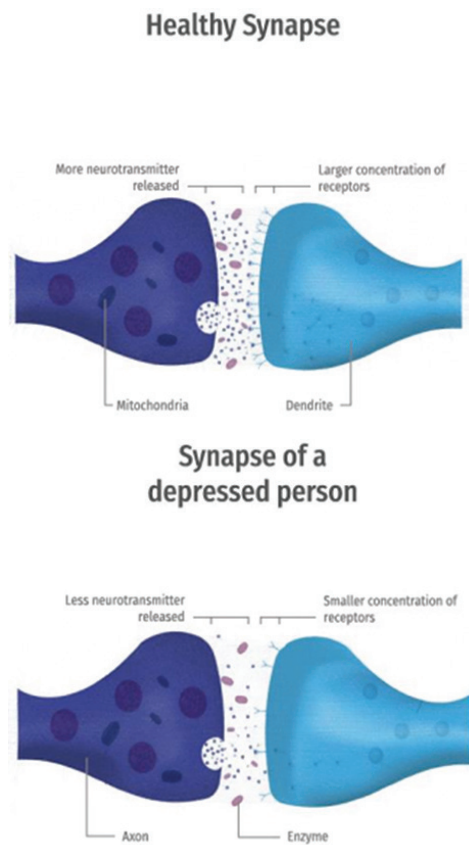
prejudice, and trauma, is now becoming more widely acknowledged [5]. A mental health condition called depression, commonly referred to as major depressive disorder, affects a person's emotions, thoughts, and behavior. Although it can strike at any age, this prevalent ailment typically manifests in early adulthood. Depression has historically been associated with moral failure or weakness, and those who experience it are frequently stigmatized and shunned.

### 1.1. CAUSES OF DEPRESSION

As depression is a complicated disorder, its root causes are still unclear. Depression is probably brought

on by a confluence of hereditary, environmental, and psychological factors. The following are some typical reasons for depression:

- **Biological factors:** Particularly, biological components may be quite important in the emergence of depression. Depression may be exacerbated by imbalances in brain neurotransmitters such as serotonin [6] and dopamine [7]. Depression tends to run in families, so genetics may also be involved. Depression may also be exacerbated by abnormalities in neurotransmitters such as serotonin, norepinephrine [8], and dopamine. These chemicals control mood, and when they are out of balance, it can result in depressive symptoms. Hormonal imbalances, persistent pain, anomalies in the structure of the brain, and chronic sickness [9] are some additional biological variables that might contribute to depression. It is crucial to remember that while biological variables can play a part in the onset of depression, they are not the only contributing element.
- **Environmental factors:** Depression development may also be influenced by environmental variables. Depression risk factors include things like traumatic experiences [10], ongoing stress [11], and social isolation. Depression can be brought on in those who are vulnerable by traumatic events including losing a loved one, experiencing physical or mental abuse, or going through a significant life transition. Persistent stress, like that brought on by work-related stress or financial anxiety, can also help depression develop. Feelings of hopelessness and despair can result from social isolation [12], which can also cause depression. Loneliness and a lack of social support are two examples of social isolation. Furthermore, being exposed to chemicals or pollutants in the environment might harm mental health and hasten the onset of depression [13]. It is crucial to understand how environmental variables may contribute to the onset of depression and to take action to address and manage these factors in order to avoid or treat the illness. This may entail reaching out to friends and family for support, limiting exposure to stressors, and engaging in self-care practices like meditation or exercise.
- **Psychological factors:** Depression might also occur because of psychological reasons. For instance, unfavorable cognitive habits, including ruminating, criticizing, and having poor self-esteem, might help depression develop and persist [14]. A feeling of hopelessness and helplessness, which are frequent signs of depression, can result from negative self-talk and persistently dwelling on unfavorable occurrences. Moreover, some psychological qualities, including neuroticism [15], may make someone more susceptible to depression. Negative emotions are more powerful and frequent in neurotic people than in non-neurotic people, which might hasten the onset of depression.
- **In addition, negative life events like abuse, neglect, or childhood trauma can raise one's risk of developing depression later in life.** These events can increase a person's sense of worthlessness and powerlessness, which increases their risk of developing depression [16]. Depressive disorders can also develop as a result of a lack of social support and harmful coping strategies like substance abuse and avoidance. Developing effective treatment strategies, such as cognitive-behavioral therapy (CBT) [17] and mindfulness-based therapies that target negative thought patterns and assist people in creating healthy coping mechanisms to manage depression symptoms, requires an understanding of the psychological factors that contribute to depression.
- **Hormonal factors:** Depression can also be influenced by hormonal changes, such as those that take place during pregnancy, menopause, or the menstrual cycle [15]. Moreover, hormones might contribute to depression, especially in women. Depression can be exacerbated by changes in hormone levels brought on by the menstrual cycle, pregnancy, and menopause. Similarly, fluctuations in hormone levels during pregnancy might bring on feelings of depression. Many new moms experience postpartum depression, a common mood illness that is also thought to be brought on by hormonal changes following childbirth. Finally, some women may have symptoms of depression during menopause [18] due to the drop in estrogen levels. Depression may also be exacerbated by hormonal imbalances brought on by illnesses like hypothyroidism or hyperthyroidism. When there is an imbalance in the thyroid hormones, which are essential for controlling the body's metabolism and energy levels, depression symptoms can result. The symptoms of sadness can also be exacerbated by hormonal imbalances brought on by specific drugs or medical procedures, such as hormonal birth control or chemotherapy [19]. In order to create efficient treatment plans that address the underlying hormonal imbalances, such as hormone replacement therapy or drug management, it is crucial to understand the hormonal aspects that contribute to depression.
- **Substance abuse:** Addiction to drugs or alcohol [20] is another potential factor in depression. This is because of a variety of things, such as the harm that substance abuse does to the brain and the emotional toll that addiction takes. Abuse of substances can cause chemical imbalances in the brain, disturb the balance of neurotransmitters, and result in depressive symptoms. Synapses from a healthy person and a depressed person are shown in Fig. 1. A sense of hopelessness and despair can also be produced by the cycle of addiction and substance misuse, which can aid in the onset of depression.



**Fig. 1.** Synapse of Healthy person and Depressed Person

## 1.2. TYPES OF DEPRESSION

There are several different types of depression, each with its own specific symptoms and causes. Here are some of the most common types:

- Major depressive disorder (MDD): MDD, commonly referred to as clinical depression, is defined by a person's loss of interest in formerly liked activities as well as persistent emotions of melancholy and hopelessness. Treatment for MDD is frequently necessary because it can make daily tasks difficult.
- Persistent depressive disorder (PDD): Chronic depression, or PDD, lasts for at least two years. Although the symptoms of depression in people with PDD are less severe than in people with MDD, they can nevertheless be disruptive to everyday living.
- Seasonal affective disorder (SAD): SAD is a form of depression that typically develops in the winter and is brought on by seasonal changes. Low mood, lack of energy, and increased tiredness are SAD symptoms.
- Postpartum depression (PPD): Postpartum depression (PPD) is a kind of depression that affects new moms. PPD symptoms include depression, anxiety, and a hard time bonding with the infant.
- Bipolar disorder: Extreme mood swings, such as those that can range from episodes of depression

to episodes of mania or hypomania, are the hallmark of bipolar disorder, a mood illness.

- Psychotic depression: A severe form of depression known as psychotic depression includes psychotic symptoms like hallucinations or delusions.
- Situational depression: Situational depression is a form of depression that is brought on by a particular event, like the death of a loved one or the loss of a job.

The paper begins with an introduction discussing the historical context and multifactorial nature of depression, emphasizing the importance of timely detection, particularly in women. It delves into various causes and types of depression, highlighting biological, environmental, psychological, hormonal, and substance-related factors. Following the introduction, a comprehensive literature review is presented, summarizing recent research on depression detection using machine learning and deep learning techniques. The methodology section describes the proposed approach for depression detection using EEG signals, detailing the CNN-LSTM model architecture and dataset used. Results and discussions analyze the performance of the model in detecting depression from EEG signals, including accuracy, sensitivity, and comparison with previous studies. The conclusion summarizes the findings, highlighting the efficacy of the proposed method for early detection of depression in women and discussing potential future research directions, such as utilizing larger and more diverse datasets and exploring multimodal data for improved diagnosis.

## 2. LITERATURE REVIEW

Burdisso et al. [21] introduce SS3, a brand-new supervised learning model for text classification, to serve as a generic foundation for addressing issues with early depression risk detection. The SS3 is intended to address three of the most difficult aspects of ERD: explainability, support for early classification, and incremental classification of sequential data. The work has limitations because the authors used words as the fundamental building blocks rather than higher-level building elements like sentences and paragraphs.

Su et al. [22] utilized long short-term memory (LSTM) and six machine learning (ML) models to forecast several depression risk indicators as well as the likelihood of depression in the older population over the following two years. The prediction accuracy of the model was assessed using decision curve analysis (DCA) and receiver operating curves (ROC). For early diagnosis and intervention, the suggested LSTM+ML model-based decision support system may be particularly beneficial for physicians, nurses, and community medical professionals. The method's need to progressively increase the retrospective waves employed in the LSTM model is one of its limitations. The high-dimensional and time-series data on risk factors for depression in the elderly can be successfully captured by the LSTM+ML model.

Zheng et al. [23] suggest a graph attention model for depression detection that is embedded with multi-modal information. In addition to teaching acceptable embeddings for knowledge graph nodes, this method makes use of medical information to enhance classification and prediction performance via the knowledge attention mechanism. The suggested approach greatly outperforms other leading state-of-the-art approaches in terms of classification and prediction performance, with guaranteed robustness with each modality of multi-modal data, according to experimental results on real-world datasets. Their approach may obtain the knowledge-attention representation vector using learned medical embeddings and a deep learning-based network, which primarily aids the detection model's performance. The findings may not be applicable to other populations or nations because the study only examined a particular population (elderly individuals in China).

Islam et al. [24] suggested employing a convolutional neural network (CNN) for a computer-aided detection (CAD) system (ConvNet). The authors employed transfer learning to build the architecture of ConvNet since the CAD system used in clinical practice and the architecture of ConvNet that was developed through trial and error should both be constructed on the foundation of local databases. The technique is used to diagnose depression without the use of harmful procedures or drugs. Only a few electroencephalogram (EEG) channels were employed in the investigation, which may not have been enough to gather sufficient information for a thorough analysis.

Facebook can be used as a tool for assessing and identifying serious depression among its users, according to Shah et al. [25], who have demonstrated this potential. The purpose of this study is to analyze depression using Facebook data that was gathered from an online open source. The authors suggest using scalable and effective machine learning technology to study the impact of depression identification. For the dataset, the authors used Linguistic Inquiry and Word Count (LIWC). The study's use of non-invasive techniques to gather information from social media platforms lessened the strain and expense of conventional data collection techniques. Due to the possibility that user-sensitive information may be present in social media data, the study raises ethical questions surrounding privacy and confidentiality.

Chiong et al. [26] have suggested a model that may identify depression from user posts. The training data were used to train deep learning algorithms, while the test data were used to evaluate their performance. With the use of this deep learning technique, authors have attempted to overcome the constraints of conventional machine learning, which only allows for text classification. The study showed the potential for early depression detection using deep learning techniques, which could result in early intervention and treatment,

thereby lessening the severity of the condition and improving outcomes. The study showed great accuracy in the diagnosis of depression, indicating that the deep learning methodology may be a potential way for social network data depression identification. Users could not have mentioned all symptoms of depression or may not have used some keywords associated with depression; therefore, the study may not have gathered all pertinent information about depression.

In order to suggest a generalized strategy for depression diagnosis using social media texts, Katchapakin et al. [27] study several text preprocessing and textual-based feature methods coupled with machine learning classifiers, including single and ensemble models. The authors first train and test the machine learning models using two publicly available, labeled Twitter datasets, and then they assess the performance of the learned models with three non-Twitter depression-class-only datasets. According to experimental findings, the suggested method may successfully identify depression in social media texts even when the training datasets don't include specific keywords (like "depression" and "diagnosis") and when unrelated datasets are used for testing. In contrast to other feature sets, the study's text-based methodology can collect more in-depth and complex information regarding depression. The study might have only used a few feature extraction techniques, which could have reduced the approach's accuracy.

In order to create a depression detection algorithm for the Thai language on Facebook, where people utilize it as a platform for exchanging ideas, emotions, and life events, Alghowinem et al. [28] used natural language processing (NLP) approaches. Findings from 35 Facebook users showed that the degree of depression may be predicted by their behavior on Facebook. The findings of the trial indicate that depression might be predicted using Facebook behavioral data, including messages and actions. However, because Facebook has restricted their ability to acquire personal information and the procedure for doing so has grown more challenging, the sample size of this paper is quite small. As a result, not all significant aspects may be included in the study's findings. Also, as the language-related elements had to be translated from Thai to English in order to analyze the process, there may have been some mistakes as a result of the translation process because some crucial sentimental polar words may have been lost.

Arora et al. [29] analyze the frequently used features and their impact on the classification outcomes in order to provide an interpretation of the depression detection model. The created framework gathers and chooses the 38 feature selection algorithms of various categories' most promising characteristics for modeling depression detection. In order to create more accurate and effective models, the paper employs feature selection techniques that can help identify the most crucial elements for depression identification.



The framework's aggregated features are generalized by the authors using various datasets. Despite the fact that the paper's main focus is interpretability, the feature selection techniques may not fully explain how the models produce predictions.

Lam et al. [30] offer a method of separating depression- and anxiety-related health tweets from all other mixed tweets, which puts us one step closer to determining the health state of a living environment. It's a new platform for patient contact, with many of them taking part in decision-making for better treatment outcomes. Many crucial actions are carried out after the retrieval of the health tweets.

Genevieve et al. [31] focus on machine learning ways to automatically identify depression from clinical interviews, utilizing the models that have been trained on multimodal data. The authors propose a method that incorporates a data augmentation procedure based on topic modeling using a transformer and deep CNN for acoustic feature modeling.

Chiu et al. [32] provide a method for automatically detecting sadness that involves examining a person's environment at home, their behavior, and the social media posts they make. Recurrent neural networks are used in the proposed method to compute each person's post-representation. Next, using deep neural networks, the representations are integrated with additional content-based, behavioral, and living environment variables to forecast the individual's diagnosis of depression.

The existing literature primarily focuses on utilizing either CNN or LSTM networks for depression detection from EEG signals, but there is a noticeable gap in research exploring the fusion of these two deep learning architectures. This research gap prompted my interest in investigating the potential synergistic effects of combining CNNs and LSTMs to improve the accuracy and reliability of depression detection from EEG data. The man's contributions to the proposed work are as follows:

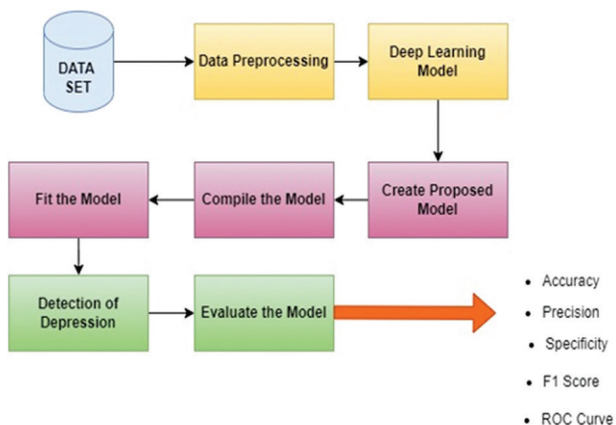
- Develop a novel approach utilizing CNN and LSTM fusion with EEG data for the timely detection of depression in women.
- Evaluate the performance of the proposed CNN-LSTM model in accurately identifying depression from EEG signals, aiming to provide a non-invasive and effective method for early detection of depression in female individuals.

### 3. METHODOLOGY

EEG signals are utilized in this study for automatic depression identification in women. CNNs are utilized for feature extraction from the EEG data, capturing spatial patterns and local dependencies within the signals. These extracted features are then fed into LSTM networks, which specialize in learning temporal dependencies, enabling the model to capture long-range de-

pendencies in the EEG data over time. Additionally, the study utilizes Keras, a deep learning library in Python, to implement the CNN-LSTM architecture, ensuring the computational efficiency and scalability of the model.

The feature maps and output of CNN are supplied to the LSTM as inputs, and these signals are subjected to sequence learning. The LSTM outputs are transmitted through fully connected layers. Fig. 2 illustrates the study's methodology in a block diagram format.



**Fig. 2.** The block diagram of the proposed architecture

While learning sequential information is difficult for the CNN model, it is good at extracting temporal aspects. Sequential information is generally used in the context of CNNs to describe data that has a temporal or sequential structure, such as speech signals, movies, or sensor data. Convolutional filters are applied to the sequence by the CNN, which searches for patterns at various scales. The patterns or traits that a CNN has trained to recognize and extract from the time dimension of incoming data are referred to as temporal features. CNNs often identify temporal features in sequential or time-series data, such as speech signals, movies, or sensor data. The CNN scans the input data for patterns at various time scales by applying convolutional filters across the temporal dimension. The local features of the input EEG signals are learned by CNN (representation learning), the long-term dependencies are learned by LSTM, and these features are processed sequentially by LSTM (sequence learning).

The CNN network is made up of numerous layers of connected neurons, each of which is in charge of identifying particular aspects of the input image. Edge detection is normally carried out in the first layer, with more intricate feature detection carried out in later layers. Pooling layers, which minimize the spatial size of the feature maps to increase the network's computational efficiency, are also used by CNNs. A fully connected layer serves as the network's final layer and maps the output of the convolutional layers to the class labels of the input images. The mathematical action that a CNN takes when processing an input picture or signal is represented by this equation.

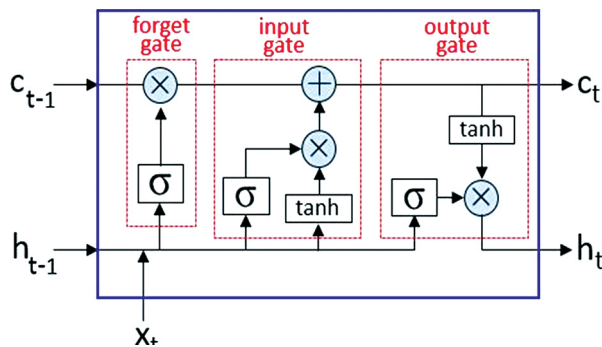
$$Y_c(t)=(X_c*W_c)(t)=\int X_c(\tau)W_c(t-\tau)D_\tau \quad (1)$$

Where impulse response is  $W_c$ , input signal is  $X_c$  and  $Y_c(t)$  denotes the output signal at time  $t$ . The output signal  $y$  is created by convolving the input signal  $X_c$  with a filter or kernel  $W_c$ . By integrating the input signal  $\int X_c(\tau)$  and the filter  $W_c(t-\tau)$  throughout the input signal domain, one may determine the outcome of the convolution operation at each point in time  $t$ .

$$F_c(k)=\max(0,k) \quad (2)$$

Where

$F_c(k)$  is a function that returns the highest value between  $b$  and  $a$ , and  $k$  is the input to the ReLU function.  $Y_c(t)$  is the output of the ReLU function, given an input  $x$ . If the input value is positive, this method returns  $k$ ; if it is negative, it returns 0. Because it is computationally effective and reduces the vanishing gradient issue in deep neural networks, the ReLU function is preferred over other activation functions like sigmoid. The behavior of some types of neurons in the brain that only fire when the input signal exceeds a particular threshold is modeled by the ReLU function, which is also biologically inspired. The new inputs for the following layer will be the extracted feature maps. The MaxPooling1D layer employs the maximum pooling strategy to minimize the size of the input. To prevent the suggested network from having an overfitting issue, we additionally applied the dropout approach. Because it enables the model to learn from data sequences like time series, sentences of natural language, and DNA sequences, sequence learning is a crucial step in deep learning. Transformers in DL models with LSTM networks are made to identify patterns and dependencies in data sequences. With the aid of sequence learning, the model may take the features from each element in the sequence and use them to forecast the next element or categorize the entire sequence. Deep learning models may accurately predict or classify data even when it is complicated and noisy by including temporal information. The vanishing gradient problem, which makes it challenging to train the prior layers and learn the network, is the main issue encountered while training artificial neural networks with back propagation.



**Fig. 3.** Architecture of LSTM

Unlike the RNN architecture, the LSTM architecture shown in Fig. 3 has unique memory cells for storing the previous input for a considerable amount of time.

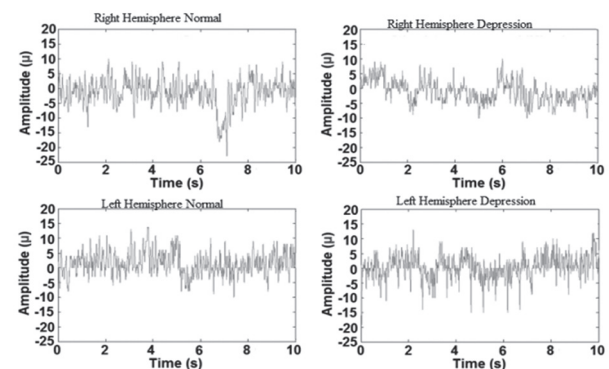
### 3.1. DESCRIPTION OF DATASET

We used the identical normal and depressive EEG signals in this analysis as those used by Acharya et al. [34]. The left and right halves of the brains of 15 healthy patients and 15 depressed subjects were used to collect the EEG data. A 50 Hz notch filter was used to remove power line noise, while at 256 Hz, the EEG signals were captured. Experts painstakingly removed muscle and eye movement-related artifacts from the EEG readings. The EEG data from 30 patients were divided into two groups: normal and depressed classes for the left and right hemispheres, respectively. It consists of 4318 normal EEG records and 4798 depression records.

**Table 1.** Details of the EEG Dataset

Class	EEG signal of right hemisphere	EEG signal of left hemisphere	Total
Depressed	2398	2400	4798
Normal	2159	2159	4318
Total	4557	4559	9116

In this study, automatic depression identification is carried out using EEG signal samples of the brain. The data are randomly split into training and testing groups when using the random splitting technique. As shown in Fig. 5, the dataset for each group in this study is split into 80% training sets and 20% test sets. The deep learning model is trained using the training datasets. The data that the model never used during training makes up the test set. As a result, test data can be used to more effectively observe the trained model's test performance. 4557 data files were collected from the right hemisphere. We have utilized 3645 of these files for training and 912 for testing. Similar to this, there are 4559 files overall in the left hemisphere's data. 911 files were used for testing after 3648 were used for training. Python is used to implement the model, together with the Keras deep learning tools. Fig. 4 displays examples of a few signal samples found in the data.



**Fig. 4.** Sample EEG signal

Table 2 list the parameters and the layers used in this approach. Using the values obtained using the brute-force method, the model's filter number and parameter selection, such as kernel size is carried out.

Table 2 list the parameters and the layers used in this approach. Using the values obtained using the brute-force method, the model's filter number and parameter selection, such as kernel size is carried out.

**Table 2.** Details of the proposed CNN-LSTM model

Layers	Type	Output shape
0	Inputs	8 x 512
1-6	Time-distributed convolutional block	8 x 256 x 32
7-12	Time-distributed convolutional block	8 x 128 x 32
13-18	Time-distributed convolutional block	8 X 64 X 32
19-24	Time-distributed convolutional block	8 x 32 x 32
25	Time-d distributed flattened	8 x 1024
26	Time-distributed fully connected	8 x 32
27	Time-distributed batch normalization	8 x 32
28	Time-distributed dropout 50%	8 x 32
29	LSTM	32
30	Batch normalization	32
31	Dropout 50%	32
32	Fully connected	16
33	Batch normalization	16
34	Dropout 50%	16
35	Outputs	4

#### 4. RESULT AND DISCUSSION

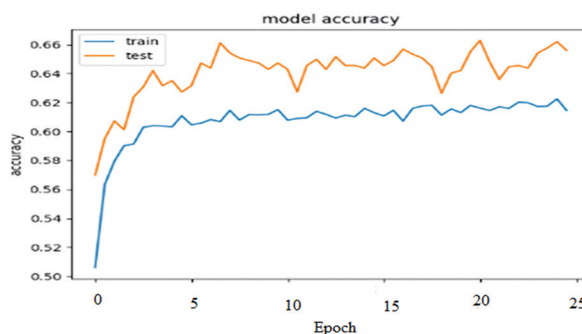
For right- and left-hemisphere EEG data, independent experimental tests were carried out. The normal and depression classes were included in the training set of the CNN-LSTM model. We ensure that overfitting does not take place during training by monitoring the model's performance. At the conclusion of training, the right hemisphere EEG signal model had achieved an accuracy of 98.06%, and the loss value had dropped from 0.42 to 0.03. Also, each epoch of training took an average of 52 seconds. By the end of the 25<sup>th</sup> epoch, the left hemisphere EEG signal model's training performances had attained an accuracy of 99.02% during training with a loss of 0.04.

**Table 3.** Obtained performance values

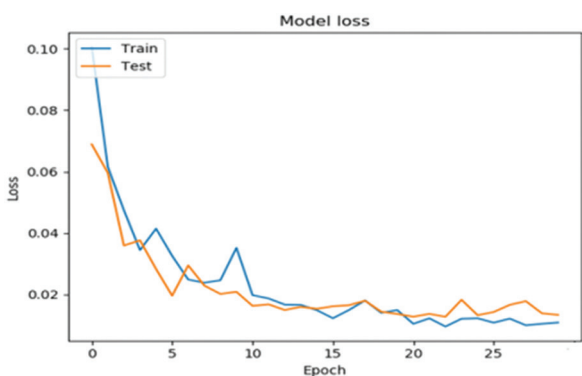
	Accuracy	Precision	Specificity	F1-Score
Using EEG of Right Hemisphere	98.06	98.01	98.45	99.01
Using EEG of Left Hemisphere	99.02	98.23	98.49	99.41

Only six input data points for the right hemisphere EEG signals were incorrectly classified by the CNN-LSTM model, which achieves 98.06% accuracy. The sensitivity of the normal class is 99.70%, whereas the sensitivity

of the depressed class is 98.55%. For left-hemisphere EEG signals, the model's accuracy is 97.66%. The classes for depression and normal have sensitivity values of 96.02% and 97.19%, respectively. It is clear from the results of the preceding analysis that the CNN-LSTM model performed best at identifying depression in EEG signals. Also, according to the results of our investigation, left hemisphere (Fp2-T4) EEG signals perform better than right EEG signals. The trained CNN-LSTM model classified a single EEG input in about 0.005 seconds. The experimental results demonstrate the efficacy of the proposed CNN-LSTM model in detecting depression from EEG signals. Specifically, the model achieved high accuracy rates of 98.06% and 99.02% for right and left hemisphere EEG signals, respectively. This indicates the model's robust performance in accurately classifying depression status based on EEG data. Additionally, the sensitivity values for both depression and normal classes further validate the model's effectiveness in correctly identifying individuals with depression while minimizing false positives. The performance graph of the proposed model is shown in the below figure.

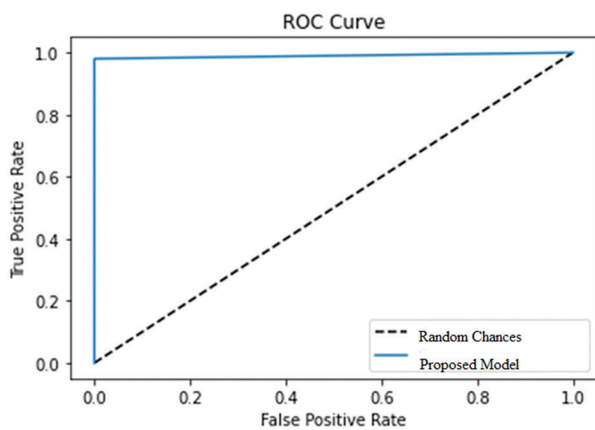


**Fig. 5.** Performance measure Accuracy plot



**Fig. 6.** Performance measure Loss plot

Also, ROC curves were used to evaluate the model. ROC curves are widely used in healthcare decision-making and are regarded as a powerful measure of a diagnostic test. They are extremely helpful for developing classifiers and visualizing their performance. On the X-axis of the ROC curve are graphs showing the false positive rate, and on the Y-axis are plots showing the real positive rate. The ROC for the depression detection algorithm is shown in Fig. 7.



**Fig. 7.** ROC of the proposed system

It is evident that the new method has higher accuracy than the earlier investigations. The accuracy of Acharya et al. [33] for right hemisphere EEG signals was 95.49%, whereas our accuracy using the same database was 99.02 and 98.06%. Although the LSTM network learns sequences from these attributes, the CNN network learns local properties. A few cutting-edge investigations have used feature extraction approaches. These methods demand expertise, are difficult, and take a lot of time. Yet, in our investigation, the model itself does the feature extraction. In addition, a deep detection model that uses layers rather than simple classifiers like SVM and ANN is proposed. As a result, the CNN LSTM model is reliable and accurate at identifying depression from EEG signals. The suggested model quickly and precisely detects unidentified EEG signals. As a result, the suggested model is practical and simple to apply for clinical applications. The biggest flaw in this study is the small number of participants—15 healthy individuals and 15 depressed individuals. The developed model also requires a lot of computing. A future study will expand on this research by adding more people from various backgrounds and utilizing graphics processing units (GPUs) to simplify the computation. Further improvements will be made using more sophisticated deep learning techniques and more datasets. Moreover, classifiers like SVM will be used to evaluate the CNN-LSTM layers.

**Table 4.** Comparison of the proposed method with other methods

Reference	Methodology	Accuracy
[34]	CNN	87.95%
[35]	CNN	92.66%
[36]	CNN+LSTM	98.54%
[37]	DepAudioNet	66.07%
[38]	Time series classification	91.67%
Proposed	CNN+LSTM	99.02%

The results corroborate findings from previous studies, indicating the viability of utilizing deep learning models such as CNN and LSTM for depression detection from EEG signals. The high accuracy and sensitivity achieved by the proposed CNN-LSTM fusion model

align with the working hypothesis, affirming its potential as an effective tool for the early identification of depression in women. These results suggest that integrating CNN and LSTM networks can enhance the discriminative power of EEG-based depression detection models, thus contributing to advancements in mental health screening methodologies.

## 5. CONCLUSION

In summary, the work suggests a unique method for detecting depression in women early on by combining CNN and LSTM with EEG data. The suggested model showed excellent accuracy in identifying female depression using EEG data. Deep learning models have a number of benefits over conventional diagnostic techniques for using EEG data to identify depression. It offers unbiased assessments of brain activity, enabling more accurate mental health disorder diagnosis and monitoring. In order to provide remote diagnosis and monitoring of mental health disorders, the proposed model can be connected with telehealth systems. The comparatively tiny dataset utilized for model training and testing is one of the study's drawbacks. In order to enhance the generalizability of the model, future research should take into account utilizing larger and more varied datasets. More research into the use of multimodal data for a more precise diagnosis of mental health disorders such as speech or facial expressions would be intriguing. These findings suggest that the CNN-LSTM approach offers a promising and reliable method for the early detection of depression in women. Thus, the proposed method can be deemed acceptable for clinical application, providing a valuable tool for mental health screening and intervention. Future research could explore integrating additional modalities, such as speech or facial expressions, with EEG data to develop a more comprehensive diagnostic tool for mental health disorders.

## 6. REFERENCES

- [1] O. F. Kernberg, "Challenges for the future of psychoanalysis", *The American Journal of Psychoanalysis*, Vol. 81, 2021, pp. 281-300.
- [2] D. Talevi et al. "Mental health outcomes of the CoViD-19 pandemic", *Rivista di Psichiatria*, Vol. 55, No. 3, 2021, pp. 137-144.
- [3] R. A. Champagne, "Sérotonine by Michel Houellebecq", *The French Review*, Vol. 93, No. 3, 2020, pp. 203-204.
- [4] L. Finkelstein-Fox, C. L. Park, "Control-coping goodness-of-fit and chronic illness: a systematic review of the literature", *Health Psychology Review*, Vol. 13, No. 2, 2020, pp. 137-162.



- [5] S. Groeneweg, F. S. van Geest, R. P. Peeters, H. Heuer, W. E. Visser, "Thyroid hormone transporters", *Endocrine Reviews*, Vol. 41, No. 2, 2020, pp. 146-201.
- [6] K. Itriyeva, "Premenstrual syndrome and premenstrual dysphoric disorder in adolescents", *Current Problems in Pediatric and Adolescent Health Care*, 2022, p. 101187.
- [7] E. Akyuz, A. K. Polat, E. Eroglu, I. Kullu, E. Angelopoulou, Y. N. Paudel, "Revisiting the role of neurotransmitters in epilepsy: An updated review", *Life Sciences*, Vol. 265, 2021, p. 118826.
- [8] W. Tschacher, D. Meier, "Physiological synchrony in psychotherapy sessions", *Psychotherapy Research*, Vol. 30, No. 5, 2020, pp. 558-573.
- [9] J. Bell, J. Strang, "Medication treatment of opioid use disorder", *Biological Psychiatry*, Vol. 87, No. 1, 2020, pp. 82-88.
- [10] C. Liu, P. S. Kaeser, "Mechanisms and regulation of dopamine release", *Current Opinion in Neurobiology*, Vol. 57, 2020, pp. 46-53.
- [11] E. Saboory, M. Ghasemi, N. Mehranfard, "Nor-epinephrine, neurodevelopment and behavior", *Neurochemistry International*, Vol. 135, 2020, p. 104706.
- [12] P. A. Coventry et al. "Psychological and pharmacological interventions for posttraumatic stress disorder and comorbid mental health problems following complex traumatic events: Systematic review and component network meta-analysis", *PLoS Medicine*, Vol. 17, No. 8, 2020, p. e1003262.
- [13] N. Rohleder, "Stress and inflammation—The need to address the gap in the transition between acute and chronic stress effects", *Psychoneuroendocrinology*, Vol. 105, 2020, pp. 164-171.
- [14] D. M. Campagne, "Stress and perceived social isolation (loneliness)", *Archives of gerontology and geriatrics*, Vol. 82, 2020, pp. 192-199.
- [15] E. Beurel, M. Toups, C. B. Nemeroff, "The bidirectional relationship of depression and inflammation: double trouble", *Neuron*, Vol. 107, No. 2, 2020, pp. 234-256.
- [16] M. Daly, A. R. Sutin, E. Robinson, "Depression reported by US adults in 2017-2018 and March and April 2020", *Journal of Affective Disorders*, Vol. 278, 2020, pp. 131-135.
- [17] F. Zhang et al. "Causal influences of neuroticism on mental health and cardiovascular disease", *Human Genetics*, Vol. 140, 2021, pp. 1267-1281.
- [18] T. van der Zweerde, L. Bisdounis, S. D. Kyle, J. Lancee, A. van Straten, "Cognitive behavioral therapy for insomnia: a meta-analysis of long-term effects in controlled studies", *Sleep Medicine Reviews*, Vol. 48, 2019, p. 101208.
- [19] H. Pan et al. "The perception of menopause among women in Taiwan", *Maturitas*, Vol. 41, No. 4, 2002, pp. 269-274.
- [20] K. D. Mochrie, M. C. Whited, T. Cellucci, T. Freeman, A. T. Corson, "ADHD, depression, and substance abuse risk among beginning college students", *Journal of American College Health*, Vol. 68, No. 1, 2020, pp. 6-10.
- [21] S. G. Burdisso, M. Errecalde, M. Montes-y-Gómez, "A text classification framework for simple and effective early depression detection over social media streams", *Expert Systems with Applications*, Vol. 133, 2019, pp. 182-197.
- [22] D. Su, X. Zhang, K. He, Y. Chen, "Use of machine learning approach to predict depression in the elderly in China: A longitudinal study", *Journal of Affective Disorders*, Vol. 282, 2021, pp. 289-298.
- [23] W. Zheng, L. Yan, C. Gou, F. Y. Wang, "Graph attention model embedded with multi-modal knowledge for depression detection", *Proceedings of the IEEE International Conference on Multimedia and Expo*, London, UK, 6-10 July 2020.
- [24] F. M. Shah et al. "Early depression detection from social network using deep learning techniques", *Proceedings of the IEEE Region 10 Symposium*, Dhaka, Bangladesh, 5-7 June 2020, pp. 823-826.
- [25] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot, Y. Kaewpitakkun, "Facebook social media for depression detection in the Thai community", *Proceedings of the 15th International Joint Conference on Computer Science and Software Engineering*, Nakhonpathom, Thailand, 11-13 July 2018, pp. 1-6.
- [26] P. Arora, P. Arora, "Mining twitter data for depression detection", *Proceedings of the international*

- conference on signal processing and communication, Nodia, India, 7-9 March 2019, pp. 186-189.
- [27] G. Lam, H. Dongyan, W. Lin, "Context-aware deep learning for multi-modal depression detection", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Brighton, UK, 12-17 May 2019, pp. 3946-3950.
- [28] A. H. Orabi, P. Buddhitha, M. H. Orabi, D. Inkpen, "Deep learning for depression detection of twitter users", Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic, New Orleans, LA, USA, 5 June 2018, pp. 88-97.
- [29] M. R. Islam et al. "Depression detection from social network data using machine learning techniques", Health information science and systems, Vol. 6, 2018.
- [30] R. Chiong, G. S. Budhi, S. Dhakal, F. Chiong, "A textual-based featuring approach for depression detection using machine learning classifiers and social media texts". Computers in Biology and Medicine, Vol. 135, 2021, p. 104499.
- [31] S. M. Alghowinem, T. Gedeon, R. Goecke, J. Cohn, G. Parker, "Interpretation of depression detection models via feature selection methods", IEEE Transactions on Affective Computing, Vol. 14, No. 1, 2020, pp. 133-152.
- [32] M. Y. Wu, C. Y. Shen, E. T. Wang, A. L. Chen, "A deep architecture for depression detection using posting, behavior, and living environment data", Journal of Intelligent Information Systems, Vol. 54, 2020, pp. 225-244.
- [33] C. Y. Chiu, H. Y. Lane, J. L. Koh, A. L. Chen, "Multi-modal depression detection on instagram considering time interval of posts", Journal of Intelligent Information Systems, Vol. 56, 2021, pp. 25-47.
- [34] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, H. Adeli, D. P. Subha, "Automated EEG-based screening of depression using deep convolutional neural network", Computer Methods and Programs in Biomedicine, Vol. 161, 2018, pp. 103-113.
- [35] C. Uyulan et al. "Major depressive disorder classification based on different convolutional neural network models: Deep learning approach", Clinical EEG and Neuroscience, Vol. 52, No. 1, 2021, pp. 38-51.
- [36] M. A. Wani, M. A. ELAffendi, K. A. Shakil, A. S. Imran, A. A. Abd El-Latif, "Depression screening in humans with AI and deep learning techniques", IEEE Transactions on Computational Social Systems, Vol. 10, No. 4, 2022, pp. 2074-2089.
- [37] Y. Lin et al. "A deep learning-based model for detecting depression in senior population", Frontiers in Psychiatry, Vol. 13, 2022, p. 1016676.
- [38] A. Rafiei, R. Zahedifar, C. Sitaula, F. Marzbanrad, "Automated detection of major depressive disorder with EEG signals: a time series classification using deep learning", IEEE Access, Vol. 10, 2022, pp. 73804-73817.

# Sentiment Mining in E-Commerce: The Transformer-based Deep Learning Model

Original Scientific Paper

## Tahani Alsaedi

Department of computer science and information  
Applied College, Taibah University,  
Madinah, Saudi Arabia  
tbsaedi@taibahu.edu.sa

## Asif Nawaz

University Institute of Information Technology,  
PMAS-Arid Agriculture University,  
Rawalpindi, Pakistan,  
asif.nawaz@uaar.edu.pk

## Abdulrahman Alahmadi

Department of Computer Science and Information,  
Taibah University,  
Medina 42353, Saudi Arabia  
aahmadi@taibahu.edu.sa

\*Corresponding author

## Muhammad Rizwan Rashid Rana\*

University Institute of Information Technology,  
PMAS-Arid Agriculture University,  
Rawalpindi, Pakistan,  
rizwanrana315@gmail.com

## Ammar Raza

Herschel grammar school,  
Slough, United Kindom  
raza.ammar1029@gmail.com

**Abstract** – Sentiment analysis is crucial for comprehending customer feedback and enhancing workplace culture, as well as improving products and services. By employing natural language processing (NLP) techniques to meticulously analyze this feedback, organizations can identify specific areas that require improvement, address employee issues, and cultivate a positive work environment. These deep learning models powered by NLP offer invaluable tools for HR and sales departments in the e-commerce sector, enabling them to track sentiment trends among employees and users over time and implement targeted interventions. Focusing on the e-commerce industry, this study employs NLP-driven deep learning methodologies to analyze both employee and user feedback, with the objective of identifying underlying sentiments. The proposed framework leverages these advanced techniques to categorize user feedback into positive, negative, or neutral sentiments. This approach aims to develop a robust and effective system for sentiment analysis, providing significant insights that can help drive organizational improvements and enhance customer satisfaction. The key steps of this framework include data collection, NLP-enhanced feature extraction, sentiment detection, and final classification using finite-state automata. The effectiveness of this NLP-centric approach was tested on diverse datasets of customer feedback collected from an e-commerce industry. Evaluation metrics such as accuracy, precision, and recall were utilized to assess the performance of the system. The results demonstrate the effectiveness of the proposed framework, achieving a 93.75% accuracy rate and surpassing existing benchmark methods. The outcomes of this study are particularly consequential for the e-commerce sector, offering them a strategic advantage in refining their product portfolios and cultivating a more dynamic workplace culture

**Keywords:** sentiment analysis, machine learning, natural language processing, ecommerce, products

Received: March 3, 2024; Received in revised form: July 8, 2024; Accepted: July 8, 2024

## 1. INTRODUCTION

In the dynamic and competitive e-commerce industry, the role of workplace culture in driving an organization's success is more crucial than ever. A positive culture within e-commerce companies leads to higher

employee engagement, increased productivity, and better retention rates, essential in an industry known for its fast pace and high turnover [1]. Conversely, a negative workplace environment can be particularly detrimental in this sector, affecting not just employees

but also customer satisfaction and sales. E-commerce organizations, therefore, actively seek out employee feedback to continually adapt and improve their cultures [2]. However, the challenge lies in making sense of the vast amounts of unstructured data, which in e-commerce can be even more diverse and subjective due to the wide range of roles and interactions in this sector. This is where sentiment analysis, empowered by Natural Language Processing (NLP) and Deep Learning (DL), becomes an indispensable tool. It assists in effectively analyzing and understanding employee feedback, a task that is essential for maintaining a robust and adaptive culture in the fast-paced world of e-commerce [3].

Sentiment analysis, a technique that involves analyzing textual data to identify underlying sentiments or opinions, becomes significantly powerful with the integration of machine and deep learning technologies [4]. Aspect-based sentiment analysis (ABSA) delves deeper by focusing on specific aspects or entities mentioned in the text, making it particularly effective for extracting opinions on distinct features such as design, performance, and customer service, as often seen in product reviews. Driven by natural language processing (NLP), sentiment analysis enables organizations to swiftly analyze vast amounts of employee feedback, decoding perceptions, attitudes, and sentiments related to workplace culture. This tool is essential for understanding and enhancing organizational dynamics.

The field of sentiment analysis is fraught with challenges due to the complex nature of human language. Employees often express their opinions in unique and subjective ways, which can obscure the true meaning of their sentiments. Additionally, personal biases, past experiences, and cultural factors shape an employee's perspective [5]. Despite these significant obstacles, sentiment analysis offers numerous benefits to organizations. Analyzing employee feedback provides insights into areas of workplace culture that need improvement, such as work-life balance, communication, and recognition practices. By addressing these areas, organizations can enhance employee engagement and productivity, leading to increased profitability and sustainable growth. This paper delves into the complexities of analyzing employee feedback, thoroughly examining the challenges involved. Moreover, it explores how machine learning and deep learning technologies address these issues, providing effective solutions.

This study conducted an in-depth NLP-based analysis of a comprehensive collection of sentiment data, encompassing reviews from e-commerce. Initially, we fine-tuned the BERT model, a cutting-edge NLP tool, and subsequently employed an aggregating layer to generate text embeddings, crucial for understanding the nuances in language. Ultimately, we developed an intricately integrated model, BiGRU-Senti-GNN, for a more accurate classification of sentiments. To further enhance our NLP approach, we incorporated finite au-

tomata for refined sentiment detection. The following points annotate the main contribution of this paper:

- This work has proposed the hybrid NLP-driven deep framework using BERT-BiGRU-Senti-GNN for the classification of sentiments.
- A novel approach is introduced for building a graph used in graph convolutional networks over dependency trees, incorporating effective knowledge. This method aims to capture emotional relationships related to particular aspects.
- The experimental evaluation claims that the proposed technique produces better accuracy with a value of almost 93.75% compared to baseline approaches.

The structure of the rest of the paper is as follows: Section 2 presents a brief review of relevant literature on sentiment analysis. Section 3 outlines the methodological framework used for review classification. The results of our experimental setup and parameters are discussed in Section 4. Lastly, Section 5 summarizes the findings, contributions, and key insights of this study.

## 2. LITERATURE REVIEW

This section discusses the existing studies of sentiment analysis of customer reviews in multiple domains.

Deng et al, introduce a novel BERT-ETextCNN-ELSTM model for sentiment analysis on social media comment data [6]. This model combines BERT for word embedding and encoding, followed by optimized CNN for local feature extraction and LSTM for capturing long-term dependencies. Compared to baseline models, the proposed approach achieves superior performance, with accuracy, F1 value, and macro-average F1 value reaching up to 0.89, 0.88, and 0.86, respectively, on the datasets. The suggested model proves its efficacy in sentiment analysis, particularly excelling in review sentiment tasks compared to other comparable models. In another research, Diao et al. propose an innovative approach for detecting sentiment in social media text [7]. The research acknowledges the limitations of existing sentiment detection methods, which often struggle to capture the complex nature of sentiment and the contextual information necessary for accurate detection. To address these limitations, the authors introduce a multi-dimensional question Answering Network (MDQAN) incorporating various linguistic features and contexts to identify sarcastic utterances effectively.

Another research introduces a novel method for sentiment detection in a social media text. This method addresses challenges like figurative language and irony [8]. It involves three main steps: sentiment classification using a convolutional neural network (CNN), feature extraction by combining lexical, syntactic features, and individual expression habits, and final classification using a support vector machine (SVM). The proposed approach surpasses existing methods on SemEval-2018 and sentiment Corpus datasets, showing strong perfor-

mance measured by F1 scores. Notable strengths are its use of sentimental context and individual expression habits, which enhance interpretability. However, the method relies on sentiment classification and feature extraction modules and may have limitations in specific contexts or expressions.

Pandey and Singh address sentiment detection challenges in code-mixed social media by introducing a BERT-LSTM model that combines contextual information from BERT with sequential dependencies captured by LSTM networks [9]. Their model demonstrates superior performance in sarcasm detection over existing methods. Similarly, Pandey et al. (2021) propose an attention-based LSTM network enhanced with a CNN for sentiment identification, achieving the highest F1-score among various models and highlighting the model's utility in social media and customer reviews [10-15]. Shrawankar and Chandankhede (2019) focus on sarcasm detection in workplace communication, employing a two-stage approach integrating machine learning and rule-based methods to improve accuracy, contributing to the context of workplace stress management [11].

Another study addresses the importance of sentiment analysis in the age of abundant online comments, utilizing AI to understand public opinion [12]. The research introduces an innovative approach by enhancing traditional TF-IDF word representation with sentiment information, yielding weighted word vectors. These vectors are then processed through a bidirectional LSTM to capture context effectively, resulting in improved comment representation. Liang et al., introduce "Sentic GCN," a graph convolutional network that utilizes SenticNet's emotional relationships for analyzing sentences based on specific aspects [13]. The approach integrates affective knowledge from SenticNet to enhance sentence dependency graphs, incorporating contextual and aspect words' dependencies. Experimental findings on various benchmark datasets demonstrate Sentic GCN's superiority over existing methods, showcasing its effectiveness in ABSA.

This paper presents an effective sentiment analysis Deep Learning Modified Neural Network (DLMNN) technique applied to Twitter data [14]. First, the data is transferred to the Hadoop Distributed File System (HDFS) where duplicate words are identified and removed using the MapReduce technique. Then, classification is carried out using a Deep Learning Modified Neural Network (DLMNN). Another research introduces a Two State GRU (TS-GRU) model with an integrated feature attention mechanism [15]. The TS-GRU model uses sequential modeling to determine sentiment polarity by examining word features in detail. It incorporates a pre-feature attention phase to explore complex word relationships and emphasize significant sentiments, and an attention layer to identify key keywords. Authors proposed Target Word Transferred ABSA (WordTransABSA) for ABSA [16]. WordTransABSA generates context-specific semantics and predicts affective

tokens at positions aligned with aspect terms. The final sentiment polarity for each aspect term is determined using a selection of carefully chosen sentiment identification strategies.

The studies discussed offer innovative sentiment detection techniques, yet they encounter several challenges. These include the inherent ambiguity of language, diverse forms of sarcasm, the need for robust contextual understanding, dependencies in sentiment analysis, biases in training data, and limitations in context-specific applicability. Additionally, some methods depend on pre-trained embeddings that may not adequately capture all contextual nuances. Despite the valuable contributions of each study to the field, these issues highlight the continued complexity in accurately identifying sentiments in natural language and emphasize the need for ongoing research and methodological refinement.

### 3. PROPOSED METHODOLOGY

This section introduces the BERT-Senti-Graph-Convolutional Network (BERT-BiGRU-Senti-GCN) framework. Initially, the framework begins with data preprocessing, encompassing several steps to clean user text reviews. Subsequently, word embeddings are generated using Bidirectional Encoder Representations from Transformers (BERT). Feature extraction is carried out using BiGRU, followed by classification with Senti-GCN. Finally, finite automata are employed for sentiment ranking. The proposed framework is illustrated in Fig 1.

#### 3.1. DATA PREPROCESSING

Effective data preprocessing holds a pivotal role in the realm of text data analysis [17]. As seen in tweets, blogs, reviews, and similar content, textual data often exhibit repetitions and redundancies, introducing complexities. To mitigate these challenges, data normalization serves as a foundational filtration technique. The normalization process encompasses various procedures such as word tokenization [18], elimination of stop words, trimming extra spaces, padding, converting text to lowercase, and eradicating hashtags, among other preprocessing steps.

#### 3.2. WORD EMBEDDINGS

Each word is carefully converted into a numerical vector representation through a method called word embedding. This technique allows for the generation of accurate textual representations for words with similar meanings. It is a part of unsupervised learning, focusing on capturing semantic relationships. In the proposed framework, the application of RoBERTa for creating these word embeddings is crucial [19].

RoBERTa (Robustly Optimized BERT Approach) is an advanced natural language processing model developed by Facebook AI, designed to surpass the capabilities of



the original BERT (Bidirectional Encoder Representations from Transformers) model. BERT revolutionized NLP by employing bidirectional training to understand the context of words from both directions [20]. Building on this, RoBERTa incorporates several significant enhancements. It utilizes larger mini-batches during training, which stabilizes and improves the optimization process by processing more data in each training step, resulting in better gradient estimates and enhanced convergence. Additionally, RoBERTa employs longer training sequences, enabling it to handle longer text inputs and better capture long-range dependencies and contextual information within the text. A notable improvement is the elimination of the next sentence prediction task, present in BERT, allowing RoBERTa to focus solely on masked language modeling, leading to more accurate contextual word representations. By leveraging larger datasets and extensive computational resources, RoBERTa achieves superior performance, making it a more robust and powerful model for a wide array of NLP applications.

### 3.3. FEATURE EXTRACTION

BiGRU (Bidirectional Gated Recurrent Unit) is a type of recurrent neural network that processes sequences

both forwards and backwards, effectively capturing contextual information for tasks like sentiment analysis [21]. In the proposed framework, the BiGRU layer follows the word embeddings layer, processing tokenized and padded input sequences to extract features that reflect the complex relationships between words [22]. These features can be utilized for sentiment analysis by either using the final hidden states of the BiGRU or combining outputs from different timesteps using methods like max-pooling or averaging. The framework utilizes a three-layer BiGRU (BiGRU-3) where each layer's output is the input for the next, enhancing the model's ability to capture deeper linguistic features.

$$G_u = \alpha (W_i * [h_{t-1}, y_t]) \quad (1)$$

$$G_r = \alpha (W_r * [h_{t-1}, y_t]) \quad (2)$$

$$h_t = \tanh (W_c * [G_r, h_{t-1}, y_t]) \quad (3)$$

Here,  $\alpha$  represents the sigmoid function and ' $*$ ' signifies the dot product. In this context,  $y_t$  denotes the input vector at time  $t$ , while  $h_t$  represents the hidden state.  $G_r$  acts as a reset gate, responsible for discarding control information; together, they collaboratively influence the resulting output of the hidden state.

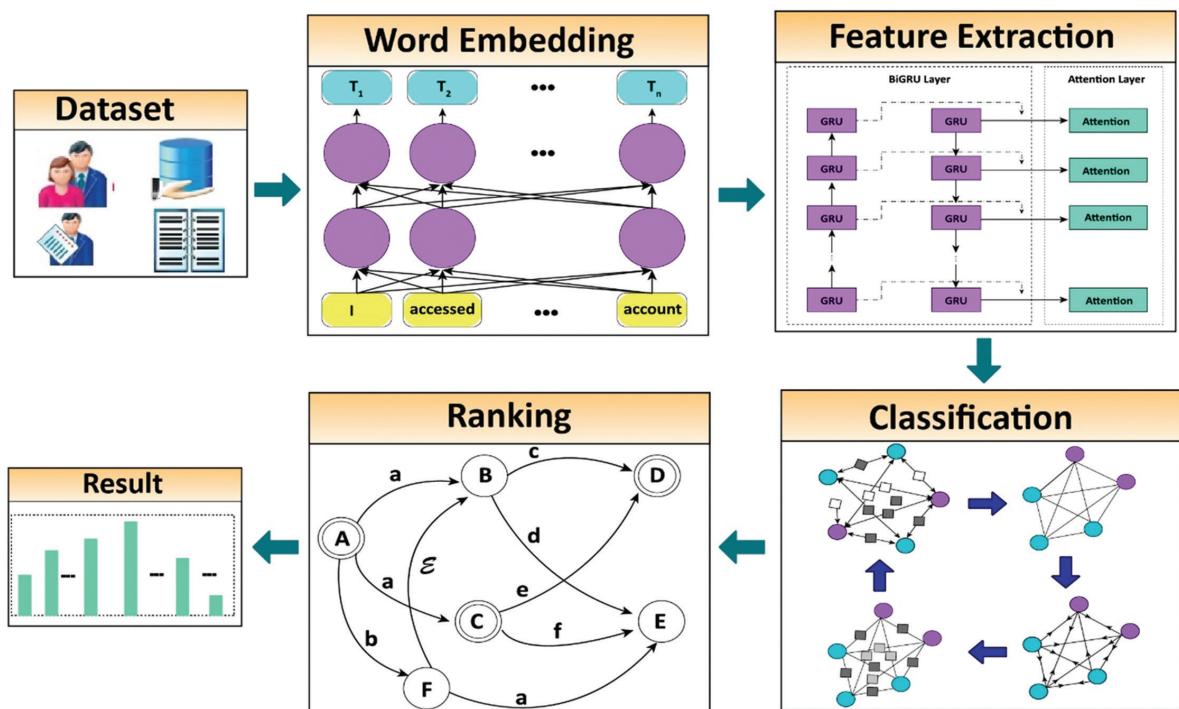


Fig. 1. Proposed framework for sentiment classification

### 3.4. CLASSIFICATION

Graph Convolutional Networks (GCNs) are designed to process graph-structured data, such as social networks, citation networks, and molecular graphs [23-25]. They adapt convolutional operations—commonly used in image recognition—to graphs, allowing the network to learn features based on a node's neighborhood. In this approach, the GCN layers utilize an

enhanced adjacency matrix, which integrates affective scores from SentiWordNet. This integration helps capture the emotional nuances of words, enriching the matrix and enabling more precise analysis of affective interdependencies between contextual and aspect words [26-28]. The goal is to deepen the understanding of the emotional relationships inherent in the language, improving the effectiveness of the network in handling tasks involving complex emotional contexts..

$$Sc_{x,y} = SWN(weight_x) + SWN(weight_y) \quad (4)$$

where SentiWordNet ( $weight_i$ )  $\in [-1, 1]$  shows the affective score of the word  $weight_i$  in SentiWordNet. Here, SentiWordNet ( $weight_i$ ) = 0 shows the word  $weight_i$  is a neutral word or inexistent in SentiWordNet. The sentiment association between two related words is computed by aggregating their emotional scores, leveraging words with strong emotional connotations to enhance sentiment analysis. Current graph convolutional network (GCN) models for aspect sentiment analysis often fail to emphasize the designated aspect in graph construction. This study addresses this limitation by enhancing the emotional interdependencies between contextual and aspect words using SentiWordNet, improving the model's effectiveness in aspect-focused sentiment analysis.

To highlight the key features of aspect words, we utilize aspect-specific masking. This technique involves masking non-aspect words in the output vectors obtained from the final Graph Convolutional Network (GCN) layer, while keeping the original representations of the aspect words intact:

$$h_t = \begin{cases} h_{tt} & \text{if } \tau < t < \tau + k \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Here,  $h_{tt}$  denotes the learned representation of the word at the  $t$ -th position generated by the final output of the GCN layers. In this context, " $\tau$ " indicates the starting index of words associated with a specific aspect within the sentence, while " $k$ " represents the length of this aspect-related word group. Later, an attention mechanism rooted in retrieval is harnessed to assess both the emotional content and semantic connections between the words in the context and the aspect-related words. This strategy encompasses the construction of a graph using affective common-sense knowledge sources like SentiWordNet, which reveals emotional dependencies within the sentence. Additionally, a specialized operation that centers around the aspect-related words is applied, thereby amplifying the significance of these words. The process involves using the softmax function to obtain the output distribution for the sentiment classifier.

$$Y_{output} = softmax(Wr + b) \quad (6)$$

### 3.5. RANKING FOR SENTIMENT DETECTION

Detecting sentiment in text is inherently complex due to the multifaceted and context-dependent nature of user reviews. This task requires a deep understanding of emotions, contextual nuances, and intricate linguistic patterns. While finite automata can be utilized for elementary string-matching tasks, they are inadequate for capturing the subtleties of sarcasm, which is inherently nuanced and context-sensitive. Finite automata are mathematical models consisting of states and transitions between these states based on input symbols. They are commonly used to identify patterns in strings, such as detecting specific words or phrases. To apply fi-

nite automata to sentiment detection, one would need to devise a set of patterns or phrases that signify sarcasm and irony. These patterns would then be encoded into a finite automaton, which would be employed to analyze input text for occurrences of sentiments.

Identification of patterns is very important in sentiment analysis. These patterns can include specific words, phrases, or grammatical structures frequently used sarcastically. For example, phrases like "Great job!" to indicate dissatisfaction or "Oh, that's just what I needed!" to convey frustration. Construct a finite automaton based on the sarcastic patterns identified in the previous step. The automaton will consist of states and transitions representing the characters or sequences of characters in the pattern.

- **State Definition:** In the automaton, each state signifies a particular character or sequence of characters within the sarcastic pattern. For instance, one state might represent the letter "O" in the phrase "Oh, sure."
- **Transition Definition:** Transitions between states are triggered by characters in the input text. For example, if the current state corresponds to the letter "O" and the subsequent character is "h," the automaton will transition from the "O" state to the "h" state.
- **Accepting States:** Identify the states that signify the completion of the sarcastic pattern. These accepting states indicate that the automaton has successfully detected sarcasm.

The steps for matching the strings of reviews using finite automata are:

- **Initialize the automaton:** Set the initial state of the automaton to the starting state, representing the beginning of the sarcastic pattern.
- **Process the input text:** Iterate through each character in the feedback instance and follow the transitions in the automaton accordingly. Move from one state to another based on the encountered characters.
- **Handle invalid transitions:** If no transition is defined for a particular character, handle it appropriately. You can choose to ignore the character or reset the automaton to the starting state.
- **Track the final state:** After processing the entire feedback instance, note the final state of the automaton.

Examine the final state of the automaton to determine if sentiment has been detected.

- **Accepting state:** If the final state is accepting, it indicates that the feedback instance matches the sarcastic pattern, and sarcasm is present.
- **Non-accepting state:** If the final state is not accepting, it implies that the feedback instance does not match the sarcastic pattern and is non-sarcastic.

An example of the pattern "ababaca" is shown in Fig. 2.

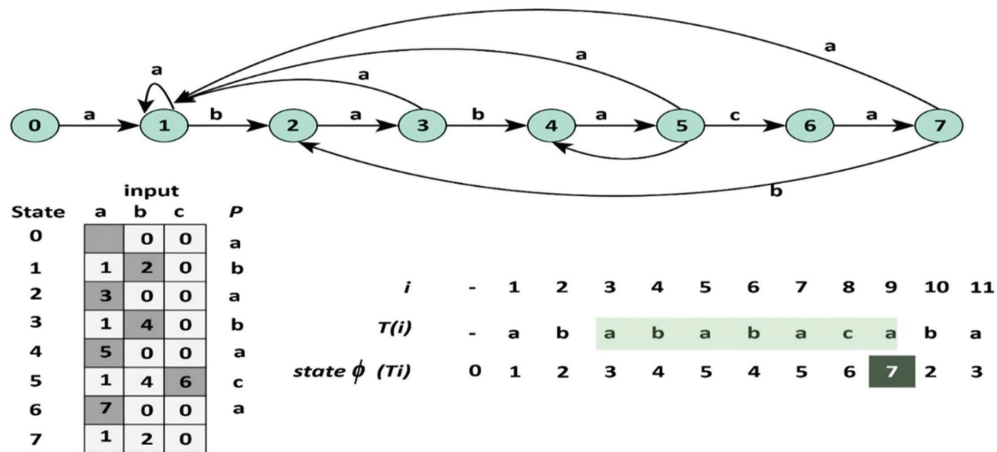


Fig. 2. Example of pattern matching

#### 4. EXPERIMENTAL EVALUATION

In this part, we give a thorough examination of the experimental analysis as well as an assessment of the accuracy and efficiency of the proposed framework through a series of rigorous experiments. The results show that the proposed framework outperforms the existing cutting-edge techniques and beats state-of-the-art models in terms of performance. The descriptive explanation of each experiment is provided in the below subsections.

##### 4.1. DATASET DESCRIPTION

Describing the datasets utilized in this research is essential before getting into the in-depth experimental evaluation. Three multidomain datasets have been used for experiments. Each dataset was meticulously annotated with positive and negative classifications. The first dataset is a benchmark dataset containing different customer reviews. The second dataset, the "SemEval dataset" comprises a wide range of ratings and comments from different consumers along the product images, concentrating on computers and restaurants. The third dataset is a database that provides product metadata, reviews and image vectors for 9.4 million Amazon goods. However, we'll concentrate on the clothes, shoes and jewelry categories, which comprise over 1.5 million goods. In addition to text, it also has audio and visual components. The dataset breakdown is shown in Table 1.

Table 1. Datasets breakdown

Dataset ID	Dataset Name	Reviews	Web Link
DS-I	Customer Review	2,270	<a href="https://shorturl.at/q9ZQa">https://shorturl.at/q9ZQa</a>
DS II	SamEval Dataset	11,577	<a href="https://shorturl.at/Wx3vs">https://shorturl.at/Wx3vs</a>
DS III	Amazon Dataset	109,102	<a href="https://shorturl.at/cGuAp">https://shorturl.at/cGuAp</a>

##### 4.2. BASELINE METHODS

The following baseline reference models have been considered for comparison to evaluate the efficiency of the proposed framework.

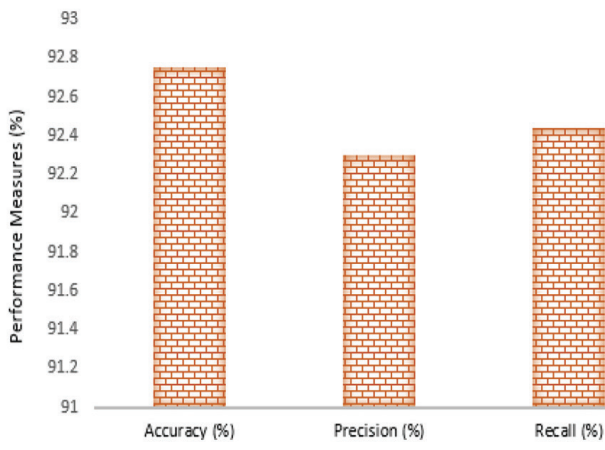
- DGEDT [29] employs a dual-transformer network along with a dependency graph to capture both flat and graph-based representations.
- BERT+GNN [30] combines BERT with a Graph Convolutional Network (GCN) and dependency trees, using selective attention to identify key words for aspect representations.
- KSCB [31] is a hybrid model that integrates K-means++, SMOTE, CNN, and Bi-LSTM for text sentiment analysis.
- TNet-LF [32] uses a CNN for targeted sentiment classification at the aspect level
- POS+SentiGA [33] calculates polarity weights based on the specific aspect of user text reviews for sentiment analysis.

##### 4.3. RESULT

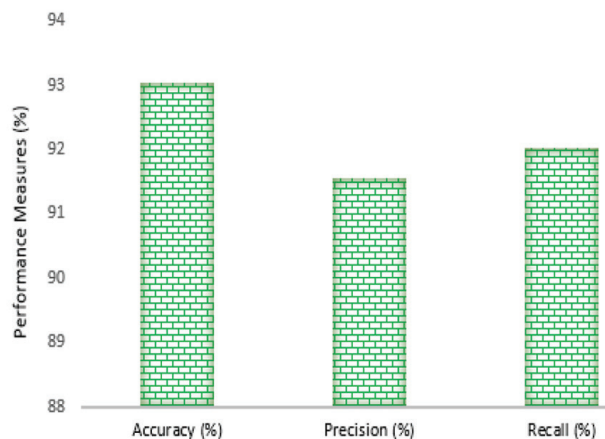
The first experiment assesses the proposed framework's precision, accuracy and recall for three datasets. The graph in Fig. 3 shows that the proposed method achieved spectacular results on all datasets, suggesting its outstanding precision, accuracy, and recall performance across all datasets. On the customer dataset, the proposed approach produced an impressive precision of 92.25%, recall of 92.44%, and accuracy of 92.44%. Similarly, when evaluated on the SamEval dataset, the findings were even more promising, with 91.45% precision, 91.95% recall, and 93.16% accuracy. In the third dataset of Amazon reviews, the proposed framework achieved 95.64% accuracy, 95.02% precision and 95.32% recall.

A performance-based confusion matrix has also been portrayed to measure the efficiency of the proposed framework in terms of True- +ve (TP), False- +ve, True- -ve (TN), and False- -ve (FN). Fig. 4 shows a schematic illustration of predicated and actual values datasets.

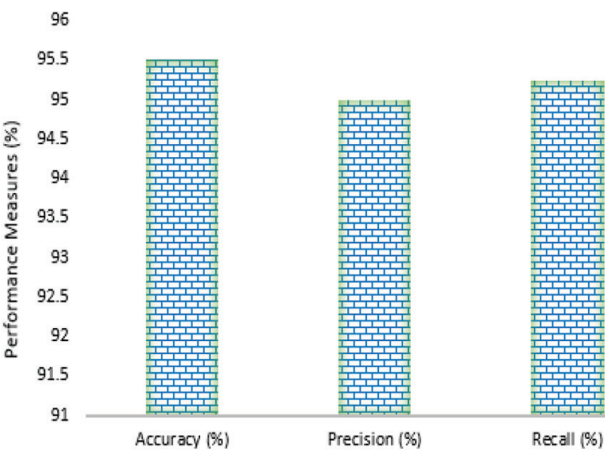




(a) Customers Dataset



(b) Amazon Dataset

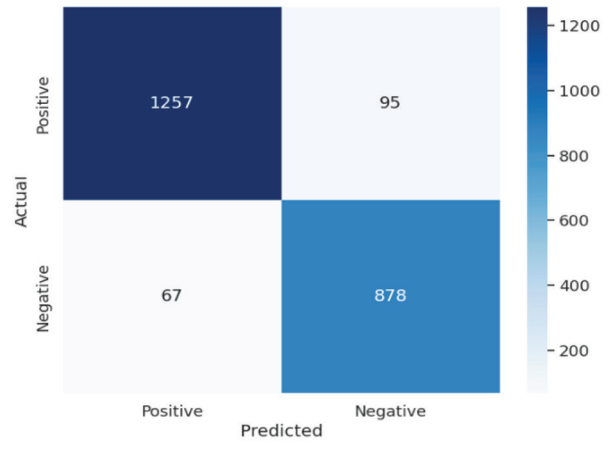


(c) SamEval Dataset

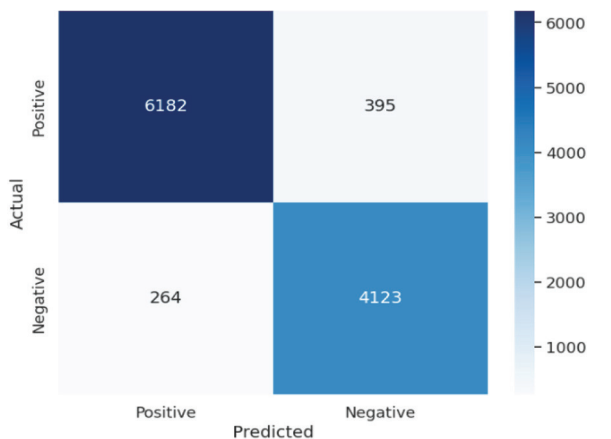
**Fig. 3. Experimental Results**

The next assessment compares the effectiveness of the proposed model with three different traditional deep learning configurations, each designed for feature extraction and sentiment analysis.

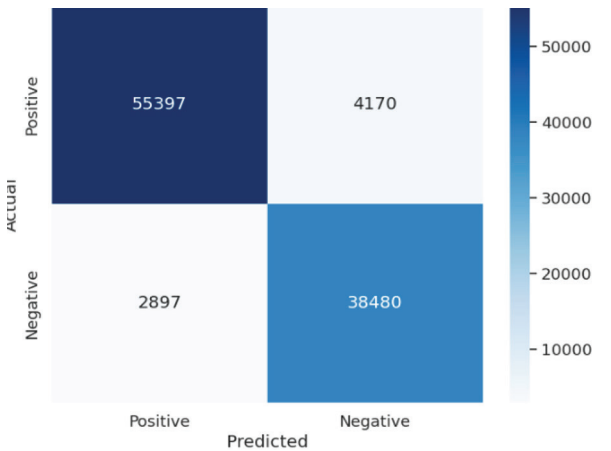
These configurations are named CNN-CNN, RNN-RNN, and LSTM-LSTM. The findings show that, although models incorporating CNN and RNN layers perform admirably, they do not surpass the proposed model, which consistently achieves superior results across all evaluation metrics. The results are shown in Fig. 5. As shown in Fig. 6, the empirical results clearly dem-



(a) Customers Dataset



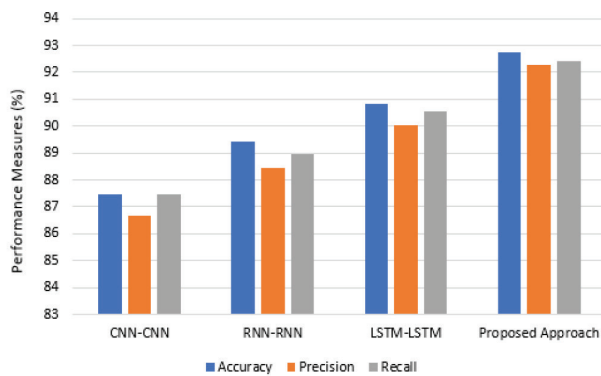
(b) Amazon Dataset



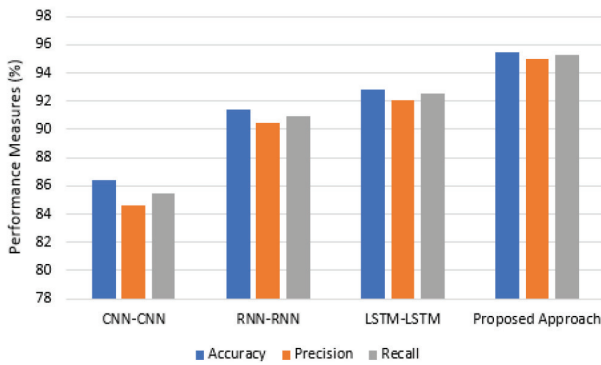
(c) SamEval Dataset

**Fig. 4. Confusion Matrix**

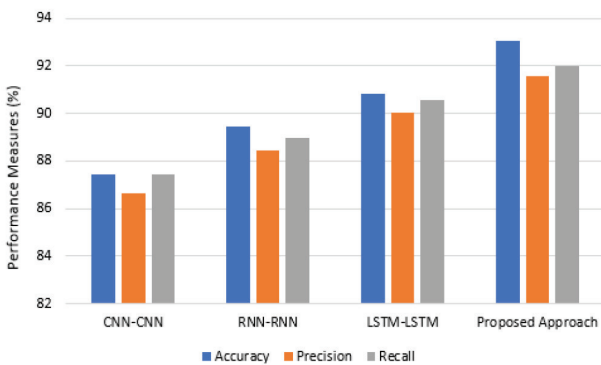
onstrate that our innovative framework significantly outperforms all other models on three well-known benchmark datasets. These datasets include deep neural networks, graph networks, and BERT-based models. This highlights the effectiveness of our proposed model in the ABSA domain. Our new SentiWordNet-GCN surpasses previous models, especially GCN-based models that build graphs over sentence dependency trees. This provides strong evidence of the benefits and effectiveness of enhancing word dependencies within a sentence by leveraging the emotional insights from SentiWordNet.



(a) Customers Dataset

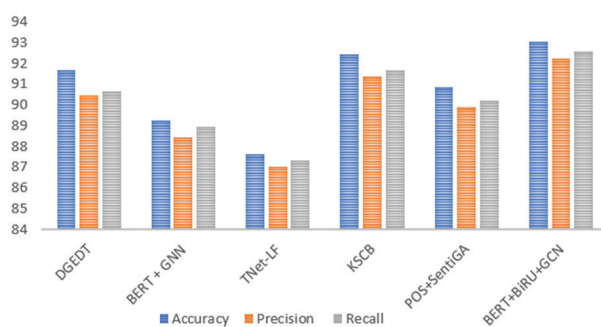


(b) Amazon Dataset



(c) SamEval Dataset

**Fig. 5.** Comparison with CNN-CNN, RNN-RNN and LSTM-LSTM



**Fig. 6.** Comparison of proposed approach with baseline approaches

The success of this method can be attributed to the significant interdependence of parent-child links within sentence dependency trees. In natural language processing, understanding the syntactic relationships

between words is crucial for accurate sentiment analysis. By leveraging these relationships, a unidirectional graph within convolutional networks can effectively capture the intricate dependency structure, leading to more precise sentiment predictions.

Our innovative Affective GCN model capitalizes on this by constructing a sentence graph that integrates affective words from SentiWordNet-GNN. This integration allows the model to utilize not only the syntactic dependencies but also the emotional context provided by SentiWordNet. As a result, the Affective GCN model exhibits superior performance compared to the BERT + GNN and DGED models. It demonstrates that incorporating affective information into the GCN framework significantly enhances the model's ability to understand and predict sentiments. Furthermore, our model delivers results on par with more sophisticated graph-based models, underscoring the efficacy of our approach. By embedding sentiment information from SentiWordNet into the GCN framework, we enhance the model's capability to capture nuanced emotional subtleties in text. This proves particularly beneficial for Aspect-Based Sentiment Analysis (ABSA), where understanding the sentiment associated with specific aspects of a product or service is crucial.

In conclusion, the integration of syntactic dependencies and affective insights from SentiWordNet within a unidirectional graph structure offers a powerful approach to sentiment analysis. Our Affective GCN model's outstanding performance demonstrates the value of this integration, paving the way for more accurate and nuanced sentiment analysis in various NLP applications. This approach not only improves the precision of sentiment predictions but also contributes to a deeper understanding of the emotional context in textual data, making it a valuable tool for enhancing sentiment analysis in ABSA and beyond.

## 5. CONCLUSION

This paper introduces BERT-BiGRU-Senti-GCN, an innovative NLP methodology tailored for analyzing customer feedback in the e-commerce industry. The framework integrates BERT-BiGRU for robust feature extraction with Senti-GNN for precise sentiment classification, and employs finite automata for efficient pattern recognition. By weaving emotional data into a graph structure, it significantly enhances the analysis of relationships between aspect-specific and contextual words, thereby refining sentiment analysis. Demonstrated through rigorous evaluation on several multidomain benchmark datasets, the model achieves leading performance metrics in the field of NLP. Future initiatives will focus on enhancing sentiment diffusion within the graph structure, improving the model's responsiveness to domain-specific variations via transfer learning and domain adaptation, and incorporating advanced pattern recognition using deep learning techniques. Moreover, the potential for real-time

implementation and scalability of this model will be explored to ensure its applicability across diverse e-commerce platforms.

## 6. REFERENCES

- [1] R. J. K. Almahmood, A. Tekerek, "Issues and Solutions in Deep Learning-Enabled Recommendation Systems within the E-Commerce Field", *Applied Sciences*, Vol. 12, No. 21, 2022, p. 11256.
- [2] M. R. R. Rana, A. Nawaz, T. Ali, A. M. El-Sherbeeney, W. Ali, "A BiLSTM-CF and BiGRU-based Deep Sentiment Analysis Model to Explore Customer Reviews for Effective Recommendations", *Engineering, Technology & Applied Science Research*, Vol. 13, No. 5, 2023, pp. 11739-11746.
- [3] L. Deng, B. Liu, Z. Li, "Multimodal sentiment analysis based on a cross-modal multihead attention mechanism," *Computers, Materials & Continua*, Vol. 78, No.1, 2024, pp. 1157-1170.
- [4] K. H. Leong, D. P. Dahnail, "Classification of Healthcare Service Reviews with Sentiment Analysis to Refine User Satisfaction", *International Journal of Electrical and Computer Engineering Systems*, Vol. 13, No. 4, 2022, pp. 323-330.
- [5] M. R. R. Rana, A. Nawaz, T. Ali, G. Mustafa, "Enhancing healthcare services recommendation through sentiment analysis", *Acta Universitatis Sapientiae, Informatica*, Vol. 15, No. 2, pp. 330-344.
- [6] L. Deng, T. Yin, Z. Li, Q. Ge, "Sentiment Analysis of Comment Data Based on BERT-ETextCNN-ELSTM", *Electronics*, Vol. 12, No. 13, 2023, p. 2910.
- [7] Y. Diao, H. Lin, L. Yang, X. Fan, Y. Chu, K. Xu, D. Wu, "A Multi-Dimension Question Answering Network for Sarcasm Detection", *IEEE Access*, Vol. 8, 2020, pp. 135152-135161.
- [8] Y. Du, T. Li, M. S. Pathan, H. K. Teklehaimanot, Z. Yang, "An Effective Sarcasm Detection Approach Based on Sentimental Context and Individual Expression Habits", *Cognitive Computation*, Vol. 14, No. 1, 2021, pp. 78-90.
- [9] R. Pandey, J. P. Singh, "BERT-LSTM model for sarcasm detection in code-mixed social media post", *Journal of Intelligent Information Systems*, Vol. 60, No. 1, 2022, pp. 235-254.
- [10] R. Pandey, A. Kumar, J. P. Singh, S. Tripathi, "Hybrid attention-based Long Short-Term Memory network for sarcasm identification", *Applied Soft Computing*, Vol. 106, No. 1, 2021, p. 107348.
- [11] M. Shrivastava, S. Kumar, "A pragmatic and intelligent model for sarcasm detection in social media text", *Technology in Society*, Vol. 64, No. 1, 2021, p. 101489.
- [12] G. Xu, Y. Meng, X. Qiu, Z. Yu, X. Wu, "Sentiment Analysis of Comment Texts Based on BiLSTM", *IEEE Access*, Vol. 7, 2019, pp. 51522-51532.
- [13] B. Liang, H. Su, L. Gui, E. Cambria, R. Xu, "Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks", *Knowledge-Based Systems*, Vol. 235, 2022, p. 107643.
- [14] D. Paulraj, P. Ezhumalai, M. Prakash, "A Deep Learning Modified Neural Network (DLMNN) based proficient sentiment analysis technique on Twitter data", *Journal of Experimental & Theoretical Artificial Intelligence*, Vol. 6, No. 3, 2024, pp. 415-434.
- [15] M. Zulqarnain, R. Ghazali, M. Aamir, Y. M. N. Hasim, "An efficient two-state GRU based on feature attention mechanism for sentiment analysis", *Multimedia Tools and Applications*, Vol. 83, No. 1, 2024, pp. 3085-3110.
- [16] W. Jin, B. Zhao, Y. Zhang, J. Huang, H. Yu, "WordTransABSA: Enhancing Aspect-based Sentiment Analysis with masked language modeling for affective token prediction", *Expert Systems with Applications*, Vol. 238, 2024, p. 122289.
- [17] H. T. Duong, T. A. Nguyen-Thi, "A review: preprocessing techniques and data augmentation for sentiment analysis", *Computational Social Networks*, Vol. 8, No. 1, 2020.
- [18] A. Naresh, P. V. Krishna, "An efficient approach for sentiment analysis using machine learning algorithm", *Evolutionary intelligence*, Vol. 14, No. 2, 2021, pp. 725-731.
- [19] U. Sirisha, S. C. Bolem, "Aspect based sentiment & emotion analysis with ROBERTa, LSTM", *International Journal of Advanced Computer Science and Applications*, Vol. 13, No. 11, 2022.
- [20] K. L. Tan, C. P. Lee, K. S. M. Anbananthen, K. M. Lim, "RoBERTa-LSTM: a hybrid model for sentiment

- analysis with transformer and recurrent neural network”, *IEEE Access*, Vol. 10, 2022, pp. 21517-21525.
- [21] Y. Han, M. Liu, W. Jing, “Aspect-level drug reviews sentiment analysis based on double BiGRU and knowledge transfer”, *IEEE Access*, Vol. 8, 2020, pp. 21314-21325.
- [22] Z. Gao, Z. Li, J. Luo, X. Li, “Short text aspect-based sentiment analysis based on CNN+ BiGRU”, *Applied Sciences*, Vol. 12, No. 5, 2022, p. 2707.
- [23] Y. Tian, Y. Chen, Y. Song, “Aspect-based sentiment analysis with type-aware graph convolutional networks and layer ensemble”, *Proceedings of the Conference of The North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, Online, 6-11 June 2021, pp. 2910-2922.
- [24] B. Liang, R. Yin, L. Gui, J. Du, R. Xu, “Jointly learning aspect-focused and inter-aspect relations with graph convolutional networks for aspect sentiment analysis”, *Proceedings of the 28th International Conference on Computational Linguistics*, Online, 8-13 December 2020, pp. 150-161.
- [25] J. Zhou, J. X. Huang, Q. V. Hu, L. He, “Sk-gcn: Modeling syntax and knowledge via graph convolutional network for aspect-level sentiment classification”, *Knowledge-Based Systems*, Vol. 205, 2020, p. 106292.
- [26] P. Zhao, L. Hou, O. Wu, “Modeling sentiment dependencies with graph convolutional networks for aspect-level sentiment classification”, *Knowledge-Based Systems*, Vol. 193, 2020, p. 105443.
- [27] S. Zhang, H. Tong, J. Xu, R. Maciejewski, “Graph convolutional networks: a comprehensive review”, *Computational Social Networks*, Vol. 6, No. 1, 2019, pp. 1-23.
- [28] M. Husnain, M. M. S. Missen, N. Akhtar, M. Cousataty, S. Mumtaz, V. S. Prasath, “A systematic study on the role of SentiWordNet in opinion mining”, *Frontiers of Computer Science*, Vol. 15, No. 4, 2021, p. 154614.
- [29] H. Tang, D. Ji, C. Li, Q. Zhou, “Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification”, *Proceedings of the 58th Annual Meeting of The Association for Computational Linguistics*, Online, 6-8 July 2020, pp. 6578-6588.
- [30] X. Hou, J. Huang, G. Wang, X. He, B. Zhou, “Selective attention based graph convolutional networks for aspect-level sentiment classification”, *arXiv:1910.10857*, 2019.
- [31] W. Jiang, K. Zhou, C. Xiong, G. Du, C. Ou, J. Zhang, “KSCB: a novel unsupervised method for text sentiment analysis”, *Applied Intelligence*, Vol. 53, No. 1, 2022, pp. 301-311.
- [32] X. Li, L. Bing, W. Lam, B. Shi, “Transformation networks for target-oriented sentiment classification”, *arXiv:1805.01086*, 2018.
- [33] M. R. R. Rana, S. U. Rehman, A. Nawaz, T. Ali, M. Ahmed, “A conceptual model for decision support systems using aspect based sentiment analysis”, *Proceedings of the Romanian Academy, Series A*, Vol. 22, No. 4, 2022, pp. 371-380.

# End-to-End Fully Automated Lung Cancer Volume Estimation System

Original Scientific Paper

## Pushkar Sathe\*

Department of Electronics and Telecommunications Engineering, Mukesh Patel School of Technology Management & Engineering, SVKM's Narsee Monjee Institute of Management Studies (NMIMS) Deemed-to-University, Mumbai-400056, India  
pushkarsathe@gmail.com

## Alka Mahajan

JK LakshmiPat University, Jaipur-302026, India.  
alka.mahajan@jklu.edu.in

\*Corresponding author

## Deepak Patkar

Department of Radiology, Nanavati Superspeciality Hospital, Mumbai-400056, India  
drdppatkar@gmail.com

## Mitusha Verma

Department of Radiology, Nanavati Superspeciality Hospital, Mumbai-400056, India  
drmitusha@gmail.com

**Abstract** – The volume of the tumor plays a very crucial role in deciding the stage of lung cancer which in turn helps in deciding the best treatment and its schedule. Currently used computer-based volume estimation techniques are semi-automatic with limited accuracy. For any automatic lung cancer segmentation system, lung CT scans of hundreds of patients are required along with their corresponding annotated segmentation masks. It is difficult to get accurately annotated data as cancer segmentation of CT scans done by the radiologists, is a time-consuming manual process. Also, it is subjective and prone to intra and inter-observer variability. Further, owing to the irregular shape of the cancerous tumor, accurate volume estimation becomes a challenge with regular convolution models. This paper proposes an end-to-end automatic tumor volume estimation model that estimates volume using the GPR (Gaussian Process Regression) interpolation method. The proposed modified cancer segmentation model uses deformable convolutions. This modification offers a higher segmentation accuracy in terms of IoU (Intersection over Union) and clearly defined nodule boundaries with correct retention of the nodule shape. The research was undertaken in collaboration with Nanavati Hospital, Mumbai, and all the models were validated on a real dataset obtained from the hospital. The proposed model gives a mean segmentation IoU (Intersection over Union) of 0.9035 and a volume estimation accuracy of 93.13% which are almost 5% and 3% higher than 0.8548 and 90.51% which are the corresponding results obtained using a standard U-net++ algorithm.

---

**Keywords:** Cancer segmentation, Deep learning, Deformable convolution, Lung cancer volume estimation, U-net++

---

Received: March 13, 2024; Received in revised form: June 4, 2024; Accepted: July 2, 2024

## 1. INTRODUCTION

Lung cancer is the most widespread type of cancer and is the second most common cancer after prostate cancer in males and breast cancer in

females. The early detection of lung cancer plays an effective role in diagnosis and leads to an early treatment increasing the likelihood of patient survival rate [1, 2]. Chemotherapy, which is the main treatment for lung cancer requires knowledge of the accurate location of the cancer along with its spread along all the 3 axes. Accurate volume estimation is required for determining the stage of the cancer. Volume estimation can

be done using 3D or 2D CT scan images. Though 3D data gives more accurate volume estimation it suffers from a lack of annotated data required for training and higher complexity of the segmentation algorithms. So generally 2D data is used to estimate the slice-wise area of cancer spread and then volume is obtained using these area values. It is very difficult to get accurately annotated data as radiologists usually perform manual segmentation. Human intervention leads to errors arising out of fatigue and subjectivity. To overcome this problem, cancer segmentation models based on deep learning were proposed but they require a large dataset along with corresponding annotations of every CT



scan for training to get acceptable segmentation accuracy that leads to precise area calculation. Also, the majority of cancer segmentation algorithms fail to define the boundaries of cancerous nodules clearly. While many software programs are available for calculating volume automatically, they are semi-automatic as they need demarcations of cancerous portions to be done by radiologists in every or at most on alternate slices. In this paper, a fully automatic tumor volume estimation model is proposed that performs segmentation of cancerous portions by modifying the U-net++ algorithm using deformable convolutions. The volume is calculated using GPR interpolation. The proposed model offers highly accurate volume estimation and requires a lesser number of original CT scans.

The primary contributions of our proposed research are as follows:

- 1) Designed an improved segmentation algorithm based on a modified convolution that achieved an improved segmentation accuracy in terms of well-defined nodule boundaries and the retention of actual nodule shape.
- 2) Developed the cancer volume estimation system using interpolation techniques that provided better accuracy of volume estimation.

## 2. RELATED WORK

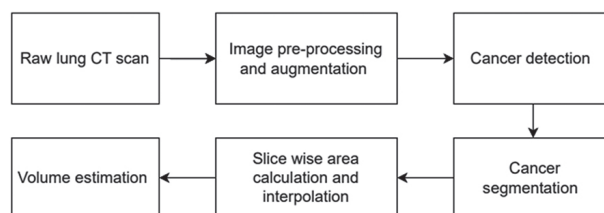
In recent times CT scans have been dominantly used for cancer detection and a lot of research on developing AI (Artificial Intelligence) based cancer detection algorithms is reported in the literature. Matt Daykin et al. used a One-Class Support Vector Machine (OCSVM) based model to detect lung abnormality [1]. The method, however, does not use complete images but patches of lung CT scans majorly to increase the size of the data set. Automatic feature extraction or the use of deep learning techniques is also not explored fully. Other reported abnormality detection algorithms [2, 3] majorly dealt with cancers other than lung cancer. Irigoien et al. applied OCC (One Class Classification) to medical data for the detection of various diseases like breast cancer, liver disorders, leukemia, etc. [2]. The authors compared the performance of four different algorithms namely Gaussian, mixtures of Gaussian, Parzen, and typicality approach based on their average AUC (Area Under Curve). The paper reported the best results with an average AUC of 77.4% with the typicality approach. Tarassenko et al. [3] also applied OCC to investigate normality using a large number of available mammograms which do not show any evidence of mass-like structures. The recent advancement in deep learning techniques allows for automatically extracting features from the images, thereby improving the overall performance as compared to conventional CAD (Computer Aided Diagnosis). The use of transfer learning techniques is also reported in literature where fixed feature extraction is done using a pre-trained network. Ardimento et al. [4] applied three transfer learning models viz; VGG (Visual Geometry

Group), Xception, and ResNet for feature extraction and combined the results using ensemble architecture to classify the scans as cancerous and noncancerous. Similarly, the state-of-the-art transferable architectures such as VGG-16, VGG-19, GoogLeNet, Inception-V3, ResNet-18, ResNet-50, ResNet-101, InceptionResNet-V2 and 3D multipath VGG like network have also been used for lung cancer feature extraction [5-9] and the performance of SVM (Support Vector Machine) and AdaBoostM2 classifier is analyzed on the deep features extracted from publicly available datasets. In [10] the performance of various machine learning algorithms was evaluated on a lung cancer detection task. The fusion of DenseNet201 with color histogram techniques was used to extract a hybrid feature set. Similarly, an innovative deep-learning model for lung cancer detection by integrating markers from mRNA, miRNA, and DNA methylation was developed [11]. The principal components analysis (PCA) was implemented to streamline features and the synthetic minority over-sampling technique (SMOTE) algorithm was applied to ensure class balance. The PCA-SMOTE model achieved an F1 score of 0.97. Lung nodule segmentation methods can be categorized into traditional approaches and deep learning-based techniques. Traditional methods encompass threshold and region-growing methods, clustering methods, active contour models, and mathematical models [12-20]. On the other hand, deep learning-based methods for segmentation can be further divided into 2D and 3D segmentation networks [21]. While traditional segmentation methods do not necessitate a substantial amount of labeled data for model training, they heavily rely on human intervention and are more focused. These methods primarily rely on shallow image features such as grayscale and texture. In contrast, 2D segmentation networks utilize 2D convolution to extract features from images. The original Fully Convolutional Neural Network (FCN) [22], has found extensive use in various image segmentation fields but it lacks global context information which is very important in the case of medical image segmentation. Ronneberger et al. [23] introduced a U-Net network based on the FCN architecture for medical image segmentation. U-Net incorporates both low-resolution and high-resolution information through skip connections, which is particularly advantageous for segmenting medical images with blurred boundaries. In this approach, low-resolution information is used for target identification, while high-resolution information aids in the localization of the segmentation. Consequently, U-Net has served as a foundation for several improved algorithms. The average IoU of 77.5% is achieved but the algorithm failed to consider the multiscale information required for accurate segmentation. So, to further improve segmentation results, some researchers have explored the integration of Atrous Spatial Pyramid Pooling (ASPP) [24] instead of the intermediate or output layer of U-Net. This approach enables the extraction of multi-scale image information through different perceptual fields [25-27] but lacks in retention of exact shapes. Though the average IoU of 81.3 was achieved, the algorithm lacked complete recovery

of spatial information. Some 3D segmentation networks utilize 3D convolution to extract features from volumetric images, allowing them to better capture the spatial relationship between nodules and surrounding tissues compared to 2D segmentation networks. Duo et al. [28] introduced a 3D fully convolutional neural network that automatically segments the liver and cardiac great vessels. Milletari et al. [29] incorporated residual connections inspired by ResNet [30] and employed 3D convolution in their V-Net architecture for prostate volume segmentation, departing from the 2D convolution used in U-Net. Deepseed [31] proposed 3D-squeeze-and-excitation (SE) networks, incorporating dynamic scaling with cross-entropy loss [32] to address the sample imbalance problem. Their method was evaluated on the LIDC (Lung Image Database Consortium) [33] and LUNA16 (LUng Nodule Analysis) [34] datasets, demonstrating promising results for lung nodule segmentation. However, it is worth noting that 3D networks suffer from longer training times and may not necessarily outperform 2D networks in terms of evaluation metrics. Zhou et al. [35] proposed U-net++, an extension of U-Net that incorporates dense skip connections to enable multi-scale fusion and feature acquisition at different levels. The model achieves an average IoU of 81.4% but as the algorithm involves non-integral convolution, it fails in maintaining the exact shape of nodules. In [36] the attention module is used with the U-net algorithm having residual connections for fast backpropagation. But model failed to maintain the exact shape of cancerous nodules. Similarly, different deep learning segmentation models were evaluated on different datasets, and the effects of different preprocessing methods were examined [37]. Though the TransUet model achieved the highest segmentation accuracy with an average dice coefficient of 0.81, it failed to maintain the exact shape of the cancerous nodule. Being a transformer-based model, the decision-making process and interpreting predictions was challenging. Also, in [38] a method to integrate feature information through a dual-branch network framework and multi-dimensional fusion module is proposed. By training and validating with multiple data sources and different data qualities, the method demonstrated leading performance on the LUNA16, Multi-thickness Slice Image dataset, LIDC, and UniToChest, with an average dice similarity coefficient of 82%. This method failed to maintain the cancer nodule's exact shape, leading to inaccurate segmentation. Accurate segmentation of cancerous nodules is of high importance as the accuracy of volume estimation and grading is directly depending on the same.

### 3. METHODOLOGY

Fig. 1 shows the schematic of the proposed system. Raw CT scans are first cleaned and then features are extracted from the isolated lung section. Extracted features are then applied to the cancer detection model as explained in the sections that follow.



**Fig. 1.** Schematic of the proposed system

#### 3.1 DATA COLLECTION AND PRE-PROCESSING

Normal and cancerous 2D lung CT scans were made available by the Nanavati Hospital, Mumbai. They were of DICOM (Digital Imaging and Communications in Medicine) file type. Each CT scan is of size 512 X 512. The dataset was checked, classified, and annotated independently by two radiologists and used for training and testing. A total of 872 normal and 146 cancerous scans were collected. Cancerous scans were increased to an optimum of 584 using data augmentation techniques approved by the doctors to balance the available data without the risk of overfitting. All the simulations were carried out using Python running on a server with AMD Ryzen 5 processor, 8GB RAM, and one NVIDIA GeForce GTX 1650 GPU with a compute capability of 7.5. Training of an algorithm took 14 hours and testing of applied image takes just 8 seconds.

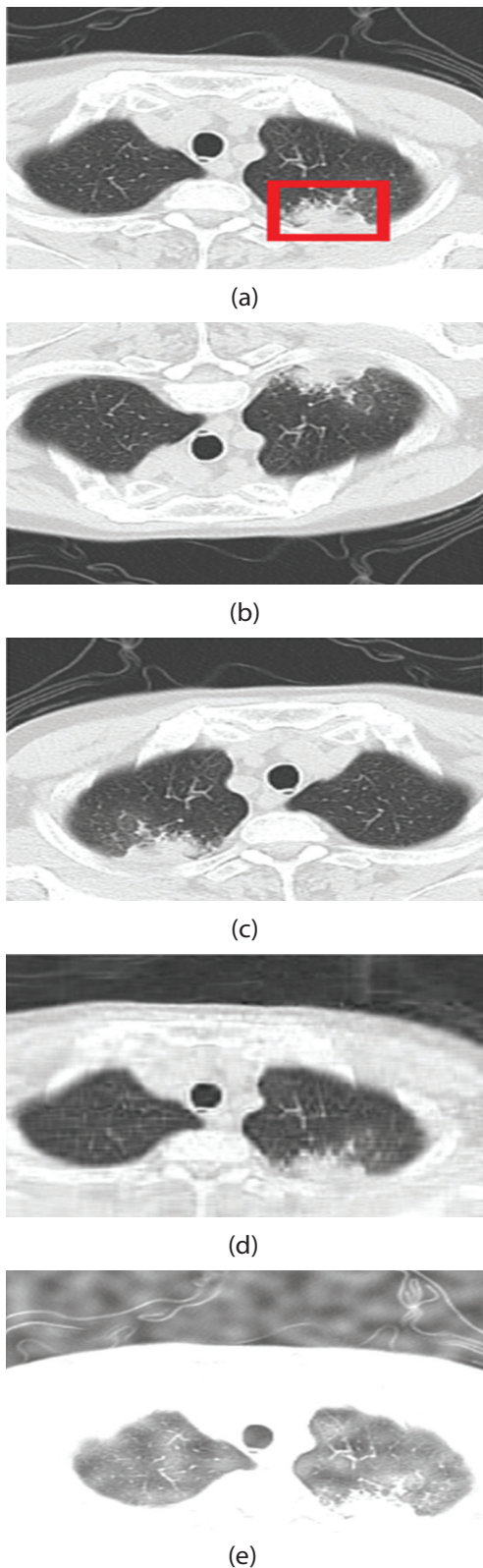
#### 3.2 DATA AUGMENTATION

A significant obstacle in training deep learning models for various tasks is the abundance of data required. This challenge is particularly pronounced in the field of medicine, where limited access to costly imaging resources or a scarcity of study subjects can hinder progress. Additionally, the requirement of annotations for every cancerous scan puts a limit on data collection. Consequently, there is a growing trend of employing data augmentation techniques in research involving small data sets [39, 40].

The demand for extensive data in deep learning networks has spurred the development of various strategies. Only those strategies are used in this work that provide a different visual impact compared to an original image. Vertical flipping, horizontal flipping, PCA (Principal Component Analysis), and image overlay techniques are used in our work for data augmentation.

Fig. 2 shows the images generated using corresponding data augmentation methods for a sample CT scan. Fig. 2(a) shows the original CT scan. The irregularly shaped cancerous nodule is highlighted using the red bounding box. Figs. 2(b) and 2(c) show the results of the flipping operation performed on the original CT scan. After vertical and horizontal flipping the position of the cancerous nodule changed, making it appear as a new image for a deep-learning model. Horizontal flipping is used because objects like cancer exhibit horizontal symmetry as it often looks similar when flipped

horizontally. Vertical flipping is useful in the medical domain where the orientation of objects like cancer varies significantly. Fig. 2(d) shows the result obtained by performing PCA with 20 components. Similarly, Fig. 2(e) shows the new image generated using the image overlay technique.



**Fig. 2.** (a) Original CT scan (b) Vertical flipping (c) Horizontal flipping (d) PCA with 20 principal components (e) image overlay

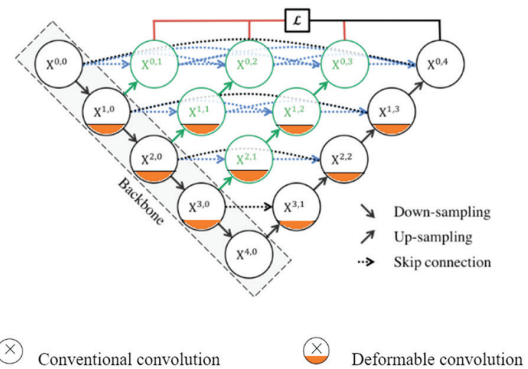
### 3.3 CANCER DETECTION

After data augmentation, the CT scans were applied to the cancer detection model. The cancer detection model was trained using both normal and cancerous CT scans. 872 normal and 584 cancerous CT scans were used in training. The features extracted using VGG were applied to three different classifiers namely SVM, Decision Tree, and Random Forest. The model was validated on a test dataset of 50 normal and 50 cancerous scans.

### 3.4 CANCER SEGMENTATION

The u-net algorithm proved to be the most significant segmentation algorithm for medical images [23]. Many variations of the basic algorithm have been developed in recent years. The basis of the U-net algorithm is the duplication of feature maps of the encoder to the decoder section to improve the quality of the upsampling feature map. U-net++ further improves the segmentation performance of the U-net algorithm by using the concept of deep supervision and nested skip connections that use convolutional layers [35]. However, in conventional convolutional layers, the receptive field remains unchanged for a given network structure, regardless of the object's size. But use of the same receptive field for objects of different sizes is not ideal. In the field of medical image segmentation, the lesions to be segmented often exhibit irregular shapes and sizes. The use of conventional convolution results in less accurate demarcation of cancer nodules [41, 42].

Fig. 3 shows the schematic of the proposed architecture consisting of 5 layers in the encoder and decoder.



**Fig. 3.** The schematic architecture of the proposed algorithm

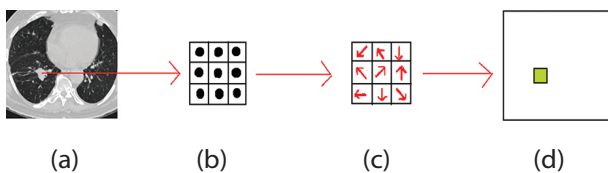
The encoder performs downsampling by capturing the contextual information and reduces spatial dimensions while the decoder performs upsampling and constructs the segmented image based on encoded features. The skip connections concatenate feature maps from the encoder to the decoder at the same spatial resolution. The nested skip pathways combine feature maps from different encoder resolutions to capture multi-scale contextual information. The weighted sum of the feature maps shown by in Fig. 3 improves the aggregation of features from different levels of the encoder.



+To further improve upon the segmentation accuracy, the proposed model uses deformable convolution instead of conventional convolution in the encoder and decoder layers. The increased geometric flexibility results in improved segmentation accuracy by retaining the original shape of the cancerous nodule.

We strategically incorporated deformable convolutions [41] into different parts of the network to identify the most effective place. Considering that the first downsampled convolution block is responsible for extracting the image's basic features, adding deformable convolutions to this block significantly prolongs network training time. Also, after multiple convolution and downsampling operations as the feature map size becomes very small and the feature information becomes abstracted, the benefits of using deformable convolutions are not evident. Taking these factors into account, we included the deformable convolution operation in the basic convolution block of the second, third, and fourth layers of the architecture as shown in Fig. 3.

### 3.4.1 DEFORMABLE CONVOLUTION



**Fig. 4.** (a) lung CT scan (b) Grid of conventional convolution showing fixed locations of points (c) trainable offsets to perform convolution at variable locations instead of fixed places (d) features map extracted

Fig. 4(a) shows the lung CT scan. In the case of conventional convolution, the process of convolution is performed using a fixed grid and because of this, the receptive field remains the same. Figure 4(b) shows a fixed convolution grid of size 3 x 3. It is not appropriate to use the constant receptive field in our case as cancer nodules are of varying shape and size. Deformable

convolution solves the problem of constant receptive field by using trainable offsets that are added to a fixed grid before performing convolution. Figure 4 (c) shows the trainable offsets. This process allows the network to learn information about nodule boundaries more accurately and it leads to better segmentation accuracy with retention of the exact shape of the cancerous nodule. Figure 4(d) shows the corresponding feature map extracted. Mathematically the deformable convolution can be represented as shown in Eq. (1).

$$y_i(P_0) = \sum_{P_n \in R} w(P_n) \cdot x(P_0 + P_n + \Delta P_n) \quad (1)$$

Here, the pixel value at  $P_0$  is replaced by weighted addition performed between image pixels and convolution mask at locations  $P_n + \Delta P_n$  where

$$P_n \in R \quad (2)$$

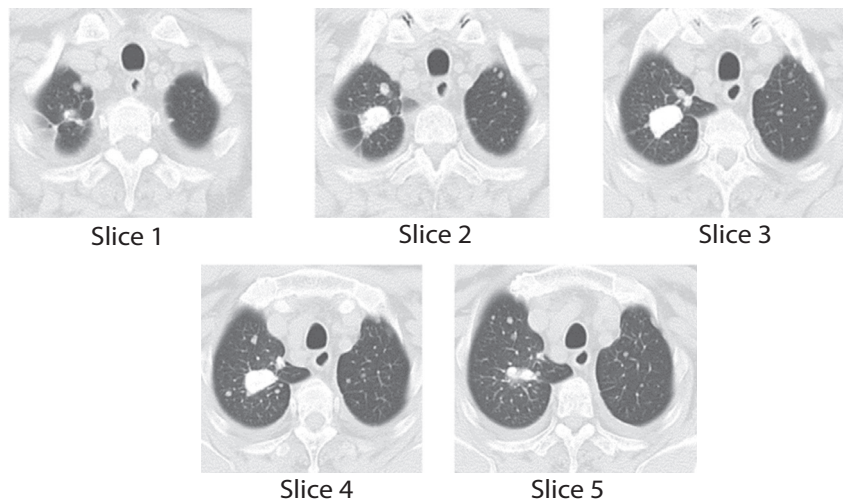
$$R = \{(-1, -1), (-1, 0), \dots, (0, 1), (1, 1)\} \quad (3)$$

for a grid of 3 x 3 size and  $\Delta P_n$  indicates the learnable distance offset for the best location to perform convolution. In equation (1),  $w$  indicates the convolution mask and  $x$  indicates the image.

### 3.5. VOLUME ESTIMATION

When a series of CT scans of the patient is applied to the cancer segmentation model, it provides corresponding masks having cancerous portions highlighted for every CT scan. These annotated masks are applied to the volume estimation model that calculates the area of the cancerous portion of every scan and uses the interpolation method to find the volume of the tumor.

From the segmented mask, several white pixels are obtained. The area of the cancerous portion is calculated by multiplying the number of white pixels by the area of each pixel which is obtained from pixel height and pixel width functions of the Python library. Likewise, the area is calculated for all consecutive CT scans and then the volume is obtained from these area values using the interpolation method. Fig. 5 shows the consecutive cancerous scans for one sample patient case. Corresponding area values are 0.435, 2.966, 4.104, 3.8, and 1.755 respectively.



**Fig. 5.** Consecutive cancerous scans of a sample patient

Each slice is 5 mm thick. This means the real area of cancer spread is known only at intervals of 5 mm each and not in between. Considering the same cancerous area throughout the slice thickness will lead to wrong volume calculation. For better volume estimation, interpolation is used to predict the area values of the cancerous portion 0.01 mm apart slice-wise. Various types of interpolations like linear, bilinear, cubic, Lagrange, and GPR were tried to determine the volume of cancer, and their results are compared with the actual volumes obtained from radiologists. GPR interpolation is found to be the best approach.

### 3.5.1. GPR INTERPOLATION

The basic idea of GPR is to model the relationship between the input data and the output values as a Gaussian process. A Gaussian process is a collection of random variables, any finite number of which have a joint Gaussian distribution. In GPR, we assume that the output values follow a Gaussian process with a mean function  $\mu(x)$  and a covariance function  $k(x, x')$  that captures the similarity between given data points and interpolation points.

The most commonly used covariance function is the Radial Basis Function (RBF) kernel. This interpolation is the most suitable for irregular data as it can better capture underlying trends and patterns. This property makes GPR the most suitable interpolation method in cancer volume estimation. The key idea of GPR interpolation is to use Bayesian inference to compute the posterior distribution over the unknown value  $\hat{y}$  conditioned on the observed data. The posterior distribution is a Gaussian distribution with a mean  $\hat{\mu}$  and a covariance  $\hat{\Sigma}$ . The mean  $\hat{\mu}$  represents the estimated value of  $\hat{y}$ , and the covariance  $\hat{\Sigma}$  quantifies the uncertainty associated with the estimation.

Mathematically GPR interpolation can be explained as follows:

First, the mean function  $\mu(x)$  and the covariance function  $k(x, x')$  were defined. Then the covariance matrix  $K(X, X)$  between the observed input points in  $X$  and the covariance vector  $k(X, \hat{x})$  between the observed input points in  $X$  and the new input point  $\hat{x}$  were computed. The covariance scalar  $k(\hat{x}, \hat{x})$  between the new input point  $\hat{x}$  and itself was then calculated. Finally, the mean vector  $\hat{\mu}$  and the covariance matrix  $\hat{\Sigma}$  of the posterior distribution were obtained using the formulae [43] given in equations (4) and (5)

$$\hat{\mu} = k(X, \hat{x})^T [K(X, X) + \sigma^2 I]^{-1} y \quad (4)$$

$$\hat{\Sigma} = k(\hat{x}, \hat{x}) - k(X, \hat{x})^T [K(X, X) + \sigma^2 I]^{-1} k(X, \hat{x}) \quad (5)$$

where  $\sigma^2$  is the noise variance parameter and  $I$  is the identity matrix.

The estimated output value  $\hat{y}$  at the new input point  $\hat{x}$  was obtained by the mean  $\hat{\mu}$ , and the uncertainty associated with the estimation was quantified by the covariance  $\hat{\Sigma}$ .

## 4. 4. RESULTS AND DISCUSSIONS

### 4.1. CANCER DETECTION

It was found that classification results are optimum for a random forest classifier for the finalized values of hyperparameters when compared to other classifiers. All the cancerous scans were correctly classified by Random Forest, with a recall (or sensitivity) of 100%, accuracy of 99.75%, and precision of 99.5%. Accuracy considers all predictions (both positive and negative) and measures the overall correctness of the model. Whereas, the precision focuses only on the positive predictions and measures how many of those predicted positives are true positives. Also, 100% recall indicates that the model has successfully identified every case who had a disease.

### 4.2. CANCER SEGMENTATION

In this work, the Intersection over Union (IoU), Dice Similarity Coefficient (DSC), Relative Volume Difference (RVD), Average Surface Distance (ASD), and Hausdorff Distance (HD) were used to evaluate the image segmentation results. IoU is calculated by finding the ratio of the area of overlap between the predicted segmentation and the ground truth segmentation to the area of union between the two. In mathematical terms, it is defined as shown in equation (6)

$$IoU = |X \cap Y| / |X \cup Y| \quad (6)$$

Here,  $X$  and  $Y$  represent the sets or regions of pixels that belong to the predicted segmentation and the ground truth segmentation, respectively. A higher  $IoU$  value indicates a better segmentation result, with a value of 1 indicating a perfect overlap between the predicted and ground truth segmentations. The Dice coefficient is a similarity measure commonly used in image segmentation to evaluate the overlap between a segmented mask and a ground truth mask. Mathematically it is defined as given in equation (7).

$$DSC = (2 * |X \cap Y|) / (|X| + |Y|) \quad (7)$$

RVD quantifies the relative difference in volume or size between the segmented region and the ground truth region.

$$RVD = (|Volume segmented - ground truth volume|) / ground truth volume$$

The RVD value is a measure of how closely the segmented region's volume matches the ground truth volume. It's typically expressed as a percentage, with values closer to zero indicating a more accurate segmentation. If RVD is exactly zero, it means the segmentation perfectly matches the ground truth. A positive RVD indicates an overestimation, meaning the segmented region is larger than the ground truth, while a negative RVD indicates an underestimation, meaning the segmented region is smaller than the ground truth.

ASD on the other hand quantifies the average distance between the surfaces of the segmented region

and the corresponding surfaces of the ground truth region. It calculates the distance between each corresponding pair of points, one from the segmented region's surface and one from the ground truth region's surface. Then it calculates the average of these distances. Smaller ASD values indicate a more accurate segmentation, as they imply that the segmented region's surface is closer to the ground truth surface.

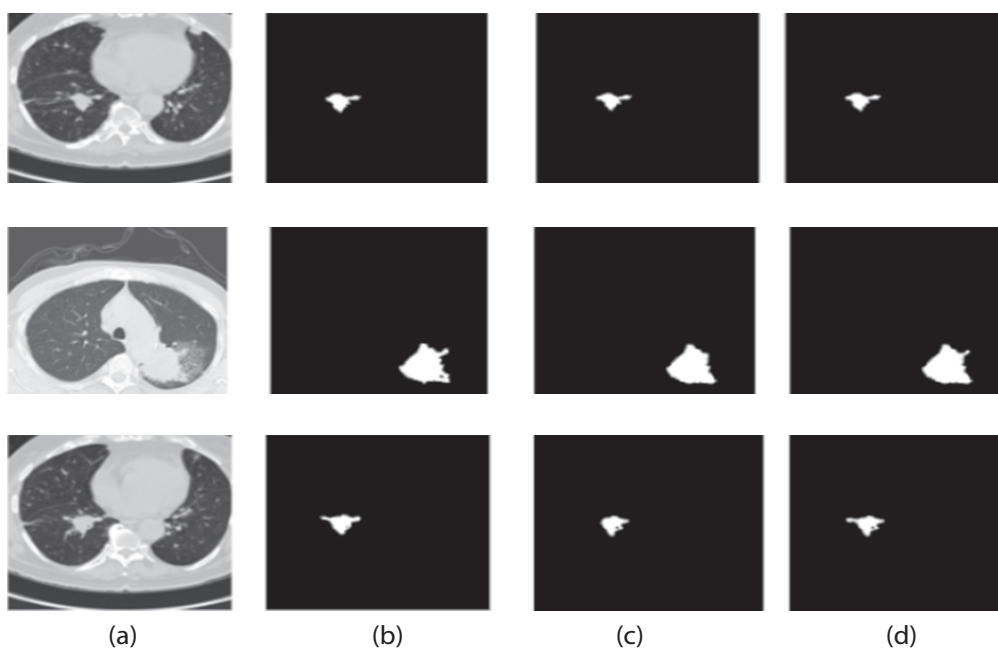
HD is also a measure of the dissimilarity between the boundary of a segmented object and a ground truth boundary. It calculates the Euclidean distance between each point in one set and its nearest neighbor in the other set. First, it calculates the Hausdorff distance from the first set to the second set (H1) as the maximum distance among the recorded Euclidean distances from the above step. Similarly, the Hausdorff Distance from the second set to the first set (H2) is calculated. The HD is then defined as the maximum of H1 and H2.

The deformable convolution was used basically to improve the accuracy of the cancerous nodule segmentation. Table 1 shows the performance metrics for the proposed algorithm with regular convolution-based U-net++. The performance metrics for 3 patients CTP1, CTP2, and CTP3 are highlighted in Table 1. It is seen that for all three patients, the RVD value is reduced

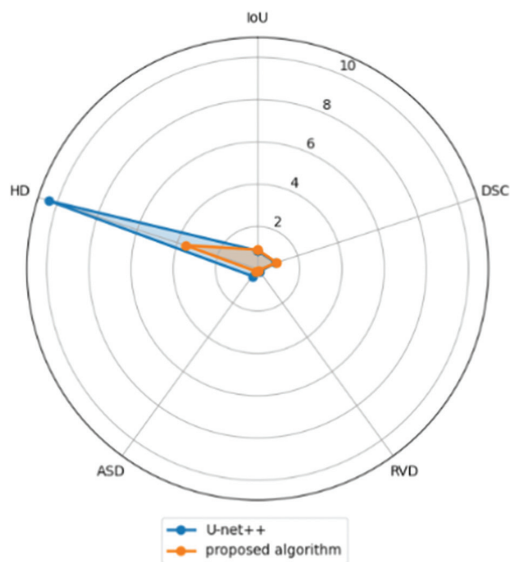
with the proposed algorithm compared to that with U-net++ which indicates improvement in segmentation. Similarly, there is a reduction in ASD value for three patients, particularly for patients CTP2 and CTP3. Also, the average values of performance parameters for 58 patients indicate that IoU and DCS show considerable improvement of 5.69% and 3.72 % respectively for the proposed model as compared to the U-net++. Similarly, other parameters RVD, ASD, and HD have been reduced for the proposed model by 41.02%, 65.67%, and 65.39% respectively which indeed indicates better segmentation. Fig. 6 shows the results of segmentation obtained using U-net++ and the proposed algorithm for patients. It is seen clearly that the proposed algorithm retains the shape of a nodule by correctly demarcating its boundaries and provides more accurate segmentation compared to U-net++. The retention of the exact shape of the cancerous nodule is of prime concern as the accuracies of the estimation of the highest dimension, volume are directly dependent on it. The spider chart showing the comparison of U-net++ and the proposed algorithm for various segmentation metrics is shown in Fig. 7. The larger difference in the areas covered by the two algorithms in the spider chart indicates the difference in their performance.

**Table 1.** Comparison of proposed algorithm with U-net++ for various segmentation metrics

Patients	Algorithm/ parameters	IoU	DSC	RVD	ASD	HD
CTP1	U-net++	0.8607	0.925	0.1256	0.1725	3.1623
	Proposed algorithm	0.8887	0.941	0.0874	0.1157	2.00
CTP2	U-net++	0.7723	0.872	0.2375	0.8460	15.6525
	Proposed algorithm	0.8327	0.909	0.1	0.1998	3.6056
CTP3	U-net++	0.9305	0.964	0.0529	0.2115	11.4018
	Proposed algorithm	0.9441	0.971	0.0371	0.0712	3.6056
Average for 58 patients	U-net++	0.8548	0.913	0.156	0.428	10.419
	Proposed algorithm	0.9035	0.947	0.092	0.1469	3.6056



**Fig. 6.** (a) Original CT scan (b) Ground truth (c) result of U-net++ (d) result of the proposed algorithm



**Fig. 7.** Spider chart of comparison of U-net++ with the proposed algorithm

#### 4.2. VOLUME ESTIMATION

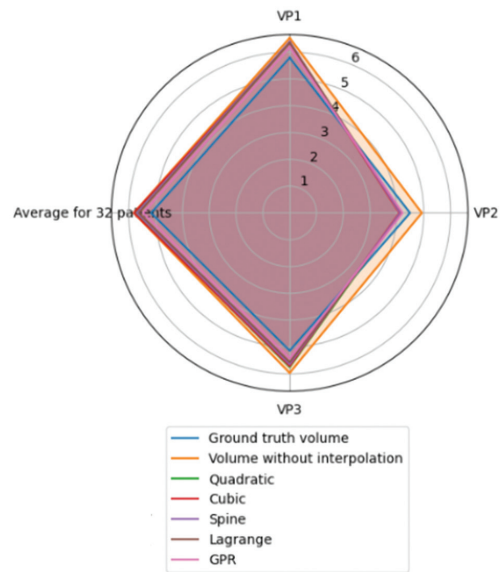
The results of the cancer segmentation model are applied to the volume estimation model. The

segmentation model generates annotations corresponding to all the cancerous scans. The cancerous portion is demarcated in white and the area corresponding to this cancerous portion is calculated in Python. Interpolation is performed on obtained area values to get area values of unknown places between the slices. Then the volume is obtained by summing up all area values. Various types of interpolations like linear, bilinear, cubic, Lagrange, and GPR (Gaussian Process Regression) are checked in this work and their results are compared. Table 2 shows the results of these interpolation methods applied to CT scans of 3 patients. VP1, VP2, and VP3 are the volumes estimated for 3 patients P1, P2, and P3 respectively. These volumes are compared with ground truth volumes calculated by the radiologists. It is seen that there is a huge difference in calculated volume and ground truth volume when interpolation is not used. In the case of VP3, volume estimation accuracies with the quadratic, cubic, spine, Lagrange, and GPR interpolations are 88.13%, 88.91%, 90.27%, 91.63%, and 94.36% respectively. Similarly, for the other patients, it is found that GPR gives the best volume estimation accuracy with an average value of 93.13%. Fig. 8 indicates the spider chart showing the comparison of various interpolation methods. It is evident from the diagram that the GPR curve is the closest to the ground truth volume.

For comparison purposes, we applied GPR interpolation on the segmentation masks obtained using the U-net++ algorithm. Table 3 indicates volumes obtained using GPR interpolation applied to the U-net++ algorithm and our proposed algorithm for VP1, VP2, and VP3. It is seen that the accuracy of the proposed algorithm is approximately 4% higher than that of U-net++.

The comparison of results obtained is shown in Fig. 9. For all volumes VP1, VP2, and VP3 as well as for average it is seen that the volume estimated using the proposed algorithm is closer to the ground truth compared to that of the U-net++ algorithm.

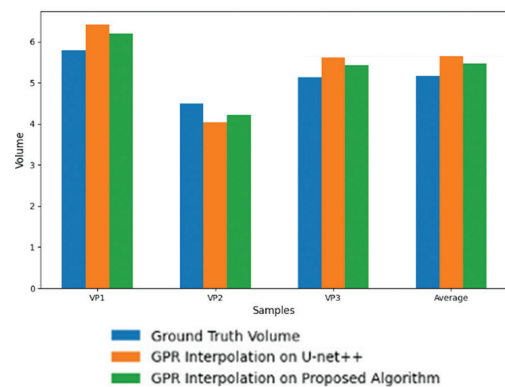
Volume is one of the important parameters in deciding the stage of the cancer. It also helps radiologists to further decide about the treatment. Though cancer grading depends on multiple factors including the spread of cancer to lymph nodes, spread of cancer to another lung, etc. volume is considered to be the most important factor.



**Fig. 8.** Spider chart of comparison of various interpolation methods

So if volume estimation goes wrong, it may lead to the inaccurate prediction of the stage of cancer and hence the corresponding dosage and the treatment.

Keeping the aim of a fully automated end-to-end volumetric estimation system, in this work we proposed the idea of deformable convolution and GPR interpolation to achieve better accuracy of cancer nodule segmentation and volume estimation respectively.



**Fig. 9.** Bar graph of comparison of volume estimation using U-net++ and proposed algorithm



## 5. CONCLUSION

This paper proposes a fully automatic system that automatically segments the cancerous portions from the applied CT scans and calculates the volume of the cancerous nodule. The proposed model uses a modified cancer segmentation algorithm which is based on deformable convolution. The modification captures the shape of the nodule more accurately as the convolution was performed at flexible locations instead of fixed locations. These locations were obtained using trainable offsets. This flexibility resulted in improved segmentation accuracy and clearly defined nodule boundaries with correct retention of the nodule shape. The average value

of segmentation, an IoU of 0.9035, obtained with the proposed segmentation algorithm is nearly 5% higher than that given by the U-net++ algorithm. The work also uses GPR interpolation resulting in better volume estimation accuracy. It proved to be the best interpolation method for irregularly shaped cancer nodules leading to a final volume estimation accuracy of 93.13%. The fully automatic lung cancer volume estimation system reported here not only removes the need for radiologist intervention and thereby the resulting subjectivity but also improves the overall accuracy of lung cancer detection, segmentation, and volume estimation, leading to an improved cancer grading.

**Table 2.** Volume estimations (in cubic cm) for various interpolation methods

Nodule volume estimation / Type of interpolation	Ground truth volume	Volume without interpolation	Quadratic	Cubic	Spine	Lagrange	GPR	Volume estimation accuracy in % for GPR
VP1	5.79	6.53	6.37	6.36	6.39	6.35	6.192	93.06
VP2	4.49	4.93	4.13	4.09	4.12	4.18	4.21	93.77
VP3	5.14	5.97	5.75	5.71	5.64	5.57	5.43	94.36
An average of 32 patients	5.17	5.82	5.78	5.81	5.69	5.63	5.52	93.13

**Table 3.** Comparison of volume estimations (in cubic cm) for U-net++ and proposed algorithm

Nodule volume estimation	Ground truth volume	Volume: GPR interpolation on U-net++	Volume accuracy in % for GPR on U-net++	Volume: GPR interpolation on the proposed algorithm	Volume accuracy in % for GPR on the proposed algorithm
VP1	5.79	6.422	89.08	6.192	93.06
VP2	4.49	4.039	89.95	4.21	93.77
VP3	5.14	5.62	90.66	5.43	94.36
An average of 32 patients	5.17	5.66	90.51	5.52	93.13

## 6. REFERENCES

- [1] M. Daykin, M. Sellathurai, I. Poole, "Comparison of unsupervised abnormality detection methods for interstitial lung disease", *Communications in Computer and Information Science*, Vol. 894, 2018, pp. 287-298.
- [2] I. Irigoien, B. Sierra, C. Arenas, "Towards application of one-class classification methods to medical data", *The Scientific World Journal*, Vol. 2014, 2014.
- [3] L. Tarassenko, P. Hayton, N. Cerneaz, M. Brady, "Novelty detection for the identification of masses in mammograms", *Proceedings of the Fourth International Conference on Artificial Neural Networks*, Cambridge, UK, 26-28 June 1995 pp. 442-447.
- [4] P. Ardimento, L. Aversano, M. L. Bernardi, M. Cimitile, "Deep Neural Networks Ensemble for Lung Nodule Detection on Chest CT Scans", *Proceedings of the International Joint Conference on Neural Networks*, Shenzhen, China, 18-22 July 2021.
- [5] M. Muzammil, I. Ali, I. U. Haq, A. A. Khaliq, S. Abdullah, "Pulmonary nodule classification using feature and ensemble learning-based fusion techniques", *IEEE Access*, Vol. 9, 2021, pp. 113415-113427.
- [6] I. Ali, M. Muzammil, I. U. Haq, A. A. Khaliq, S. Abdullah, "Deep Feature Selection and Decision Level Fusion for Lungs Nodule Classification", *IEEE Access*, Vol. 9, 2021, pp. 18962-18973.
- [7] D. Sethi, K. Arora, S. Susan, "Transfer Learning by Deep Tuning of Pre-trained Networks for Pulmonary Nodule Detection", *Proceedings of the IEEE 15th International Conference on Industrial and Information Systems*, India, 26-28 November 2020, pp. 168-173.
- [8] R. Tekade, K. Rajeswari, "Lung Cancer Detection and Classification Using Deep Learning", *Proceedings of the Fourth International Conference on Computing Communication Control and Automation*, Pune, India, 16-18 July 2018.
- [9] M. S. Rahman, P. C. Shill, Z. Homayra, "A New Method for Lung Nodule Detection Using Deep Neural

- Networks for CT Images”, Proceedings of the International Conference on Electrical, Computer and Communication Engineering, Cox'sBazar, Bangladesh, 7-9 February 2019.
- [10] N. F. Noaman, B. M. Kanber, A. Al Smadi, L. Jiao, M. K. Alsmadi, “Advancing Oncology Diagnostics: AI-Enabled Early Detection of Lung Cancer Through Hybrid Histological Image Analysis”, IEEE Access, Vol. 12, 2024, pp. 64396-64415.
- [11] T. I. A. Mohamed, A. El-Shamir Ezugwu, “Enhancing Lung Cancer Classification and Prediction with Deep Learning and Multi-Omics Data”, IEEE Access, Vol. 12, 2024, pp. 59880-59892.
- [12] S. K. Fazilov, O. R. Yusupov, K. S. Abdiyeva, “Mammography image segmentation in breast cancer identification using the OTSU method”, Web of Scientist: International Scientific Research Journal, Vol. 3, No. 8, 2022, pp. 196-205.
- [13] K. V. Shiny, N. Sugitha, “Automatic brain tumor segmentation on pre-operative MRI using region growing algorithm”, Information and Communication Technology for Competitive Strategies, Springer, 2022, pp. 653-665.
- [14] C. J. J. Sheela, G. Suganthi, “Morphological edge detection and brain tumor segmentation in magnetic resonance (MR) images based on region growing and performance evaluation of modified fuzzy C-means (FCM) algorithm”, Multimedia Tools and Applications, Vol. 79, No. 25-26, 2020, pp. 17483-17496.
- [15] H. Deng, J. P. Fitts, C. A. Peters, “Quantifying fracture geometry with X-ray tomography: Technique of iterative local thresholding (TILT) for 3D image segmentation”, Computers & Geosciences, Vol. 20, No. 1, 2016, pp. 231-244.
- [16] Y. Akbari, H. Hassen, S. Al-Maadeed, S. M. Zughair, “COVID-19 lesion segmentation using lung CT scan images: Comparative study based on active contour models”, Applied Sciences, Vol. 11, No. 17, 2021, p. 8039.
- [17] P. M. Bruntha, D. J. Rose, A. T. Shruthi, K. G. Juliet, M. Kanimozhi, “Application of selective region growing algorithm in lung nodule segmentation”, Proceedings of the 4<sup>th</sup> International Conference on Devices, Circuits and Systems, 16-17 March 2018, pp. 319-322.
- [18] P. B. Sangamithraa, S. Govindaraju, “Lung tumor detection and classification using EK-mean clustering”, Proceedings of the International Conference on Wireless Communications, Signal Processing and Networking, 23-25 March 2016, pp. 2201-2206.
- [19] M. Keshani, Z. Azimifar, F. Tajeripour, R. Boostani, “Lung nodule segmentation and recognition using SVM classifier and active contour modeling: A complete intelligent system”, Computers in Biology and Medicine, Vol. 43, No. 4, 2013, pp. 287-300.
- [20] W. J. Kostis, A. P. Reeves, D. F. Yankelevitz, C. I. Henschke, “Threedimensional segmentation and growth-rate estimation of small pulmonary nodules in helical CT images”, IEEE Transactions on Medical Imaging, Vol. 22, No. 10, 2003, pp. 1259-1274.
- [21] X. Hongtao, D. Yang, N. Sun, Z. Chen, Y. Zhang, “Automated pulmonary nodule detection in CT images using deep convolutional neural networks”, Pattern Recognition, Vol. 85, 2019, pp. 109-119.
- [22] J. Long, E. Shelhamer, T. Darrell, “Fully convolutional networks for semantic segmentation”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7-12 June 2015, pp. 3431-3440.
- [23] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional networks for biomedical image segmentation”, Proceedings of the Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5-9 October 2015, pp. 234-241.
- [24] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, “Rethinking atrous convolution for semantic image segmentation”, arXiv:1706.05587, 2017.
- [25] X. Luo, T. Song, G. Wang, J. Chen, Y. Chen, K. Li, D. N. Metaxas, S. Zhang, “SCPM-Net: An anchor-free 3D lung nodule detection network using sphere representation and center points matching”, Medical Image Analysis, Vol. 75, 2022, p. 102287.
- [26] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, J. Liu, “Ce-Net: Context encoder network for 2D medical image segmentation”, IEEE Transactions on Medical Imaging, Vol. 38, No. 10, 2019, pp. 2281-2292.

- [27] S. Bose, R. S. Chowdhury, R. Das, U. Maulik, "Dense dilated deep multiscale supervised U-network for biomedical image segmentation", *Computers in Biology and Medicine*, Vol. 143, 2022, p. 105274.
- [28] Q. Dou, L. Yu, H. Chen, Y. Jin, X. Yang, J. Qin, P.-A. Heng, "3D deeply supervised network for automated segmentation of volumetric medical images", *Medical Image Analysis*, Vol. 41, 2017, pp. 40-54.
- [29] F. Milletari, N. Navab, S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation", *Proceedings of the Fourth International Conference on 3D Vision*, Stanford, CA, USA, 25-28 October 2016, pp. 565-571.
- [30] K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA, 27-30 June 2016, pp. 770-778.
- [31] Y. Li, Y. Fan, "DeepSEED: 3D Squeeze-and-excitation encoder-decoder convolutional neural networks for pulmonary nodule detection", *Proceedings of the IEEE 17<sup>th</sup> International Symposium on Biomedical Imaging*, 3-7 April 2020, pp. 1866-1869.
- [32] J. Deng, J. Guo, N. Xue, S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition", *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, 15-20 June 2019, pp. 4690-4699.
- [33] S. G. Armato III et al. "The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans", *Medical Physics*, Vol. 38, No. 2, 2011, pp. 915-931.
- [34] A. A. A. Setio et al. "Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge", *Medical Image Analysis*, Vol. 42, 2017, pp. 1-13.
- [35] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation", *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, 2018, pp. 3-11.
- [36] J. Hou, C. Yan, R. Li, Q. Huang, X. Fan, F. Lin, "Lung Nodule Segmentation Algorithm with SMR-Unet", *IEEE Access*, Vol. 11, 2023, pp. 34319-34331.
- [37] W. Chen et al. "CT Lung Nodule Segmentation: A Comparative Study of Data Preprocessing and Deep Learning Models", *IEEE Access*, Vol. 11, 2023, pp. 94925-94931.
- [38] W. Jiang et al. "A Dual-Branch Framework With Prior Knowledge for Precise Segmentation of Lung Nodules in Challenging CT Scans", *IEEE Journal of Biomedical and Health Informatics*, Vol. 28, 2024, pp. 1540-1551.
- [39] F. Liao, M. Liang, Z. Li, X. Hu, S. Song, "Evaluate the Malignancy of Pulmonary Nodules Using the 3-D Deep Leaky Noisy-OR Network", *IEEE Transactions on Neural Networks and Learning Systems*, Vol. 30, No. 11, 2019, pp. 3484-3495.
- [40] C. Shorten, T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning", *Journal of Big Data*, Vol. 6, No. 1, 2019, p. 60.
- [41] J. Dai et al. "Deformable convolutional networks", *Microsoft Research Area*, June 2017.
- [42] X. Huang, B. Pang, T. Zhang, G. Jia, Y. Wang, Y. Li, "Improved Prostate Biparameter Magnetic Resonance Image Segmentation based on Def-Unet", *IEEE Access*, Vol. 11, 2023, pp. 43089-43100.
- [43] C. E. Rasmussen, C. K. I. Williams, "Gaussian Processes for Machine Learning", *The MIT Press*, 2006.





# Automated Surgical Wound Classification for Intelligent Emergency Care Applications

Original Scientific Paper

## Min Woo Park

Incheon National University  
Department of Computer Science and Engineering  
119, Academy-ro, Yeonsu-gu, Incheon, Republic of Korea  
yj951223@inu.ac.kr

## Mee Young Sung\*

Professor, Incheon National University  
Department of Computer Science and Engineering  
119, Academy-ro, Yeonsu-gu, Incheon, Republic of Korea  
mysung@inu.ac.kr

\*Corresponding author

**Abstract** – Developing AI-powered solutions for emergency care is critical to improving patient outcomes. This study presents an advanced AI model designed to accurately classify surgical wounds to facilitate prompt and appropriate emergency response. We used two state-of-the-art image classification models, ResNeXt and Vision Transformer (ViT), to evaluate their effectiveness in classifying wound images. These models were selected based on preliminary evaluations of several popular models, and their superior performance metrics justified their final selection. The models were trained on a combined dataset of approximately 1,000 images from the Medetec and AZH(Advancing the Zenith of Healthcare) datasets. To improve classification accuracy, the dataset was preprocessed, including image resizing, flipping, normalization, and augmentation. The ViT model demonstrated superior performance, achieving an accuracy of 92.78%, precision of 94.89%, recall of 91.87%, and an F1 score of 92.44%. These results surpass those of existing studies such as Shenoy et al. (85.1% accuracy), Rostami et al. (68.7% accuracy), and Gao et al. (76.6% accuracy). Our proposed AI system not only accelerates life-saving first aid by providing timely and accurate wound classification but also enhances the skills of emergency responders through continuous learning and real-time feedback. By integrating AI into emergency service protocols, we aim to improve response times and collaboration among medical personnel, ultimately contributing to better patient survival rates.

---

**Keywords:** surgical, wound, images, classification, emergency care, Intelligent, AI, Assistant

---

Received: May 7, 2024; Received in revised form: July 4, 2024; Accepted: July 4, 2024

## 1. INTRODUCTION

Wounds are a prevalent medical condition, and the severity and optimal treatment depend on the type of wound. Accurately identifying wound types and administering appropriate first aid in emergencies is critical for ensuring patient survival. However, current research on wound classification has primarily focused on assessing wound healing [1-4] or diagnosing diseases, such as diabetes, ulcers, and cancer [5-8], leaving a significant gap in wound classification for effective emergency response and treatment.

Machine learning and deep learning techniques in medical fields have increased, but wound classification has been neglected. Identifying the type of wound and providing the appropriate treatment quickly and ef-

fectively can significantly impact a patient's recovery. Therefore, developing AI-powered wound classification systems that can quickly and precisely assess wounds and recommend the most effective treatment is crucial.

In this study, we employed two wound image classification models: ResNeXt [9] and ViT (Vision Transformer) [10]. ResNeXt secured second place in the ILSVRC 2016 competition and has since been widely adopted for vision applications. ViT, on the other hand, has recently gained significant attention. Most of these foundation models can trace their roots back to the Transformer [11], a large language model developed by Google in 2017 with its source code made publicly available. Scaling up this model, which is specialized in natural language processing, and pre-training it with multimodal data, including text and code, has demonstrated supe-

rior performance compared to models trained solely on single-modality data. In visual recognition, reports suggest that ViT outperforms models derived from the widely used CNN family. This study contrasts the wound image classification capabilities of both ResNeXt and ViT.

To achieve the goal of developing an accurate and effective wound classification system, collecting a substantial amount of diverse data is crucial. After the data collection, we preprocess the raw data into a format the model can use. This may include tasks such as image resizing, noise reduction, and image augmentation to increase the variability of the data and improve model performance. We then generate labels based on the wound characteristics to enable the model to classify new images accurately. The optimization process involves training and fine-tuning the model to achieve the highest possible accuracy in wound classification. By employing these steps, we can create a robust and reliable system for automated wound classification that can be applied in various medical settings.

Our study aims to improve public health and medical welfare by developing an AI model for surgical wound classification that utilizes artificial intelligence technology to advise administering first aid based on the immediate and professional judgment of images of patients' wounds that occur in emergencies or daily life. The proposed AI model has the potential to significantly improve the accuracy and number of wound classifications and first aid recommendations, thereby enhancing public health and medical welfare.

Ultimately, we expect our work to contribute to the following:

- Improve public health and medical welfare by developing an AI model for surgical wound classification and an AI Assistant.
- Demonstration of the potential of AI technology to enhance the first aid process.
- Exploration of the potential for developing a virtual reality first aid training system using the proposed wound classification system to improve first aid proficiency through continuous learning.

This paper is structured as follows: Chapter 2 reviews prior research. Chapter 3 explores wound types, emphasizing the significance of first aid and the golden time. Chapter 4 introduces the dataset and preprocessing methods and compares the performance of underlying models. Chapter 5 analyzes the findings, while Chapter 6 illustrates an example of an assistant application utilizing the model. Finally, Chapter 7 discusses the contributions and proposes future research directions.

## 2. RELATED WORKS

Among existing studies, "Deepwound: Automated Postoperative Wound Assessment and Surgical Site Surveillance through Convolutional Neural Networks"

[1] shares many similarities with our research. The study employs CNNs to classify various wound images captured by smartphones into nine categories: Wound (Acc: 82%), Surgical Site Infection (84%), Granulation Tissue (85%), Fibrinous Exudate (83%), Open Wound (83%), Drainage (72%), Steri Strips (97%), Staples (95%), and Sutures (85%). Improving performance is relatively straightforward because the study categorizes wound types with distinct characteristics. However, in our study, one of the nine wound types classified by Deepwound, namely "Surgical Site Infection," is further subdivided into 14 different types. Consequently, the recognition difficulty varies considerably.

Another study titled "Multiclass wound image classification using an ensemble deep CNN-based classifier" [7] conducted two and three-class classification experiments on four types of wound images: venous, diabetic, pressure, and surgical. They employed two classifiers - per-patch and per-image - based on an ensemble DCNN. The results showed a maximum classification accuracy of 96.4% and an average of 94.28% for binary classification. The maximum and average accuracies for the three-class classification problems were 91.9% and 87.7%, respectively. Notably, this study has the potential to expand the scope of wound classification by accurately distinguishing 14 types of surgical wounds. Such distinctions can offer significant contributions to emergency medicine and first aid management.

## 3. BACKGROUND

In this section, we will discuss the medical classification of wounds and the crucial role of surgery in emergencies. Additionally, we will provide an overview of the current state of image recognition technology and its potential application in wound classification.

### 3.1. WOUND TYPES

A wound is an injury that causes a break in the skin or other bodily tissue and can be classified as either open or closed based on whether or not the skin's surface is broken [12]. Open wounds, which are the focus of this study, involve broken skin and exposed bodily tissue and can be further categorized based on their cause, shape, and level of contamination. Specifically, this study only considers open wounds that require surgical intervention. Table 1 provides a summary of the different types of open wounds.

**Table 1.** Open Wound Types

Open Wound Types	Wounds
Penetrating wounds	• Puncture wounds
	• Surgical wounds and incisions
	• Thermal, chemical, or electric burns
	• Bites and stings
Blunt force trauma	• Gunshot wounds or other high-velocity
	• Abrasions
	• Lacerations
	• Skin tears

Minor injuries can be managed using rudimentary first-aid procedures, including cleansing, hemostasis, and disinfection. Nonetheless, some circumstances necessitate hospitalization, such as penetrating wounds, amputations, and serious burns, and wounds with indeterminate extent and depth may impede decision-making. An AI system trained with appropriate data to identify appropriate first aid strategies can facilitate informed decision-making and promote timely wound care, enhancing the healing process and quality of life.

### 3.2. FIRST AID AND SURGERY

When an accident requires emergency surgery, timely and thorough treatment is crucial to success. In such situations, proper first aid can have an average success rate of 80%, with a success rate of 95% for clean amputations.

Surgical procedures are a form of life insurance that can be used in urgent situations where every minute counts. As is well known, if no action is taken within four minutes of a heart attack, it can mean death. In this way, the life and death of an injured person can depend on the quick and accurate actions of the first responders. Of course, not all diseases and injuries require surgery. Because it's not common to face these situations, most people don't know how to handle emergencies. Therefore, developing an application that recommends surgical procedures through wound image recognition is essential to make it simpler and more accurate to know which surgical procedures are needed for which diseases and wounds.

### 3.3. IMPORTANCE OF GOLDEN TIME AND HOSPITAL TREATMENT

The "Golden Time" refers to the critical period immediately following an accident when timely medical intervention, such as surgery, can mean the difference between life and death. This period is typically considered to be the first hour after a patient is critically injured. For instance, in the case of an adult who experiences cardiac arrest, surgical intervention within four minutes of the event can restore most of the central nervous system functions that existed before cardiac arrest and greatly reduce the likelihood of permanent disability [13]. Countries like the United Kingdom, the United States, Canada, Australia, and New Zealand place great value on the Golden Hour and have well-organized emergency medical systems that integrate fire and emergency rescue with requests to the police department. When an ambulance is called, fire trucks and police vehicles respond jointly. In Korea, integrating 112 and 119 services highlights the importance of the Golden Time.

A hospital visit is an essential step in receiving medical treatment. Unfortunately, many people tend to downplay their symptoms and avoid seeking medical attention, potentially worsening their condition. Helping individuals understand when to visit a doctor is critical in preventing delayed treatment and serious

health consequences. Proper guidance can encourage individuals to seek timely treatment, avoid misjudging the severity of their symptoms, and ultimately receive proper care for their condition.

### 3.4. WOUND IMAGES ANALYSIS

Recent research suggests that electronic devices such as smartphones can collect wound data, which can aid in situational analysis and decision-making during emergencies [14]. In addition, the field of surgical data science, which employs convolutional neural networks (CNN) and transfer learning models, has been introduced to recognize wound tissue images, making it possible to study wound assessment [15-17]. The recognition of wound images is a prevalent method of digital image processing that involves analyzing the color of the skin surrounding the wound using measurements of saturation, chroma, and intensity [18]. This method digitally processes the image, with color changes occurring step-by-step as the wound changes, allowing for the establishment of a scale for surgical treatment [19]. Developing algorithms for detecting and classifying wound tissue types, which play a crucial role in wound diagnosis, is beneficial to studying AI-powered wound image recognition [20].

Convolutional neural networks have gained significant attention for solving medical tasks and research [1, 7, 15, 18, 21]. CNN-based approaches are considered state-of-the-art for intelligent wound diagnosis and assessment. These networks consist of pixel-to-pixel architectures that use a large set of image data with damaged features and tunable neural network parameters. Many pre-trained models, such as VGGNet [22], ResNet [23], ResNeXt [9], DenseNet [24], EfficientNet [25], and MobileNet [26], and ViT [11] exist. Among these transfer learning models, ResNeXt and ViT are chosen and mainly compared their performance for our application because they are efficient and commonly used across various applications.

Transfer learning has already proven to be a powerful technique in machine learning. It involves taking pre-trained models and fine-tuning them on new tasks or datasets, which saves time and computing resources while achieving better performance. However, there is still room for improvement, and advancements in transfer learning are expected in the future. One area of research is the development of more efficient and effective methods for transferring knowledge between domains. Current transfer learning techniques typically rely on a fixed set of pre-trained models, limiting their ability to adapt to new or emerging tasks. Future advancements may involve the creation of more flexible and adaptable transfer learning frameworks that can better handle domain shifts and enable faster adaptation to new tasks.

### 4. WOUND CLASSIFICATION MODELING

This section describes our wound classification model, covering the dataset, preprocessing, classification model, training, optimization, and experiments.



### 4.1. WOUND IMAGE DATASET

Obtaining meaningful data is critical to the success of AI and is a key element of our research. Unfortunately, copyright-free wound images are limited, so we primarily utilized the Medetec dataset [27], a publicly available dataset of 561 wound images of 14 different types.

To improve the classification accuracy in our experiments, we merged images of the same class ("leg ulcer-1", "leg ulcer-2") into one class called "leg ulcer." We excluded other classes whose classification was unclear. This reduced the number of classes from 14 to 11, and only 502 images were used in our experiments. Table 2 shows the class-wise sample count for the 502 images from the Medetec dataset utilized in this study.

**Table 2.** Configuration per Class for Medetec Dataset with 502 Samples

Labels	Number of Samples
Abdominal	13
Burns	19
Epidermolysis-bullosa	5
Extravasation	20
Foot-ulcers	48
Haemangioma	6
Leg-ulcers	134
Malignant	9
Meningitis	24
Orthopaedic	49
Pressure-ulcer	175

Additionally, we used another public dataset, the AZH (Advancing the Zenith of Healthcare) dataset [8], which includes 538-foot ulcer images of four different wound types: Diabetic, Venous, Pressure, and Surgical. Table 3 summarizes the distribution of samples per class for the 538 images used in this study's AZH dataset.

**Table 3.** Configuration per Class for AZH Dataset with 538

Labels	Number of Samples
Diabetic	154
Venous	156
Pressure	100
Surgical	128

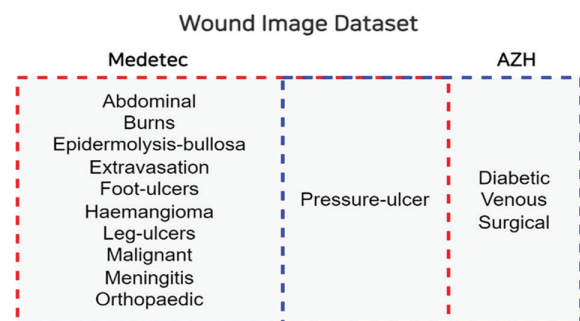
Sample images of Medetec are presented in Fig. 1. Fig. 2 shows sample images from the AZH Wound and Vascular Center database. Despite using these two datasets, they are still insufficient to achieve high classification accuracy. In this study, we used a combination of the Medetec and AZH datasets, and Fig. 3 shows the composition by class across the datasets used in this study. Fig. 3 shows the organization by class across the datasets used in this study.



**Fig. 1.** 12 classes of Medetec sample images



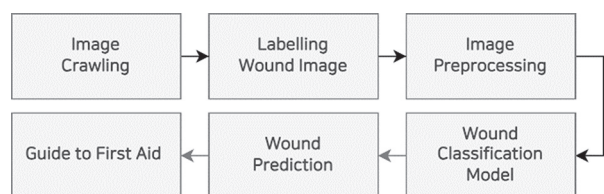
**Fig. 2.** 4 classes of AZH sample images



**Fig. 3.** Class organization of the dataset used in this study (Medetec + AZH)

### 4.2. WOUND IMAGE CLASSIFICATION PIPELINE

The first step in developing an AI application is to collect good data, which must then be preprocessed to improve performance before being inputted into the main classification model. Fig. 4 provides an overview of the wound image processing pipeline, which will ultimately predict the type of wound.



**Fig. 4.** Wound image classification pipeline



### 4.3. PREPROCESSING OF WOUND IMAGES

After collecting the wound data, the next step is to preprocess it for optimal performance before feeding it into the main classification model. Based on data analysis, preprocessing techniques can be applied to the training data to improve accuracy. The calibration of medical images is a crucial step in this process. The preprocessed training data is then classified into detailed attributes in a format suitable for the AI model. The Medetec dataset underwent preprocessing with the following steps:

- Resizing the entire image to 224x224
- Randomly flipping train images horizontally
- Randomly flipping train images vertically
- Normalizing all images

Details of the dataset preprocessing are presented in Fig. 5. The collected wound data was normalized into a format suitable for deep learning models for classification. The preprocessing involved dimensionality reduction and color correction of the wound image data. As a result, we were able to augment our dataset by a factor of 4 through data preprocessing.

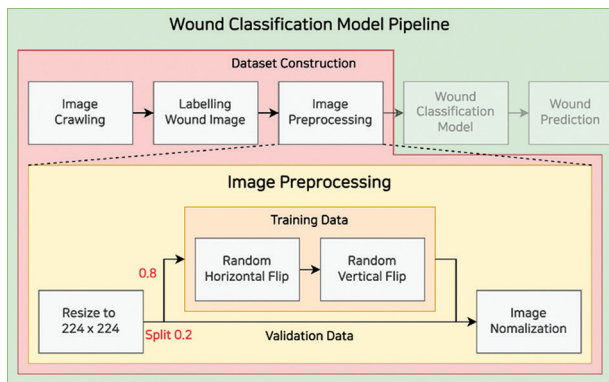


Fig. 5. Wound image preprocessing

### 4.4. WOUND IMAGE CLASSIFICATION MODELS

The preprocessing and optimization of data are important steps in the wound image classification pipeline. However, the most critical aspect is the classification model, which recognizes wounds and recommends optimal first aid. Deep neural network models are generally effective at extracting features from image data, and the final result label should correspond to one of the types listed in Fig. 3.

Before choosing the optimal classification model, we conducted a preliminary performance evaluation of several popular image classification models on the dataset used in this study. The results are presented in Table 4. For classification, we adopted two image classification models with the highest accuracy: the ResNeXt-101 32x8d architecture from “Aggregated Residual Transformation for Deep Neural Networks” [9] and the ViT architecture as the primary image classification model. In this paper, these two models will be analyzed from various perspectives.

Table 4. Accuracy by Model

Models	Accuracy
MobileNetV3-Large	81.73%
VGG19-BN	84.13%
EfficientNetV2-L	87.98%
ResNet152	87.98%
DenseNet161	88.46%
ResNeXt101-32x8d	90.38%
ViT-B/16	92.78%

#### 4.4.1. ResNeXt

ResNeXt-101 32x8d is a deep convolutional neural network architecture introduced in 2016 by Facebook AI Research (FAIR) researchers as an extension to the ResNet (Residual Network) architecture. The ResNeXt-101 32x8d architecture has 101 layers and a cardinality of 32, which refers to the number of parallel paths within each ResNeXt block. This allows the network to capture diverse features from the input images, making it more effective at image recognition tasks. The “8d” in the name of the architecture refers to the width of the bottleneck layers within each ResNeXt block, designed to reduce the number of parameters in the network and improve its computational efficiency. ResNeXt-101 32x8d has achieved state-of-the-art performance on several image recognition benchmarks, including the ImageNet dataset, which contains over 1 million images across 1,000 categories.

After preprocessing, the wound image data is input to the ResNeXt-101 32x8d model illustrated in Fig. 6. The schematic diagram of the proposed approach is presented in Fig. 7.

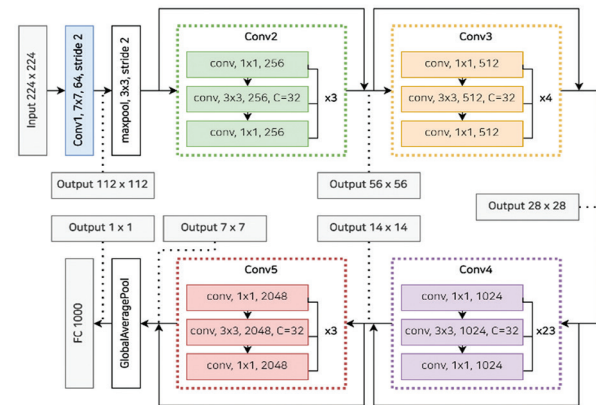


Fig. 6. ResNeXt-101 32x8d model architecture

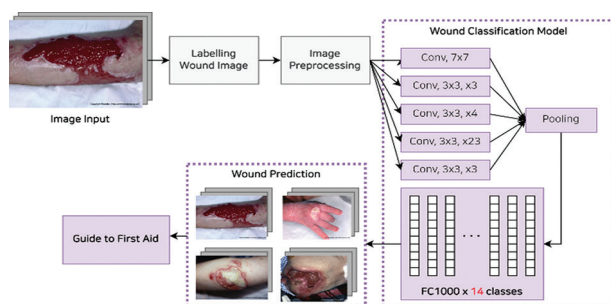
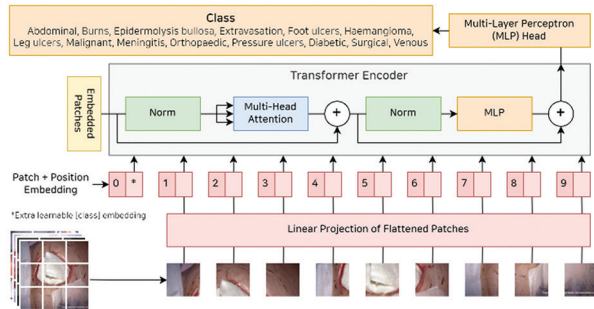


Fig. 7. Schematic diagram of the ResNeXt approach

#### 4.4.2. Vision Transformer (ViT)



**Fig. 8.** ViT model architecture

The Vision Transformer (ViT) is a novel approach that adapts the transformer architecture to the domain of image classification. Notably, ViT has demonstrated the ability to surpass the performance of conventional convolutional neural networks (CNN) and their variants while utilizing fewer computational resources. The ViT model architecture was unveiled in a research paper titled “An Image is Worth 16\*16 Words: Transformers for Image Recognition at Scale” [11]. The ViT model underwent pre-training on both the ImageNet and ImageNet-21k datasets. The recent success of transformers in NLP has paved the way for their promising performance in image classification tasks. Moreover, several transformer variants like CrossViT [28] and VIP-P [29] have been introduced, each optimized for specific applications. Within the medical realm, the ViT model has even outperformed the DenseNet model in diagnosing COVID-19 [30]. For processing, the ViT model divides an image into patches of 16x16x3. These patches are subsequently flattened and transformed into vectors of size 1x768, culminating in a patch embedding matrix with dimensions of 4x768. This matrix undergoes processing through 12 successive ViT Layers (transformer encoders) carrying location metadata. Each ViT Layer comprises a multi-head self-attention (MSA) block and a multi-layer perceptron (MLP) block.

#### 4.5. WOUND IMAGE CLASSIFICATION MODELS

To achieve optimal wound classification, it is essential to extract key features from wound images for accurate classification. This requires building an AI model that can effectively learn from wound images while preventing overfitting and over-parameterization during training and validation with real data. The optimization process then generates an optimal model that can accurately classify the cause of wounds for new wound image inputs.

During the model-building process, various components of the wound image classification model, such as hyperparameters and parameters, are adjusted and optimized to best fit the training data. Techniques such as removing the influence of similar images and preventing overfitting are applied to optimize the model. Through this process, the wound image classification

model is optimized for the training data and wound images and can accurately classify wound types.

We fine-tuned our wound image classification model using the combined dataset of Medetec and AZH and experimented with various hyperparameters to ensure optimal generalization and prevent overfitting. Table 5 displays the optimal hyperparameter values for the ResNeXt and ViT models.

**Table 5.** Hyperparameters for each model

Hyperparameters & Performance Metrics	ResNeXt101 -32x8d	Vision Transformer(ViT)
Batch size	3-100	8
Learning rate	0.01-0.00001	0.0001
Epoch	3-500	100

### 5. RESULTS

We used accuracy Eq.1, precision Eq.2, recall Eq.3, and F1 score Eq.4 for performance metrics. Table 6 shows the results for each performance metric based on the optimal hyperparameter values for both the ResNeXt and ViT models. The ViT model achieved an accuracy of 92.78%, with precision, recall, and F1 score values of 93.16%, 89.93%, and 90.79%, respectively. In contrast, the ResNeXt model yielded an accuracy of 90.38%, a precision of 94.61%, a recall of 84.10%, and an F1 score of 87.89%.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (4)$$

**Table 6.** Performance for each model

Models	ResNeXt101 -32x8d	Vision Transformer(ViT)
Accuracy	90.38%	92.78%
Precision	94.61%	94.89%
Recall	84.10%	91.87%
F1 score	87.89%	92.44%

Fig. 9 illustrates the evolution of each performance metric over 100 epochs. While ResNeXt exhibits instability across several metrics during its learning phase, the ViT model becomes increasingly stable as it learns. This stability suggests the potential for further improvement. Overall, ViT outperforms and is more consistent than ResNeXt across all metrics.

Fig. 10 and Fig. 11 show the confusion matrix of the predicted results versus the true results of the ResNeXt101-32x8d and ViT models. ViT shows equal or better accuracy for most labels compared to ResNeXt101\_32x8d. The labels Malignant, Abdominal,

and Burns, which have relatively few sample image data, show a lower accuracy of 50% for both models compared to the other labels. In terms of misclassifications, the two models are similar.

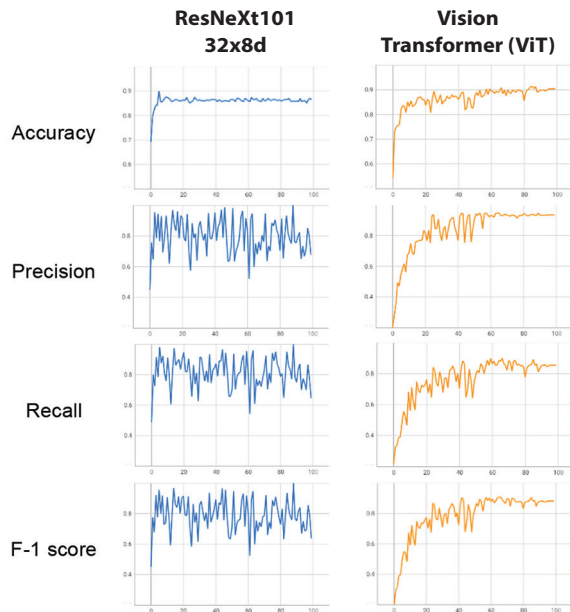


Fig. 9. Accuracy, Precision, Recall, and F1 score

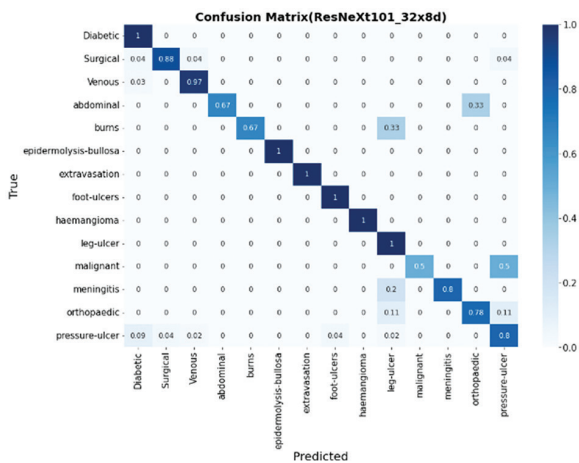


Fig. 10. Confusion Matrix of ResNeXt101-32x8d

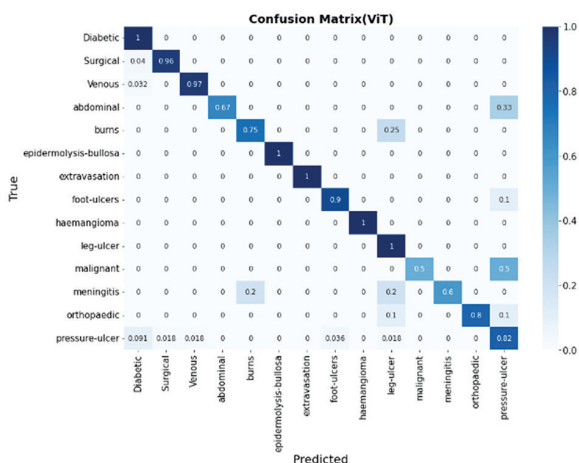


Fig. 11. Confusion Matrix of ViT

A comparison of the accuracy of our study with related studies is shown in Table 7. V. N. Shenoy et al. [1] classified wound types and wound treatment methods using images and showed an average accuracy of 85.1%. B. Rostami et al. [7] used the AZH and Milwaukee Vascular Center datasets, which were also used in this study. The study showed an accuracy of 68.7% for classifying six classes. X. Gao et al. [30] used ViT to classify COVID-19-infected and uninfected individuals and achieved an accuracy of 76.6%. Our proposed model demonstrated a high accuracy of 92.8% in classifying 14 classes, with the best performance in classifying wounds by type.

Table 7. Accuracy versus previous studies



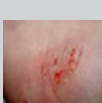
Paper	V. N. Shenoy et al.	B. Rostami et al.	X. Gao et al.	Our Paper
Accuracy	85.1%	68.7%	76.7%	92.8%

## 6. AI ASSISTANT FOR FIRST AID

There are various criteria for categorizing wounds and their corresponding labels. Wounds can be categorized based on their cause, shape, or presence/absence of infection. For instance, wounds can be classified based on their cause, such as cut, stab, laceration, scab, thermal, and shooting. Based on their shape, wounds can be categorized into linear, plate, and missing wounds. However, these criteria may not be easily understood. Therefore, in this study, wound classification uses commonly used and easily understood terms by the average person.

Table 8. illustrates seven types of wounds for first aid based on common criteria and terminology, along with brief summaries of first-aid treatment options for each type of wound.

Table 8. Seven Types of Wounds and Their Treatment Options

Types	Photos	Definition	First Aid
Contusions		Vascularized wounds caused by physical impact (edema formation, bruising)	Apply cold compresses to reduce swelling before applying warm compresses.
Burns		If the main cause of the wound is a hot object	For first-degree burns, soothe with cold water. For second and third-degree burns, cool the burned area with cold water, wrap it in gauze, and go to the hospital without forcing the blisters to pop.
Abrasions		Wounds in which the epidermis or part of the dermis has been removed (scratches)	Apply an antibiotic ointment along with simple wound cleaning.



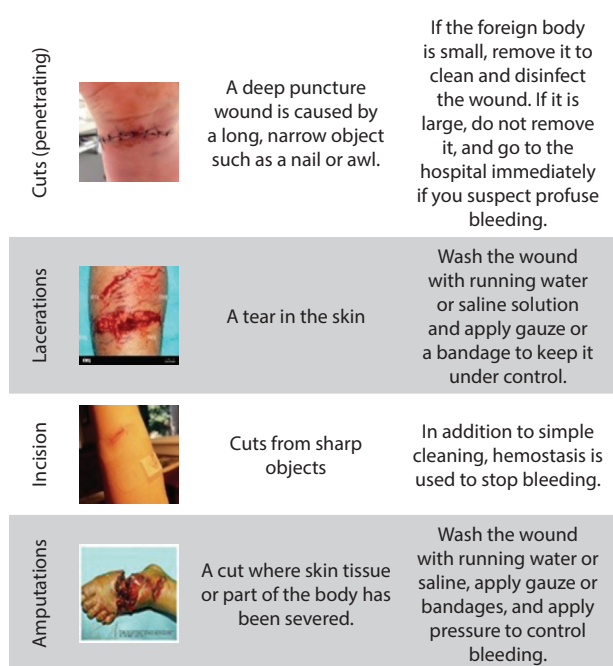


Fig. 12 shows a simple user interface for a prototyped first aid assistant. Here are some scenarios where the public can use First Aid Assistant Scenario:



Fig. 12. Application Example for First Aid Assistant

1. Accident scene: A person gets injured.
2. Utilize the AI assistant app: A bystander launches the AI assistant app on their phone and takes an image of the wound.
3. Provide first aid guidance: The app analyzes the wound image to determine the depth, type, and danger of a wound and provides immediate first aid instructions (e.g., instructions to apply pressure or elevate the wound).
4. Send information to paramedics: The app sends information about the wound and location data to the nearest paramedic.
5. Hospital referral: At the same time, the app sends patient information and the severity of the wound to the emergency room of the nearest hospital.

6. Interaction with specialists: Specialists at the hospital can view wound images directly through the app and send information to paramedics for further action or medication.
7. Patient arrival and first aid: Paramedics arrive at the scene, provide the patient with hands-on first aid based on the first aid guide provided in the app, and transport the patient to the hospital.
8. Hospital arrival: When the patient arrives at the hospital, specialists already know the patient's condition and can respond quickly.

## 7. CONCLUSION

This paper presents a study on developing an AI-powered first-aid assistant designed to classify surgical wounds and recommend appropriate treatments intelligently. We emphasize the importance of precise wound classification in emergencies and highlight the benefits of utilizing state-of-the-art AI techniques to enhance the first aid process, potentially saving lives.

The surgical wound classification system proposed here forms the core of the AI-powered first aid assistant. This system leverages recent deep-learning models to classify wound images and related data, suggesting the most fitting first-aid treatment based on factors like wound type, skin thickness, tension, and scarring. Such an approach minimizes errors, increases safety, and is time-efficient.

The key findings of this study include:

- Despite being trained on a limited dataset of roughly 1,000 images, the wound image classification using ViT models achieved an accuracy of 92.78%. The precision, recall, and F1 score values were 94.89%, 91.87%, and 92.44%, respectively, underscoring the model's potential for deployment in first-aid scenarios.
- To enhance the wound classification model, acquiring wound images from actual first aid situations is essential, highlighting the need for collaboration with medical institutions to collect such images.
- The proposed AI model can be applied beyond first aid to various medical treatments. A first aid assistant system based on this model can empower individuals with knowledge of effective first aid, potentially saving lives. It can also promote collaborative efforts, transmitting the first aid process to paramedics in real-time and thus shortening response times. The system opens up avenues for remote first aid, collaboration with paramedics, and consolidated treatment that merges diagnosis and prescription within one platform.

In this study, we tried to augment the data with preprocessing due to the small number of images in the original dataset. However, this was limited by lower accuracy for some labels. This suggests that more images of real wounds should be added to the dataset in the future.



Future research will concentrate on enhancing data augmentation techniques to boost the performance of surgical wound classification and ensure meticulous performance evaluation. We also aim to design AI systems for remote first aid, collaboration with paramedics, and real-time integration with specialist diagnostic and prescription platforms.

Another next step for this research is to develop an AI assistant for first aid that incorporates augmented reality (AR) technology. Previously, the research team conducted various haptic virtual simulation studies for medical education and training, such as suture surgery simulation [31-35] and the haptic visual discrimination test (HVDT) [36]. With the growing interest in mixed-reality technology, transitioning these simulations to a mixed-reality environment could enhance their utility. Combining this paper's research with virtual simulations is expected to enable AI surgical assistants to identify wounds in real time, suggest appropriate first-aid strategies, and guide the surgical process. This approach will likely accelerate learning and promote mastery of optimal first-aid techniques.

## 8. ACKNOWLEDGEMENT

This work was supported by the Incheon National University Research Grant in 2020 (No. 2020-0271).

## 9. REFERENCES:

- [1] V. N. Shenoy, E. Foster, L. Aalami, B. Majeed, O. Aalami, "Deepwound: Automated Postoperative Wound Assessment and Surgical Site Surveillance through Convolutional Neural Networks", Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine, Madrid, Spain, 3-6 December 2018, pp. 1017-1021.
- [2] S. Yang, J. Park, H. Lee, S. Kim, B. U. Lee, Y. Chung, B. Oh, "Sequential Change of Wound Calculated by Image Analysis Using a Color Patch Method during a Secondary Intention Healing", Plos One, Vol. 11, No. 9, 2016, p. e0163092.
- [3] C. L. Su, C. C. Chang, Y. S. Peng, M. Y. Chen, "The Predictive Factors Associated with Comorbidities for Treatment Response in Outpatients With King Classification III Diabetes Foot Ulcers", Annals of Plastic Surgery, Vol. 81, No. 6S, 2018, pp. S39-S43.
- [4] C. Chakraborty, "Computational approach for chronic wound tissue characterization", Informatics in Medicine Unlocked, Vol. 17, 2019.
- [5] L. Wang, P. C. Pedersen, D. M. Strong, B. Tulu, E. Agu, R. Ignatz, "Smartphone-Based Wound Assessment System for Patients with Diabetes", IEEE Transactions on Biomedical Engineering, Vol. 62, No. 2, 2015, pp. 477-488.
- [6] A. Heras-Tang, D. Valdes-Santiago, A. M. Leon-Mecias, M. L. B. Díaz-Romañach, J. A. Mesejo-Chiong, C. Cabal-Mirabal, "Diabetic foot ulcer segmentation using logistic regression, DBSCAN clustering and mathematical morphology operators", Electronic Letters on Computer Vision and Image Analysis, Vol. 21, No. 2, 2022, pp. 22-39.
- [7] B. Rostami, D. M. Anisuzzaman, C. Wang, S. Gopalakrishnan, J. Niezgodá, Z. Yu, "Multiclass wound image classification using an ensemble deep CNN-based classifier", Computers in Biology and Medicine, Vol. 134, 2021.
- [8] S. R. Oota, V. Rowtula, S. Mohammed, M. Liu, M. Gupta, "WSNet: Towards an Effective Method for Wound Image Segmentation", Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Waikoloa, HI, USA, 2-7 January 2023, pp. 3234-3243
- [9] S. Xie, R. Girshick, P. Dollar, Z. Tu, K. He, "Aggregated Residual Transformations for Deep Neural Networks", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, pp. 1492-1500.
- [10] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, "Attention is all you need", Advances in Neural Information Processing Systems, Vol. 30, 2017.
- [11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N. Houlsby, "An image is worth 16x16 words: Transformers for image recognition at scale", arXiv:2010.11929, 2020.
- [12] Types of Wounds, <https://www.kindredhospitals.com/our-services/ltac/conditions/wound-care> (accessed: 2023)
- [13] A. S. Peterson, "The "Golden Period" For Wound Repair", The Journal of Lancaster General Hospital, Vol. 5, No. 4, 2010, pp. 134-135.
- [14] M. L. Tolins, D. S. Hippe, S. C. Morse, H. L. Evans, W. B. Lober, M. C. Vrablik, "Wound Care Follow-Up From the Emergency Department Using a Mobile Application: A Pilot Study", The Journal of Emergency Medicine, Vol. 57, No. 5, 2019, pp. 629-636.
- [15] M. Elmogy, B. Garcia-Zapirain, C. Burns, A. Elmaghraby, A. Ei-Baz, "Tissues Classification for Pressure Ulcer Images Based on 3D Convolutional Neural Network", Proceedings of the 25<sup>th</sup> IEEE International Conference on Image Processing, Athens, Greece, 7-10 October 2018, pp. 3139-3143.

- [16] H. Nejati, H. A. Ghazijahani, M. Abdollahzadeh, T. Malekzadeh, N.-M. Cheung, K.-H. Lee, L.-L. Low, "Fine-Grained Wound Tissue Analysis Using Deep Neural Network", Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Calgary, AB, Canada, 15-20 April 2018, pp. 1010-1014.
- [17] C. Wang, D. M. Anisuzzaman, V. Williamson, M. K. Dhar, B. Rostami, J. Niezgodna, S. Gopalakrishnan, Z. Yu, "Fully automatic wound segmentation with deep convolutional neural networks", Scientific Reports, Vol. 10, No. 1, 2020.
- [18] H. Oduncu, A. Hoppe, M. Clark, R. J. Williams, K. G. Harding, "Analysis of Skin Wound Images Using Digital Color Image Processing: A Preliminary Communication", The International Journal of Lower Extremity Wounds. Vol. 3, No. 3, 2004, pp. 151-156.
- [19] M. F. A. Fauzi, I. Khansa, K. Catignani, G. Gordillo, C. K. Sen, M. N. Gurcan, "Computerized segmentation and measurement of chronic wound images", Computers in Biology and Medicine, Vol. 60, 2015, pp. 74-85.
- [20] F. J. Veredas, R. M. Luque-Baena, F. J. Martin-Santos, J. C. Morilla-Herrera, L. Morente, "Wound image evaluation with machine learning", Neurocomputing, Vol. 164, 2015, pp. 112-122.
- [21] H. Lu, B. Li, J. Zhu, Y. Li, Y. Li, X. Xu, L. He, X. Li, J. Li, S. Serikawa, "Wound intensity correction and segmentation with convolutional neural networks", Concurrency and Computation: Practice and Experience, Vol. 29, No. 6, 2017.
- [22] K. Simonyan, A. Zisserman. "Very deep convolutional networks for large-scale image recognition", arXiv:1409.1556, 2014.
- [23] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27-30 June 2016, pp. 770-778.
- [24] G. Huang, Z. Liu, L. van der Maaten, K. Q. Weinberger, "Densely Connected Convolutional Networks", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21-26 July 2017, pp. 4700-4708.
- [25] M. Tan, Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks", Proceedings of the 36<sup>th</sup> International Conference on Machine Learning, Vol. 97, Long Beach, CA, USA, 2019, pp. 6105-6114.
- [26] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications", arXiv:1704.04861, 2017.
- [27] S. Thomas, "Stock pictures and copyright free images of open wounds and surgical wound dressings", <http://www.medetec.co.uk/files/medetec-images.html> (accessed: 2023)
- [28] C. F. R. Chen, Q. Fan, R. Panda, "CrossViT: Cross-attention multi-scale vision transformer for image classification", Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10-17 October 2021, pp. 357-366.
- [29] H. Wang, Y. Ji, K. Song, M. Sun, P. Lv, T. Zhang, "ViT-P: Classification of Genitourinary Syndrome of Menopause From OCT Images Based on Vision Transformer Models", IEEE Transactions on Instrumentation and Measurement, Vol. 70, 2021, pp. 1-14.
- [30] X. Gao, Y. Qian, A. Gao, "COVID-ViT: Classification of COVID-19 from CT chest images based on vision transformer models", arXiv:2107.01682, 2021.
- [31] M. Y. Sung, B. Kang, J. Kim, T. Kim, H. Song, "Intelligent Haptic Virtual Simulation for Suture Surgery", International Journal of Advanced Computer Science and Applications, Vol. 11, No. 2, 2020.
- [32] T. Kim, C. Kim, H. Song, M. Y. Sung, "Intuition, Accuracy, and Immersiveness Analysis of 3D Visualization Methods for Haptic Virtual Reality", International Journal of Advanced Computer Science and Applications, Vol. 10, No. 11, 2019.
- [33] A. R. Choi, M. Y. Sung, "Performance improvement of haptic collision detection using subdivision surface and sphere clustering", Plos One, Vol. 12, No. 9, 2017, p. e0184334.
- [34] A. R. Choi, S. M. Kim, M. Y. Sung, "Controlling the contact levels of details for fast and precise haptic collision detection", Frontiers of Information Technology & Electronic Engineering, Vol. 18, No. 8, 2017, pp. 1117-1130.
- [35] A. R. Choi, C. W. Kim, M. Gwak, M. Y. Sung, "Haptic interactions for probing real objects in remote places", International Journal of Applied Engineering Research, Vol. 12, No. 24, 2017, pp. 14948-14954.
- [36] H. Y. Kim, M. Y. Sung, "Virtual Haptic Visual Discrimination Test", Journal of Telecommunication, Electronic and Computer Engineering, Vol. 10, No. 1, 2018, pp. 5-11.



# A Framework for 5G Network Slicing Optimization using 2-Edge-Connected Subgraphs for Path Protection

Original Scientific Paper

## Igor Begić\*

JP ELEKTROPRIVREDA HZ HB d. d. Mostar,  
Development Division, Department of Telecommunications  
Ulica kralja Petra Krešimira IV, 6-A, Mostar, Bosnia and Herzegovina  
igor.begic@ephzhh.ba

## Adrian Satja Kurdija

University of Zagreb,  
Faculty of Electrical Engineering and Computing,  
Department of Electronics, Microelectronics, Computer and Intelligent Systems  
Unska 3, Zagreb, Croatia  
adrian.kurdija@fer.hr

## Željko Ilić

University of Zagreb,  
Faculty of Electrical Engineering and Computing,  
Department of Telecommunications  
Unska 3, Zagreb, Croatia  
zeljko.ilic@fer.hr

\*Corresponding author

**Abstract** – Emerging telecommunications technologies require robust frameworks for efficient network slicing. We propose a network-slicing model that aims to optimize the deployment of virtual networks on a physical network topology. Our model ensures compliance with 5G requirements, incorporating latency and capacity constraints on virtual links. Selecting slices with cost and resource requirements on the computing nodes is optimized using a Knapsack problem with revenue maximization. We propose a path protection algorithm to deal with link failures by constructing a 2-edge-connected subgraph (or two link-disjoint Steiner trees) for each slice to provide both primary and backup paths. Simulation results include comparison with existing solutions by metrics such as latency, revenue, resource utilization, number of protected slices, and computation time, providing valuable insights for network planners operating in diverse and dynamic environments. Key contributions include efficient resource allocation using the Knapsack problem, enhanced network resilience via 2-edge-connected subgraphs for path protection, and realistic simulation experiments on SNDlib dataset topologies. The simulation results show that the proposed framework improves computational efficiency compared to the recent related solutions, particularly in large network topologies where  $k$ -connected function slicing (KC-FS) subgraph embeddings take approximately 3.5 times more computation time.

---

**Keywords:** network slicing, resource allocation, virtual networks, optimization model, 5G networks, path protection

---

Received: April 15, 2024; Received in revised form: July 4, 2024; Accepted: July 4, 2024

## 1. INTRODUCTION

5G technology has ushered in unprecedented challenges and opportunities in the rapidly evolving telecommunications landscape. As the demand for diverse and high-performance services intensifies, robust

frameworks to efficiently manage network resources become crucial. Network slicing, a pivotal concept in the realm of 5G, offers a promising avenue for achieving this efficiency by allowing the creation of virtual networks tailored to specific service requirements [1].



Ensuring compliance with 5G requirements, particularly regarding latency and capacity constraints on virtual links, becomes critical. The telecommunications industry is actively developing models and frameworks that address these challenges. Solutions that optimize the deployment of virtual networks contribute to the overall efficiency and effectiveness of 5G networks [2, 3].

As the demand for 5G services continues to escalate, network operators are confronted with strategically selecting and deploying network slices that comply with stringent cost and resource constraints on computing nodes and maximize revenue potential [4]. The cost-effective utilization of resources is pivotal to ensuring the economic viability of 5G networks, making optimization strategies a focal point in current research and development efforts [5].

Furthermore, in the 5G networks, the seamless and uninterrupted connectivity with this advanced technology promises is contingent upon robust mechanisms for handling network link failures. As 5G networks evolve to support diverse services (e.g., eMBB, uRLLC, and mMTC) and applications with stringent reliability requirements, the industry has recognized the imperative of implementing protection algorithms to safeguard against disruptions caused by link failures. The telecommunications industry has turned to solutions such as dedicated path protection algorithms to mitigate this impact. This approach involves constructing both primary and backup paths, ensuring that communication is seamlessly transmitted to the pre-established backup path in case of a link failure on the primary path [6, 7].

This paper proposes a network-slicing model designed to optimize the deployment of virtual networks on a physical network topology. Our model is crafted to ensure compliance with 5G requirements, addressing crucial aspects such as latency and capacity constraints on virtual links. The model integrates latency constraints on virtual links, ensuring that the constructed subgraphs and selected paths minimize latency, which is essential for meeting the low latency requirements of 5G real-time communications. It enhances network reliability by constructing 2-edge-connected subgraphs, ensuring that an alternative path remains available between any pair of nodes even if a single link fails, thus maintaining uninterrupted service.

By integrating cost and resource requirements considerations on computing nodes, we employ a Knapsack problem with revenue maximization to optimize the selection of slices, striking a balance between resource utilization and service quality. By formulating the slice selection problem as a Knapsack problem, the framework optimizes the allocation of resources (CPU units, memory, etc.) on computing nodes, maximizing revenue while adhering to the nodes' capacity constraints, thereby ensuring efficient use of available resources.

Recognizing the dynamic nature of telecommunications environments, our proposed model employs two variants of path protection. Our primary proposed approach for each slice constructs a subgraph with the 2-edge-connectivity property, which means that at least two distinct paths exist between each pair of slice's nodes. The alternative approach leverages the construction of two link-disjoint Steiner trees for each slice: one for primary paths and the second for backup paths. Both methods are tailored to handle link failures efficiently, providing a robust mechanism for ensuring continuous service availability and mitigating potential disruptions.

We comprehensively evaluated the proposed models through extensive simulations using network topologies from the SNDlib dataset [8]. Metrics such as cost, latency, accepted slices, and resource utilization are systematically analyzed to provide network planners with valuable insights into the performance and resilience of the network slicing framework. These insights are crucial in diverse and dynamic environments where telecommunication networks must adapt to varying demands and unexpected challenges.

This paper aims to contribute to the scientific discourse surrounding optimizing 5G networks, grappling with the complexity of modern and evolving network landscapes. By addressing deployment efficiency, compliance with stringent requirements, and resilience in the face of failures, our model seeks to advance state-of-the-art telecommunications technologies, offering tangible solutions for network planners.

The primary contributions of this paper encompass the following:

- The paper introduces a novel framework for efficient resource allocation and path protection in 5G network slicing deployment using the Knapsack problem and 2-edge-connected subgraphs.
- The paper shows the proposed framework's computational efficiency compared to the recent related solutions through extensive simulations with realistic network topologies.

The paper is organized as follows: Section II discusses the related works; Section III presents our system model and defines the problem in mathematical terms; Section IV presents the proposed algorithms; Section V presents simulation results. Finally, the conclusion of the paper is presented in Section VI.

## 2. RELATED WORK

Many papers considered 5G network slicing aspects, attempting to solve optimization problems based on the system model and the assumed constraints and requirements.

To name a few related approaches, the authors in [9] used the Multiple Choice Knapsack Problem to address the problem of maximizing the slice allocation in the

access network, i.e., a macro cell. In our paper, the 0-1 Knapsack Problem is adapted to find an optimal subset of network slices in each computing node, balancing the trade-off between revenue maximization and adherence to resource constraints.

Recently, the authors in [5] proposed a nested decomposition model for reliable slicing of Virtualized Network Functions (VNFs) running on virtual computing nodes, providing dedicated path protection with primary and backup paths being link-disjoint. Because of the enormous problem size, their solution used integer linear programming (ILP) with heuristic column generation. In [7], the authors employ a heuristic to solve a path-based ILP with bandwidth squeezing and multi-path provisioning. In another recent approach [6], latency-constrained paths between source and target nodes are considered and selected from the shortest to the longest in terms of hops.

In [10], the authors proposed a model that integrates computing and networking resources to minimize costs for the Slice Broker (SB) when purchasing resources from Infrastructure Providers (InPs). In [11], the authors introduce the Topology-Level Based Protection Scheme (TLPS) as an innovative method to improve reliability and minimize redundancy in multi-domain networks. TLPS customizes virtual networks (VNs) to meet precise reliability needs while reducing backup redundancy. It accomplishes this by crafting tailored VNs and deploying them onto the substrate network through a mixed integer linear programming (MILP) model and heuristic approach. In [12], the authors proposed an advanced approach for optimizing VNF allocation within dynamic network slices. By introducing a multilayered Service Function Chain (SFC) formation and utilizing an ILP model, the research addresses the VNF-embedding and allocation problem (VNF-EAP) on a real-world AT&T network topology.

In [13], the authors presented a hierarchical identifier (HID) 5G architecture that simultaneously supports network slicing and SFC. The HID incorporates network slice selection assistance information (NSSAI) and service path ID (SPI), allowing users to attach to specific network slices and enabling the sequential processing of flows by service functions (SFs). To mitigate degradation in quality of service (QoS) caused by unexpected increases in flow to a slice, a backup slice shared among different services is introduced to utilize network resources efficiently. In [14], the authors examined the reliability of 5G transport network slices within elastic optical networks (EON). The primary focus is slicing 5G transport networks, which involves setting up virtual networks on 5G transport infrastructure while ensuring dedicated protection.

Most existing methods solve the Virtual Network Embedding (VNE) problem by first embedding all nodes without considering link embedding information. However, such approaches can lead to suboptimal link embedding due to decisions made without explicitly considering the requirements and characteristics of

the links. This can lead to inefficient resource utilization and revenue loss for service providers.

Our approach differs from the conventional approaches that model each slice as a *chain* of virtual network functions (VNFs) implemented by the computing nodes (e.g., [5, 12]). Instead of assuming a fixed order of VNFs within a slice, our approach is more flexible: it only requires that all corresponding computing nodes are connected by paths, allowing for different orderings (chains) of VNFs within the slice. Therefore, instead of constructing separate paths for each pair of consecutive computing nodes in a slice, we build a 2-edge-connected subgraph spanning all computing nodes with minimal total latency of the used links. Path protection results from the 2-edge-connectivity property: the subgraph remains connected whenever an edge is removed, corresponding to a failed link. This approach is compared to our proposal [15], where two link-disjoint Steiner trees spanning all computing nodes are constructed, one for primary paths and the other for backup paths. It is also compared to an approach recently proposed in [16], where each service function (SF) is embedded as a *k-edge*-connected subgraph using the *k*-connected function slicing (KC-FS) algorithm. When  $k = 2$ , the KC-FS algorithm can be used for slice subgraph construction in the scenario presented here. However, the simulation results demonstrate that it is computationally much more demanding. Also, in [16] the authors do not consider selecting nodes for an SF: it assumes that the set of node instances ( $N_{f,i}$ ) for the  $i^{\text{th}}$  SF is given as part of the input. The two link-disjoint trees might not exist for network topologies with low redundancy, so the present approach is more suitable, as we will demonstrate. Another key difference from conventional approaches is that instead of using ILP formulations, which can be inefficient for large problem sizes, we use the classical algorithms for Knapsack and Steiner Tree problems, making the algorithm more straightforward to implement and guaranteeing its polynomial time complexity.

In [17], the authors highlighted the importance of virtualization in modern computing and its pervasive presence across various computing domains, emphasizing the need for efficient management of physical resources. The paper introduced the Virtual Network Embedding (VNE) problem, which entailed allocating physical resources to meet virtual resource requests while adhering to constraints and maximizing resource utilization. It presented the Improved Virtual Network Embedding using Conflict-Based Search (iVNE-CBS) as an improved algorithm for addressing the VNE problem.

In [18], the authors formulated the problem of Service Function Chain (SFC) scheduling in NFV-enabled 5G networks as a mixed integer non-linear programming challenge. They aimed to maximize the number of requests meeting latency and reliability constraints in a dynamic network environment where SFC requests arrive randomly. Additionally, the authors proposed an efficient algorithm to decide VNF redundancy while minimizing delay.

In [19], the authors tackled the challenge of optimally placing SFCs within 5G network slices, considering resource distribution across Edge and Cloud sites. They introduced solutions for the Virtual Network Functions Chain Placement Problem (VNF-CPP) using integer linear programming (ILP) and heuristic algorithms, addressing constraints like end-to-end delay, processing delay, VNF affinity, and traffic requirements between VNFs.

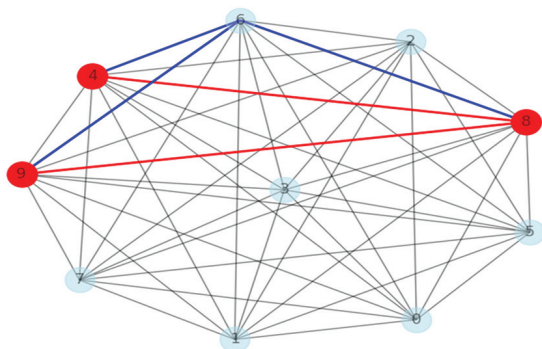
In [20], the authors proposed a virtual network embedding algorithm model incorporating a graph attention mechanism and a multi-layer perceptron. This model determines the weights between nodes by introducing the attention mechanism to measure the relationships between nodes. The proposed models and algorithms effectively reduced resource fragmentation and improved the acceptance rate of virtual network requests.

In [21], the authors employed a reinforcement learning (RL) approach to address the embedding problem. The approach integrated shareable virtual network functions into an existing RL scheme designed for virtual node embedding, achieving this with minimal additional computation.

### 3. FRAMEWORK OVERVIEW

In our framework, we focus on virtual computing nodes connected by links in the network topology. We optimize the slice selection in the computing nodes and the construction of slices' subgraphs in the network topology graph.

As an illustration of our Steiner-tree approach [15], Fig. 1 depicts the nodes selected for a specific slice (nodes 4, 8, and 9, in red color) along with the primary tree (red links) and the backup tree (blue links) connecting these selected nodes. Backup links can still connect each pair of the selected nodes in case of any primary link failure.

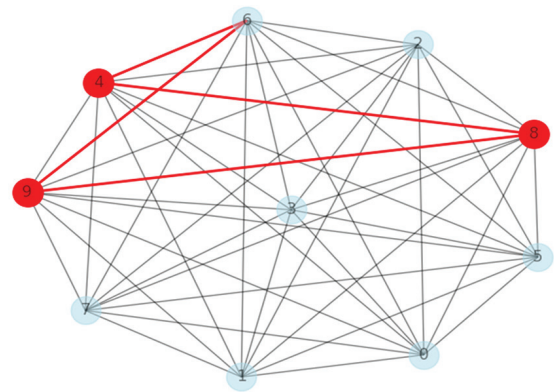


**Fig. 1.** Primary (red) and backup (blue) trees spanning the slice embedded in the nodes 4, 8, and 9

However, this paper aims for more efficient path protection. Instead of two spanning trees, one being able to replace the other, we will construct a single spanning subgraph, which is "larger" than one tree

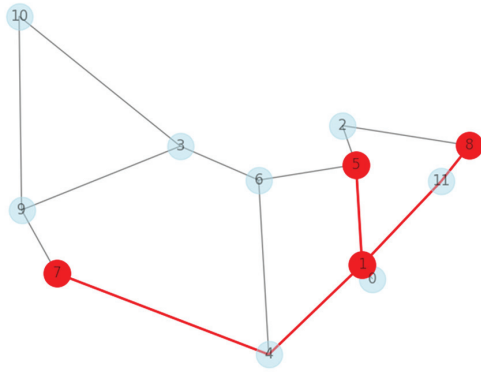
but "smaller" than two trees (in terms of the number of links). The subgraph must have the following property to allow for link failures: whenever a single edge from the subgraph is removed, there is still a path between any pair of its nodes. Therefore, the subgraph provides both primary and backup paths between the nodes. The required property of such a graph is called *2-edge-connectivity*.

To illustrate the difference between the two proposals, consider Fig. 1 and 2. In Fig. 1, the (red) primary and (blue) backup trees contain five edges. In Fig. 2, however, the single (red) subgraph includes four edges serving the same slice. The subgraph has the required property: if any of its edges is removed, it still contains paths between all its slice nodes (4, 8, and 9). For instance, if the (primary) link between nodes 4 and 8 fails, there is a backup path: 4-6-9-8. In terms of graph theory, the subgraph in Fig. 3 is 2-edge-connected. It is more efficient since it provides both primary and backup paths between its nodes and requires fewer resources than the two Steiner trees in Fig.1.

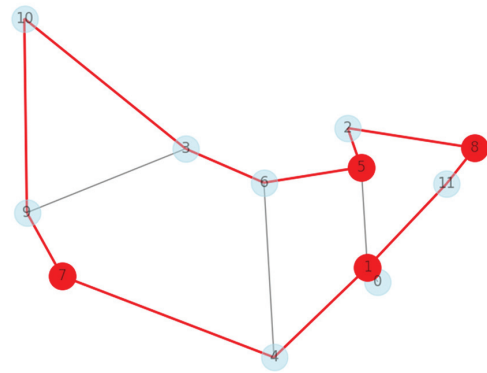


**Fig. 2.** A 2-edge-connected subgraph spanning the slice embedded in the nodes 4, 8, and 9, serving both primary and backup paths

This difference is significant in networks with low redundancy, where the number of disjoint paths between the nodes is small. In such networks, the previous proposal might not be able to construct two disjoint Steiner trees spanning each slice. For instance, Fig. 3 depicts the primary tree constructed for a slice spanning the nodes 1, 5, 7, and 8 in a topology with low redundancy ("abilene" network topology from the SNDlib dataset [11]). In this network, there is no way to construct a backup tree spanning the same nodes while avoiding the links of the primary tree. Therefore, this slice would not have path protection. On the other hand, the path protection can be achieved if a 2-edge-connected subgraph is constructed instead of a tree, serving both primary and backup paths, as depicted in Fig. 4. Such 2-edge-connected subgraph is inherently easier to construct than two spanning trees, which in most cases require more links and the existence of disjoint paths (high redundancy). The evaluation will present a numerical comparison (Sect. 5).



**Fig. 3.** Primary Steiner tree spanning the slice embedded in the nodes 1, 5, 7, and 8. A link-disjoint backup tree cannot be constructed (no path protection).



**Fig. 4.** Path protection of the slice embedded in the nodes 1, 5, 7, and 8 using a 2-edge-connected subgraph, allowing for a single link failure

To formalize the proposed framework, let  $N$  be the set of computing nodes,  $S$  the set of slices, and  $E$  the set of bidirectional links between node pairs  $(i, j)$ . We assume that each computing node  $i$  has  $R_i$  available resource elements (representing CPU units, memory units, available energy, or any other resource measure).

We assume that a slice  $s$  requires  $R_{s,i}$  resource elements when running on node  $i$  if those resources are allocated to  $s$ , in other words, if  $s$  is embedded at  $i$ . We assume that the fixed cost of embedding slice  $s$  at node  $i$  is  $F_{s,i}$  and that revenue generated per unit of resource allocated to slice  $s$  equals  $U_s$  (Table 1).

**Table 1.** Notation

Symbol	Meaning	Symbol	Meaning
$N$	set of computing nodes	$E$	set of bidirectional links $(i, j) \in N \times N$
$S$	set of slices	$L_{ij}$	latency of link $(i, j)$
$R_i$	available resource elements of node $i$	$C_{ij}$	capacity of link $(i, j)$
$R_{s,i}$	resource requirement of slice $s$ at node $i$	$R_{s,i,j}$	capacity requirement for slice $s$ on link $(i, j)$
$F_{s,i}$	cost of embedding slice $s$ at node $i$	$N_s$	set of nodes selected for slice $s$
$U_s$	revenue per unit of resource of slice $s$	$P_s$	links of the primary tree for slice $s$
$x_{s,i}$	whether slice $s$ is embedded at node $i$	$B_s$	links of the backup tree for slice $s$
$y_{s,ij}$	whether link $(i, j)$ is included in the primary subgraph for slice $s$	$z_{s,ij}$	whether link $(i, j)$ is included in the backup tree for slice $s$

We can now formalize the problem of slice selection in each computing node. Let  $x_{s,i} \in \{0,1\}$  be a binary decision variable indicating whether slice  $s$  is allocated resources at node  $i$ .

For each node  $i \in N$ , maximize the total slices' revenue:

$$\sum_{s \in S} x_{s,i} (R_{s,i} U_s - F_{s,i}) \quad (1)$$

With respect to available resources:

$$\sum_{s \in S} x_{s,i} R_{s,i} \leq R_i \quad (2)$$

Where,

- $x_{s,i}$  Binary decision variable indicating whether slice  $s$  is allocated resources at node  $i$
- $R_{s,i}$  Resource requirement of slice  $s$  at node  $i$
- $U_s$  Revenue per unit of resource of slice  $s$
- $F_{s,i}$  Cost of embedding slice  $s$  at node  $i$
- $S$  Set of slices
- $R_i$  Available resource elements of node  $i$ .

After the slice selection variables  $x_{s,i}$  are determined, the set of slice nodes is defined as the set of all nodes where slice  $s$  is embedded:

$$N_s = \{i \in N : x_{s,i} = 1\} \quad (3)$$

Where,

- $N_s$  Set of nodes selected for slice  $s$
- $N$  Set of computing nodes.

We assume that each link  $(i, j) \in E$  has known latency  $L_{ij}$  and capacity  $C_{ij}$ . Furthermore, if this link is embedded for slice  $s$ , each slice has a capacity requirement  $R_{s,i,j}$ .

We now jointly define the problems of path selection and path protection. For each slice  $s$ , a subgraph  $P_s (P_s \subset E)$  should be constructed so that all nodes in  $N_s$  are connected (directly or indirectly) by links from  $P_s$ . This subgraph should be 2-edge-connected to enable path protection in case of a single link failure. Alternatively, it can be a tree, and then it should be augmented with another backup tree  $B_s \subset E$ , which ensures path protection by providing backup paths in case of a single link fail-



ure in  $P_s$ . While constructing the subgraphs, the capacity constraints for each link should be met, and total latency should be minimal. We can formalize the path selection and path protection problem as follows.

$$\sum_{s \in S} \sum_{(i,j) \in E} y_{s,i,j} L_{i,j} \quad (4a)$$

$$\sum_{s \in S} y_{s,i,j} R_{s,i,j} \leq C_{i,j} \quad (4b)$$

Where,

- $L_{i,j}$  Latency of link  $(i, j)$
- $y_{s,i,j}$  a binary decision variable indicating whether link  $(i, j)$  is included in the primary subgraph for slice  $s$
- $R_{s,i,j}$  The capacity requirement for slice  $s$  on link  $(i, j)$
- $C_{i,j}$  The capacity of link  $(i, j)$ .

These equations correspond to a choice of 2-edge-connected subgraphs for  $P_s$ . Therefore, in the rest of the paper,  $P_s$  is the set of links  $(i, j)$  for which  $y_{s,i,j} = 1$ . In the case of trees, there are two subgraphs, so we modify the equations as follows.

Minimize:

$$\sum_{s \in S} \sum_{(i,j) \in E} (y_{s,i,j} + z_{s,i,j}) L_{i,j} \quad (5a)$$

$$\sum_{s \in S} (y_{s,i,j} + z_{s,i,j}) R_{s,i,j} \leq C_{i,j} \quad (5b)$$

Where  $z_{s,i,j}$  is a binary decision variable indicating whether link  $(i, j)$  is included in the backup tree for slice  $s$ . Therefore, in the rest of the paper,  $B_s$  is the set of links  $(i, j)$  for which  $z_{s,i,j} = 1$ .

#### 4. PROPOSED ALGORITHMS

The proposed framework is divided into two steps. The first step performs the slice selection in the computing nodes, determining the decision variables  $x_{s,i}$ . The second step performs the construction of slices' subgraphs. In both steps, we consider three variants of the proposed framework. Two of them construct primary and backup Steiner trees and are included mostly for evaluation and comparison purposes, while the third variant is our proposed approach based on 2-edge connectivity. More specifically:

- Steiner-balanced algorithm seeks to optimize the total revenue, as well as the number of accepted slices, i.e., the number of slices for which the primary tree is successfully constructed.
- Steiner-protection algorithm seeks to maximize the number of selected slices per node, as well as the number of protected slices, i.e., the number of slices for which both the primary tree and the backup tree are successfully constructed.

- Proposed variant constructs a 2-edge-connected subgraph for each slice instead of two Steiner trees, while also maximizing total revenue.

#### 4.1. SELECTING SLICES ON THE COMPUTING NODES

In the first step, since  $x_{s,i} \in \{0,1\}$ , we reduce the problem to a classical 0-1 Knapsack Problem and solve it by the corresponding dynamic programming algorithm.

Namely, for each computing node  $i$ , we consider all slices as potential items in the Knapsack with weight capacity  $R_i$ . Following Eq. (1), for an item corresponding to slice  $s$ , its value  $v_s$  equals the revenue  $R_{s,i} U_s - F_{s,i}$  while its weight  $w_s$  equals its resource requirement  $R_{s,i}$ . The goal of the 0-1 Knapsack Problem is to select items (without repetitions) to put in the knapsack to maximize the total value of selected items without exceeding the knapsack's weight capacity. If the corresponding item gets selected in the knapsack solution, then  $x_{s,i} = 1$ .

We solve the 0-1 Knapsack Problem independently for each node using the standard dynamic programming algorithm. The key idea is to build a 2D array  $dp$  where  $dp[k][w]$  represents the maximum value that can be achieved with the first  $k$  items and a knapsack capacity of  $w$ . We iterate through the items, updating the array based on whether we exclude or include the current  $k$ -th item with value  $v_s$  and weight  $w_s$  (if  $w \geq w_s$ ):

$$dp[k][w] = \max \begin{cases} dp[k-1][w] & \text{(exclude)} \\ v_s + dp[k-1][w-w_s] & \text{(include)} \end{cases} \quad (6)$$

To extend the solution to return which items are selected, we need to modify the algorithm to track the decisions made during the construction of the  $dp$  table. This reconstruction can be done by examining the  $dp$  table after it has been filled out and tracing back the choices that led to the maximum value.

The algorithm complexity is  $O(|S| \cdot |R_i|)$  for each of the  $|N|$  computing nodes. We thus obtain selected slices at each node and calculate the sets  $N_s$  using Eq. (3).

The described solution is used in the Proposed and Steiner-balanced variants. In the Steiner-protection variant, the 0-1 Knapsack is solved greedily by repeatedly selecting the item with the lowest weight, corresponding to the slice with the lowest resource requirement. This approach maximizes the number of slices embedded in the node.

After the computing nodes  $N_s$  for each slice have been determined, they should be connected by paths with respect to the latency requirements and capacity constraints. Namely, we are searching for a set of links  $P_s \subset E$  spanning all nodes in  $N_s$  with minimal total latency, which is the second step of the algorithm. The following two subsections describe the considered variants of this step, differing in the chosen structure of  $P_s$ .

## 4.2. PATH CONSTRUCTION: STEINER TREE APPROACHES

A natural idea is for  $P_s$  to be a tree because such a subgraph uses the smallest number of edges. In that case, however, path protection in case of a link failure requires one more tree for each slice – a backup tree  $B_s$ , which is link-disjoint with  $P_s$  ( $P_s \cap B_s = \emptyset$ ).

In a given graph, the construction of a tree that spans a given subset of nodes (terminals) is a standard *Steiner tree problem in graphs*. Its optimization variant which minimizes the total weight of the selected tree's edges (corresponding to link latencies  $L_{ij}$ ) can be solved by e.g. Kou's algorithm [22] which is implemented in a NetworX library [23]. For this algorithm, we consider only the links with enough capacity ( $C_{ij} \geq R_{s,ij}$ ). In this way, we obtain the primary tree  $P_s$  for each slice. Afterward, we construct the backup tree  $B_s$  in the same manner, considering only the links not selected in  $P_s$  to ensure link disjointness. The algorithm complexity is  $O(|N_s| \cdot |E|2)$  for each of the  $O(|S|)$  Steiner trees.

Whenever a primary or a backup tree is constructed, we must update the total link capacities by deducting the occupied capacity of the corresponding slice:

$$C_{i,j} := C_{i,j} - R_{s,i,j} \quad (7)$$

For this reason, the order in which the Steiner trees are constructed is not irrelevant. To maximize the number of accepted slices, we first sort the slices ascendingly by the number of embedding nodes  $|N_s|$ , since it is easier to construct a spanning tree for a smaller number of terminals.

A slice is *accepted* if its primary tree is successfully constructed and protected if its backup tree is successfully constructed as well. In the Steiner-balanced variant, we attempt to construct all primary trees for slices in the order described above, and then we attempt to construct all backup trees in the same order, optimizing the number of accepted slices by prioritizing the construction of primary trees. In the Protection-based variant, instead of constructing the backup trees after all primary trees, each backup tree is constructed immediately after the corresponding primary tree, attempting to maximize the number of protected slices.

## 4.3. PATH CONSTRUCTION: 2-EDGE-CONNECTED SUBGRAPH

The subgraph  $P_s$  spanning the nodes (terminals) selected for slice  $s$  does not have to be a tree. If  $P_s$  already contains multiple paths between each pair of terminals, path protection is ensured without constructing additional backup paths. This is the 2-edge-connectedness property.

To find such a subgraph, the following algorithm is applied for each slice:

1. We initialize  $P_s$  to be maximal, consisting of all edges  $(i, j) \in E$  with enough capacity for slice  $s$  ( $C_{ij} \geq R_{s,ij}$ ).

2. If this subgraph  $P_s$  is not 2-edge-connected, the slice is rejected.
3. Otherwise, we go over all edges in  $P_s$ , sorted by latency in the decreasing order.
  - 1) We check if  $P_s$  would still be 2-edge-connected if the current edge was removed.
  - 2) If the answer is yes, we remove the current edge and thus reduce the total latency of  $P_s$ .
4. Return the resulting 2-edge-connected subgraph  $P_s$ .

Since edges with higher latency are removed first, the subgraph obtained after removing all unnecessary edges will have minimal latency. It remains to describe how to check if a graph is 2-edge-connected. Verifying the 2-edge-connectedness of  $P_s$  is done in the following way:

1. For each pair of terminals  $u$  and  $v$  in  $P_s$ , the function *edge connectivity* from the NetworX library [23] is called to compute the minimum number of edges that must be removed to break all paths from  $u$  to  $v$ .
2. If the obtained edge connectivity number is greater than 1 for all pairs of terminals  $u$  and  $v$ , the subgraph  $P_s$  is 2-edge-connected.

After constructing each  $P_s$ , link capacities are updated by Eq. (7). The order in which the slices are processed is the same as in the previous variants (increasing  $|N_s|$ ). Since each subgraph provides both primary and backup paths for the corresponding slice, the number of accepted and protected slices is equal in this variant. It equals the number of successfully constructed 2-edge-connected subgraphs.

## 5. EVALUATION

This section describes the simulation results. All experiments were implemented in Python 3 on Intel(R) Core(TM) i7-10510U CPU @ 1.80GHz. We used the network topologies from the SNDlib dataset [8]. We note that most topologies were not suitable for the Steiner-based approaches because of their low redundancy which disabled the construction of link-disjoint trees. The suitable high-redundancy topologies were *pdh*, *dfn-bwin*, *newyork*, and *pioro40*, denoted by T1, T2, T3, and T4 in the rest of this section. The link capacity values (in Mbit/s) by individual topologies are: T1: 30, T2: 80000, T3: 1000 and T4: 155.

Apart from using realistic network topologies, we used the link capacities from the same SNDlib files and artificially generated other parameters. We assumed  $|S| = 8$  slices, the resource capacity of each node was 100 elements, while the resource requirement per slice was a random integer uniformly chosen from [10, 100]. Revenue per resource unit and embedding costs were random integers uniformly chosen from [10, 100] and [100, 500], respectively. Link latencies were uniformly chosen between 1 and 20 ms. Slices' capacity require-

ments were uniformly chosen between 1 and the average link capacity divided by 2. Each experiment was repeated 30 times, and the results were averaged.

Fig. 5 compares the algorithm variants and the KC-FS algorithm [16] across several measures. Since the KC-FS performs the second step of the present framework and not the first, the same knapsack-based selection of slice nodes (or SF candidates) is performed as the first step of KC-FS. For its second step, the slices are sorted descend-

ingly by the number of embedding nodes, following the recommendation from Algorithm 4 in [16] which starts from the SF with the most candidates in the embedding network. The main KC-FS algorithm is implemented according to the pseudocode of Algorithm 3 in [16], including Algorithm 1 (*k-Connected Network Slicing Technique*) and Algorithm 2 (*i-Based Node Degree Balancing*) as subroutines. The following paragraphs describe the obtained results by subfigures of Fig. 5.

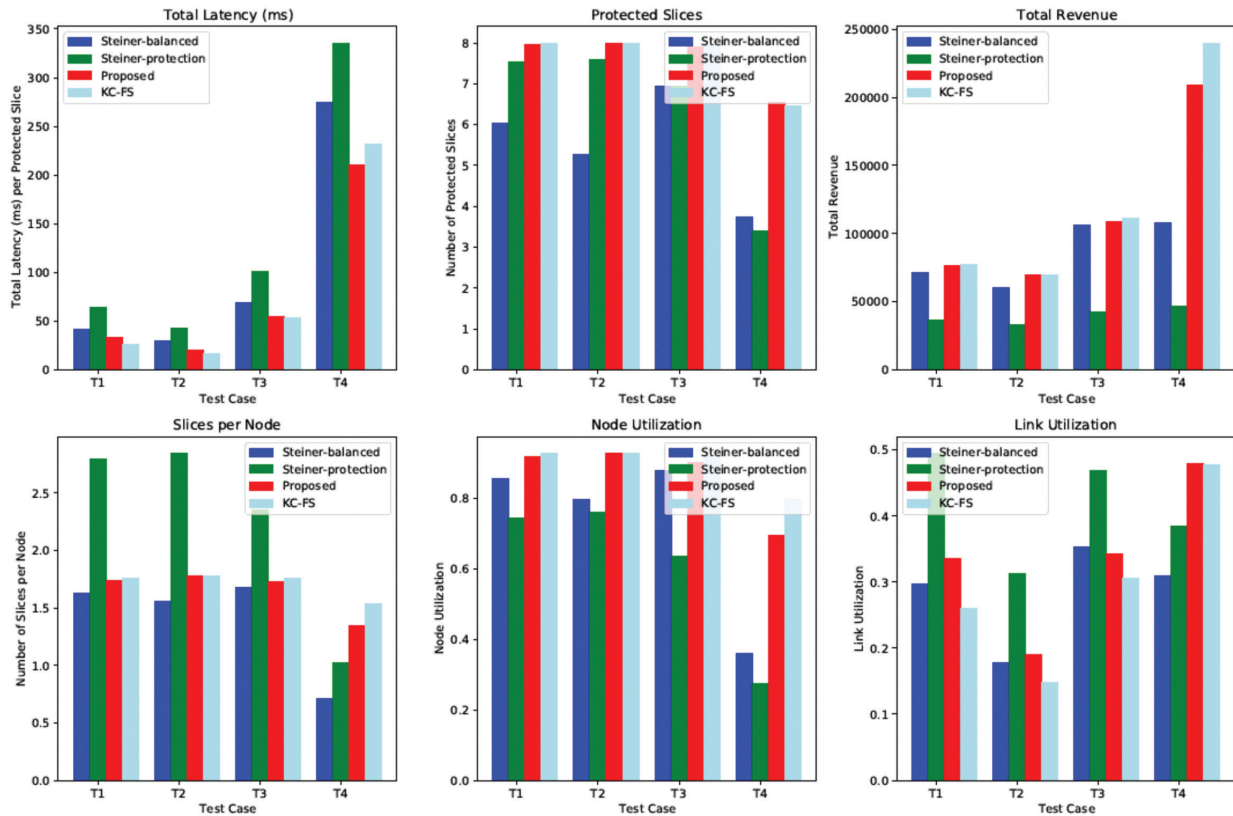


Fig. 5. Comparison between different approaches

The *Total Latency* for each subgraph was calculated by adding up the latencies of its links, and the results were averaged for all protected slices. The Proposed and KC-FS approaches achieve the best results in terms of total latency, with KC-FS giving slightly lower latencies in T1 and T2 and higher latencies in T4. The Steiner-protection variant achieves higher latencies than the Balanced-protection variant because it has the highest number of *slices per node* in all test cases, resulting from its approach to 0-1 Knapsack.

The Steiner-protection algorithm, however, has a higher number of *Protected Slices* in T1 and T2 compared to the Steiner-balanced. The Proposed and KC-FS approaches are the best in this respect, behaving similarly and enabling the highest number of protected slices.

*Total revenue* (of accepted slices) is expectedly lowest for the Steiner-protection approach, which does not take revenue into account when solving the 0-1 Knapsack. The Proposed approach is more successful than

Steiner-balanced, especially in T4, because of the much higher number of protected slices. The KC-FS approach is even more successful in T4 because it starts from the slices with the most nodes.

*Node utilization* in Fig. 5 was calculated as the average percentage of utilized required resource elements by the slices embedded in each node  $i \in N$ . Link utilization was calculated as the average percentage of utilized capacities by the slices for which a link  $(i, j) \in E$  is chosen for the subgraph/tree. Steiner-protection variant achieves lower node utilization compared to other approaches because of its selection of low-requirement slices when solving the 0-1 Knapsack. The opposite is true for link utilization since the Protection-based approach constructs larger trees because of the higher number of nodes per slice. The Proposed variant achieves slightly higher results than the Steiner-balanced with respect to node and link utilization. The Proposed and KC-FS approaches again behave similarly, with KC-FS utilizing slightly more nodes because of its slice ordering.

The most important difference between Proposed and KC-FS is depicted in Fig. 6 which shows the average execution time per test case. For small networks such as T1, T2, and T3 the execution times are negligible, but for a larger network such as T4 the KC-FS subgraph embeddings take  $\approx 3.5\times$  more computation time in comparison to Proposed. This is in line with the complexity of the procedures in Algorithms 1-3 of [16].

The results show the overall dominance of the Proposed approach in terms of execution time, latency, number of protected slices, total revenue, and node utilization. In terms of other measures, the Steiner-protection variant has the highest edge utilization (except in T4), as well as the highest number of slices per node and, equivalently, nodes per slice. However, it is dominated by the Steiner-balanced approach in terms of total revenue. In cases where execution time is irrelevant, the KC-FS approach with knapsack-based node selection can be the best choice in terms of total revenue (T4).

To visualize the effect of network topology, Fig. 7 depicts the resulting subgraphs of specific slices on 12 different

low-redundancy topologies. Only the Proposed variant achieved path protection on these topologies.

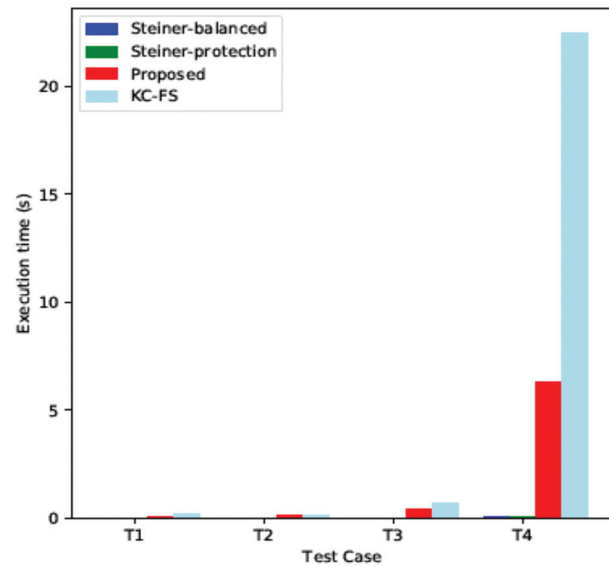


Fig. 6. Average execution time

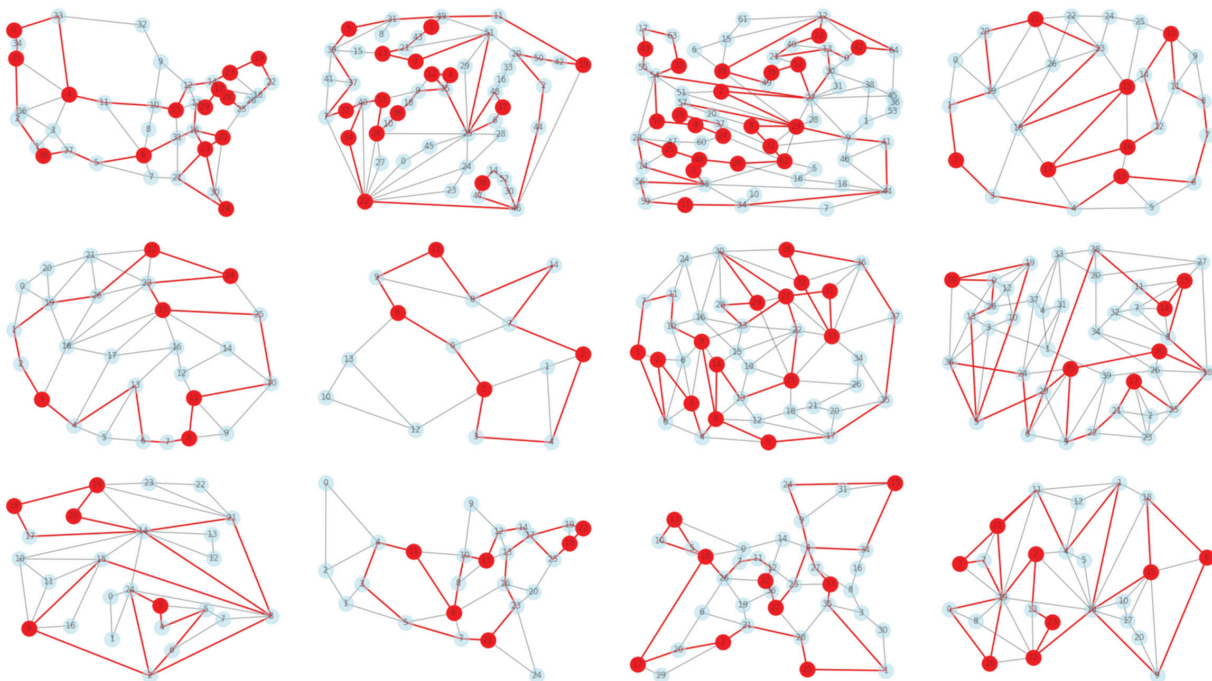


Fig. 7. Slice subgraphs on low-redundancy topologies (row by row: *janos-us-ca*, *zib54*, *ta2*, *norway*, *sun*, *atlanta*, *giul39*, *pioro40*, *france*, *janos-us*, *cost266*, and *ta1* from SNDlib dataset [8]).

## 6. CONCLUSION

In this paper, we presented a framework with several variants for slicing optimization and path protection in 5G networks, focusing on balancing factors like execution time, revenue, latency, slice protection, and resource utilization.

The proposed approach based on 2-edge-connected subgraphs emerged as the most effective, achieving the lowest total latency, highest number of protected

slices, highest node utilization, and highest total revenue along with the existing KC-FS approach which has a significantly higher execution time. This shows that the proposed algorithm created slice subgraphs more efficiently than the recent related approach. The Steiner-protection approach created larger and more protected slices than the Steiner-balanced approach, emphasizing the importance of prioritizing network resiliency with dedicated path protection. However, this comes at the expense of the total revenue, which is dominated by the



Steiner-balanced approach, underscoring the efficacy of dynamic programming in optimizing network resources and showing a tradeoff between revenue maximization and other optimization aspects, such as slice protection and the number of nodes per slice.

## 7. REFERENCES

- [1] D. Irawan, N. R. Syambas, A. A. N. Ananda Kusuma, E. Mulyana, "Network Slicing Algorithms Case Study: Virtual Network Embedding", Proceedings of the 14th International Conference on Telecommunication Systems, Services, and Applications, Bandung, Indonesia, 4-5 November 2020, pp. 1-5.
- [2] M. Aitaba, M. Elkael, B. Jouaber, H. Castel-Taleb, A. Araldo, D. Olivier, "A two-stage algorithm for the Virtual Network Embedding problem", Proceedings of the 46th Conference on Local Computer Networks, Edmonton, AB, Canada, 4-7 October 2021, pp. 395-398.
- [3] M. K. Singh, S. Vittal, A. A. Franklin, "SERENS: Self Regulating Network Slicing in 5G for Efficient Resource Utilization", Proceedings of the 3rd 5G World Forum, Bangalore, India, 10-12 September 2020, pp. 590-595.
- [4] V. Balasubramanian, M. Aloqaily, M. Reisslein, "Mutes: Multi-Tenant Switching for 5G Network Slice Revenue Maximization", Proceedings of the International Wireless Communications and Mobile Computing, Dubrovnik, Croatia, 30 May - 3 June 2022, pp. 590-595.
- [5] B. Jaumard, Q. H. Duong, "A Nested Decomposition Model for Reliable NFV 5G Network Slicing", IEEE Transactions on Network and Service Management, Vol. 20, No. 3, 2023, pp. 2186-2200.
- [6] C. Raffaelli, E. Amato, P. Monti, F. Tonini, "Reliable Slicing in Optical Metro Networks with Reconfigurable Backup Resources", Proceedings of the International Symposium on Communication Systems, Networks and Digital Signal Processing, Porto, Portugal, 20-22 July 2022, pp. 863-866.
- [7] N. Shahriar et al. "Reliable Slicing of 5G Transport Networks with Bandwidth Squeezing and Multi-Path Provisioning", IEEE Transactions on Network and Service Management, Vol. 17, No. 3, 2020, pp. 1418-1431.
- [8] S. Orłowski, R. Wessälly, M. Pióro, A. Tomaszewski, "SNDlib 1.0—Survivable Network Design Library", Networks, Vol. 55, No. 3, 2010, pp. 276-286.
- [9] M. K. Lee, C. S. Hong, "Efficient Slice Allocation for Novel 5G Services", Proceedings of the 10th International Conference on Ubiquitous and Future Networks, Prague, Czech Republic, 3-6 July 2018, pp. 625-629.
- [10] A. Sarah, G. Nencioni, "Resource allocation for cost minimization of a slice broker in a 5G-MEC scenario", Computer Communications, Vol. 213, No. 3, 2024, pp. 331-344.
- [11] Y. Xiao, J. Zhang, P. Zhu, H. Wu, C. Zhang, "Customized Topology-Level Protection for Reliable Slicing in 5G/B5G Metro Access/Aggregation Networks", Journal of Lightwave Technology, Vol. 42, No. 9, 2024, pp. 3068-3080.
- [12] D. Basu, S. Kal, U. Ghosh, R. Datta, "DRIVE: Dynamic Resource Introspection and VNF Embedding for 5G Using Machine Learning", IEEE Internet of Things Journal, Vol. 10, No. 21, 2023, pp. 18971-18979.
- [13] H. Ko, J. Lee, H. Choi, S. Pack, "Hierarchical Identifier (HID)-based 5G Architecture with Backup Slice", Proceedings of the 21st Asia-Pacific Network Operations and Management Symposium, Daegu, Korea, 22-25 September 2020, pp. 291-293.
- [14] N. Shahriar et al. "Reliable Slicing of 5G Transport Networks with Dedicated Protection", Proceedings of the 15th International Conference on Network and Service Management, Halifax, NS, Canada, 21-25 October 2019, pp. 1-9.
- [15] I. Begić, A. S. Kurdija, J. Matuško, "A Framework for 5G Network Slicing Optimization with Path Protection", Proceedings of the 47th International Convention on Information and Communication Technology, Electronics and Microelectronics, Opatija, Croatia, 22-26 May 2024.
- [16] D. Zheng, G. Shen, Y. Li, X. Cao, B. Mukherjee, "Service Function Chaining and Embedding with Heterogeneous Faults Tolerance in Edge Networks", IEEE Transactions on Network and Service Management, Vol. 20, No. 3, 2023, pp. 2157-2171.
- [17] Y. Zheng, S. Ravi, E. Kline, L. Thurlow, S. Koenig, T. K. S. Kumar, "Improved Conflict-Based Search for

- the Virtual Network Embedding Problem", Proceedings of the 32nd International Conference on Computer Communications and Networks, Honolulu, HI, USA, 24-27 July 2023, pp. 1-10.
- [18] L. Yang, J. Jia, H. Lin, J. Cao, "Reliable Dynamic Service Chain Scheduling in 5G Networks", *IEEE Transactions on Mobile Computing*, Vol. 22, No. 8, 2023, pp. 4898-4911.
- [19] R. Mohamed, A. Leivadeas, I. Lambadaris, T. Morris, P. Djukic, "Online and Scalable Virtual Network Functions Chain Placement for Emerging 5G Networks", Proceedings of the IEEE International Mediterranean Conference on Communications and Networking, Athens, Greece, 5-8 September 2022, pp. 255-260.
- [20] H. Liu, "Research on Virtual Network Embedding Model and Algorithm Based on Graph Attention Network and Multi-Layer Perceptron", Proceedings of the 5th International Conference on Artificial Intelligence and Computer Applications, Dalian, China, 2023, pp. 812-818.
- [21] A. Samar, K. M. Sivalingam, "RL-based Virtual Network Embedding using VNF Sharing for Network Slicing in 5G Networks", Proceedings of the IEEE/IFIP Network Operations and Management Symposium, Miami, FL, USA, 8-12 May 2023, pp. 1-7.
- [22] L. Kou, G. Markowsky, L. Berman, "A Fast Algorithm for Steiner Trees", *Acta Informatica*, Vol. 15, No. 2, 1981, pp. 141-145.
- [23] A. A. Hagberg, D. A. Schult, P. J. Swart, "Exploring network structure, dynamics, and function using NetworkX", Proceedings of the 7th Python in Science Conference, Pasadena, CA, USA, 19-24 August 2008, pp. 11-15.



# Automated vs. Semi-Automated Hydroponics: Quantifying Automation Effects on Plant Growth

Original Scientific Paper

**Swati Jain\***

Sharda University,  
School of Engineering and Technology, India  
jainswati3107@gmail.com

**Mandeep Kaur**

Sharda University,  
School of Engineering and Technology, India  
Mandeep.kaur@sharda.ac.in

\*Corresponding author

**Abstract** – The issue of hydroponic farming is the ongoing requirement to maintain and control the artificial growing environment to enable optimal plant growth. The quality of the plants growing on the farm can be significantly impacted by changes in the climate, natural light, and fertilizer solution at any time during the plant's growth cycle. According to studies, every 3°C increase in ambient temperature from the optimal range reduces crop productivity by 10 to 40%. Therefore, by automatically supplying ideal growth conditions throughout the growing cycle, IoT-based automation is crucial to achieving optimal plant growth. In this work, lettuce is grown simultaneously in two distinct hydroponic Nutrient Film Technique (NFT) systems, one of which is fully automated and the other semi-automated. In terms of plant growth metrics, such as plant elevation, maximum plant length, maximum plant breadth, and weight of both fresh and dry-farmed lettuce, this paper compared a fully automated farm and a semi-automated NFT farm for growing lettuce. The outcomes demonstrate that the ariel weight of fresh and dry lettuce and root weight of fresh and dry lettuce had average improvements of 10.4 gm, 0.7 gm, 0.11 gm, 0.7 gm, and 3, respectively in the fully automated setup. Findings also demonstrated that IoT-based automation enhances the growth of lettuce plants on farms when comparing a completely automatic hydroponic farm to a semi-automatic hydroponic farm, in terms of plant height, width, and total leaf count.

---

**Keywords:** hydroponics, automation, IoT, nutrient control, dosing, weather control

---

Received: March 26, 2024; Received in revised form: June 13, 2024; Accepted: June 13, 2024

## 1. INTRODUCTION

For many nations around the world, farming is an essential component of their economies. However, the capacity of conventional soil-based farming (SBF) to produce crops has been negatively impacted by the recent linear increases in population, urbanization, and food demands, as well as the concurrent drop in the amount of land and water available for farming. The disparity between the availability and demand of food is gradually widening, over time [1]. One of the biggest problems SBF has is that there are not enough resources to meet the growing demand for food, such as fertile land and water, which are essential for farming. Therefore, using the resources at hand, a smart agriculture technique is required that can produce higher yields than SBF. The majority of traditional farming's issues can be resolved with hydroponics, a branch of precision farming [1, 2].

Traditional agriculture faces challenges like climate change, soil degradation, water scarcity, labour shortages, and pathogen contamination, which impact crop yields and farming practices. Hydroponic farming addresses these issues by using soil-less systems that conserve water, eliminate soil-related problems, and allow for precise climate control. Automation and IoT technologies enhance labour efficiency, while controlled environments reduce exposure to pathogens, ensuring safer produce. This makes hydroponics a sustainable alternative to traditional farming [3]. Throughout the crop's growth cycle, HF must continuously monitor and maintain artificial climate conditions to replicate traditional weather and soil conditions. Manually monitoring and controlling weather, light, and nutrient conditions could lead to human error and a divergence from the ideal requirements. Taking this into account, to attain the required crop quality, the process of main-



taining the artificial growing environment must be automated. One such method is the Internet of Things (IoT), which enables the HF to employ sensors, controls, and actuators to maintain its ambient conditions [3-6]. By continuously monitoring growth parameters and sending on/off signals to actuators connected to devices whenever any parameter goes outside of the ideal range, IoT technology enables precision agriculture in HF.

The research questions that motivated the basis of this study were: i. Effectiveness of the manual and semi-automated farm in maintaining artificial growth environment, ii. The effectiveness of a fully automated IoT farm in maintaining the artificial growth environment and iii. The difference is the crop quality harvested from semi-automated and fully automated farms. The objective of this study is to experimentally compare and study the effectiveness of IoT automation on the hydroponic farm for maintaining the growth parameters and the eventually harvested crop products. This study compares the crop production quality attained in two types of hydroponic farms: a fully automated hydroponic farm setup (FAHS), where the IoT setup maintains all climate conditions without the need for human intervention, and a semi-automated hydroponic farm setup (SAHS), where sensors are installed for remote farm monitoring but the farm owner is still responsible for controlling and maintaining the conditions [6]. A comparison of the following growth parameters was conducted after a 30-day growth cycle of lettuce crops planted simultaneously in FAHS and SAHS: Plant Elevation (PE), Maximum Plant Breadth (MPB), Ariel Weight of Fresh Lettuce (AW-FL), Root Weight of Fresh Lettuce (RW-FL), Ariel Weight of Dry Lettuce (AW-DL), Root Weight of Dry Lettuce (RW-DL), and Total Leaf Count (TLC). Following that, an analysis of the data from SAHS and FAHS was conducted to investigate the effects of semi-automation and complete automation on plant development [7].

The structure of this paper is as follows: The FAHS and SAHS systems utilized in this study are introduced in Section 1. In the second section, a thorough analysis of previous research papers that compare various automated HFs has been provided. Section 3 discusses the experimental setup's approach as well as the comparison parameters. The experimental analysis and comparison of the observed growth parameters are presented in Section 4. This section compares the quantitative findings for the growth metrics for both mediums, Finally, section 5 concludes the study's findings and future directions.

## 2. LITERATURE REVIEW

An extensive review of the literature on automated and semi-automated farms was conducted, focussing on evaluating the efficacy of artificial vs substrate media for hydroponic farms. Experimental analysis has been carried out by researchers and developers to in-

vestigate the impact of varying degrees of automation on the quality of hydroponic crop production.

To produce lettuce, the research work [12] compared the performance of conventional soil-based growing farms and commercial hydroponic farms. An analysis was conducted on the productivity of lettuce in hydroponic systems and alternative soil-based systems, as well as the financial advantages of a greenhouse farm. Even though the hydroponic farm used less than 50% of the water utilized by the soil-based farm, the output acquired by the former was 134% of the yield obtained by the latter throughout two growth cycles of lettuce grown in identical greenhouses. The hydroponic farm's initial investment and overall running costs [8] were, respectively, 21.76 times and 47% greater than those of the conventional farm. The hydroponic farm produced lettuce with superior overall quality and performance than the traditional farm, despite the automated farm requiring higher capital and operational costs.

The growth performance of a crop of romaine lettuce in a fully automated hydroponic setup was compared to that of a crop grown in an uncontrolled, non-automated hydroponic plant by the authors in their study [9]. The following metrics were used to measure and assess the plant growth in a single harvest cycle, or 30 days: plant height, maximum leaf length and width, and the weight of the farmed plant. The controlled and automated arrangement produced a greater yield and better-quality growth in terms of weight, height, and leaf count, according to the growth findings collected from both setups. The parameters recorded in the automated setup remained in the optimal range throughout the 30-day development cycle, while the non-automated configuration required manual interference to keep the growth environment parameters (temperature, humidity, pH, and TDS) in the optimal range.

A low-cost Arduino system was used in [3] to track and maintain the aquaponics setup's growth parameters automatically. The farm's irrigation, humidity, and temperature controls were automated to maximize growing production and make frugal use of the resources at hand. When compared to the traditional arrangement, the results showed a boost in leaf yield of over 40% in the automated farm and a 400-fold increase in cultivable area density in the automated setup.

Using fuzzy logic and IoT technologies, a hydroponic farm with parameter monitoring and precision system was put into operation. Fuzzy logic was used to ensure a steady supply of water and nutrients while monitoring the nutrient levels and intake of the bok choy and lettuce crops [5, 6]. An automated vertical hydroponic farm utilizing robotic technology and IoT was established in the study [4]. The height and width of the leaf crop cultivated in the automated vertical farm were compared in detail with a traditional vertical farm that did not have any robotic automation. The automated farm's average plant height and height were found to be 12.08 cm and 5.5 cm, respectively, while the con-

ventional farm's average plant height and height were 10.58 cm and 3.9 cm, respectively. The findings indicate that automation has a greater impact on plant development than conventional growth media.

In their investigation [7], scientists grew lettuce plants in two distinct media: a soil-based substrate medium and a hydroponic system based on NFT. Based on the length of the development cycle, leaf length, leaf count, and leaf perimeter, the crop's growth results in both media were compared and examined [7]. The results from the AquaCrop simulator were also checked and validated with the substrate medium's actual growth outcomes. According to the data, growing lettuce in a hydroponic medium yields a higher quality and quantity of crop than growing it in a substrate media, and it does it in a way that is resource-competent and sustainable. The study brought to light the difficulty of hydroponic farming, which is that automated farms use about 70 times more energy than the substrate medium.

An analysis of the benefits and drawbacks of automated hydroponic farming with traditional agriculture can be found in [10]. The benefits of automated hydroponics included: (i) yields that were 11 times higher than those from conventional media; (ii) no stubble is produced; this prevents pollution from burning stubble; and (iii) improved crop quality in terms of plant height and weight. The study identified the following difficulties with the hydroponic medium: (i) a hydroponic farm's energy consumption was roughly 82 times greater than that of a traditional soil medium; (ii) a hydroponic farm required a larger initial investment and ongoing maintenance costs than a soil-based medium [10]. Smart Grow, an automated hydroponic system with minimal operating costs, was introduced in [11] as a solution to the drawbacks of high investment and operational expenses for hydroponic farms. The farm's fundamental growth characteristics, including pH, EC, and water levels, may be tracked by the system. The growth characteristics were then examined and the Smart Grow system was contrasted with the conventional soil medium. The ensabi planted in Smart Grow and soil medium had an average height of 12 cm and 9 cm, respectively.

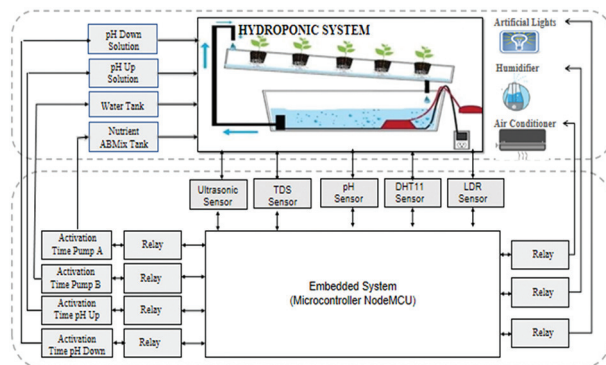
The literature review made it clear that HF and conventional SBF had been the subject of extensive prior research, as well as the effects of both growth media on plant growth [4-7, 9, 12-14]. It is also evident that a significant amount of study on the effects of automated and manual hydroponic farms on plant development has not yet been conducted thus creating a research gap. The results of this study will aid in the subsequent investigation of how various IoT automation levels in a hydroponic farm affect plant development.

### 3. PROPOSED AUTOMATED AND SEMI-AUTOMATED HYDROPONIC NFT SETUP

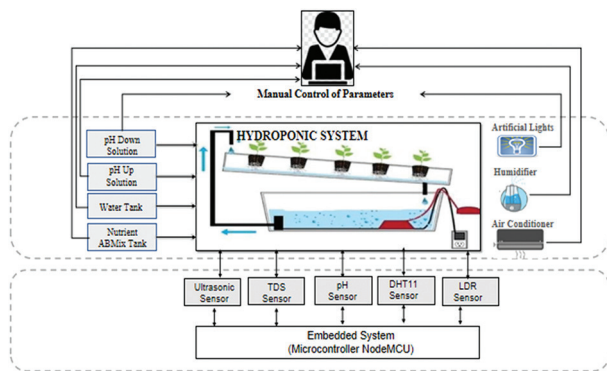
We have grown lettuce crops for one growth cycle, or thirty days, in this comparative study using an IoT-based fully automated hydroponic system (FAHS) and semi-automated setup (SAHS). The two configurations were running concurrently using the algorithm implemented in the study [7] to give the crop on both farms the same environmental conditions.

The lighting and climate automation suggested in [6, 8] and the nutrient control and automation from the setup used in [7, 11] were used to create the hydroponic farm structure to reach 100% automation. The nutrient film technique (NFT) is the foundation of the hydroponic system developed for this investigation. The system includes water pumps, an air conditioner, a humidifier, artificial lighting, a water tank, a pH up/down solution, polyvinyl chloride (PVC) hydroponic pipes, and a nutrient solution container. Plants in NFT have their roots immersed in water that has dissolved fertilizer solution. The plant's roots acquire all the nutrients needed for their growth from the water supplied by the hydroponic system. The concentration of nutrients in the water passing through the pipes is essential to the hydroponic plant's growth, and the pH of the water affects how much of the nutrients the roots can absorb.

Additionally, this hydroponic system was fitted with sensors and actuators to allow automated, human-free nutrient supply, watering, and environment management. The DHT11 temperature and humidity sensor, the pH sensor track the pH level of the nutrient water passing through the NFT system, the ultrasonic sensor monitors the water level in the irrigation tank, and the TDS sensor records the TDS level of nutrient water were the sensors utilized in FAHS. Actuators automatically monitor and maintain the growth parameters in the ideal range based on sensor data, which functions as a feedback loop. When any of the farm's development parameters deviate from their ideal range, actuators allow us to continuously manage them without the assistance of a human. Actuators were utilized in FAHS to run water pumps, humidifiers, and air conditioners. Fig. 1 shows the FAHS block structure, and Table 1 provides information on each sensor, including its ideal ranges.



**Fig. 1.** Block Diagram of Fully Automated Hydroponic Setup (FAHS)



**Fig. 2.** Block Diagram of Semi Automated Hydroponic Setup (SAHS)

The next stage was to put in place semi-automated (SAHS), in which a human source will manage all parameter handling rather than actuators. All of the actuators in the FAHS are removed (Fig. 1), and the manual device control is moved to a human source, providing us with the necessary semi-automated configuration for comparative analysis. As shown in Fig. 2, the NFT-based hydroponic system, NodeMCU Micro-controller, and sensor configurations were the same in SAHS and FAHS. The growth parameters are continuously monitored by the sensors in SAHS. If any of the parameters deviates from the required ideal range, the micro-controller unit uses WiFi to send a warning message to the mobile application. Via the Android app, the person in charge of the farm can view the warning message and take appropriate action. The person as mentioned earlier is responsible for managing the outside world in this system. For example, SAHS will issue an alarm via the mobile app if the DHT11 sensor detects a temperature higher than 26°C (Table 1). To return the temperature to the ideal range, the person will have to manually turn on the air conditioner. This temperature sensing and decision-making process—which for FAHS involves turning on the air conditioner—will be automated. Without any assistance from a human, the air conditioner will turn on automatically when the micro-controller unit sends the relay attached to it the on signal. Thus, there are no delays in FAHS when managing growth parameters.

**Table 1.** Sensors Deployed In FAHS And SAHS

Functionality	Sensor	Growth Parameter & Optimal Range for Lettuce Crop
Climate Monitoring & Control	DHT11	Temperature: 20 to 25°C Humidity: 40% to 75%
	pH Sensor	pH level of the Nutrient Water in NFT 5.7 to 6.5
Irrigation and Nutrient Control	TDS Sensor	TDS of the Nutrient Water in NFT 1-10 Days: 550 - 850 ppm 11 - 20 Days: 680 - 900 ppm 21 - 30 Days: 800 - 900 ppm
	Ultrasonic Sensor	Water Level in the Irrigation Tank Water level >= 85%

#### 4. EXPERIMENTAL ANALYSIS OF THE PROPOSED SETUP

The lettuce crop was grown for one growth cycle, or thirty days, in both the fully automated hydroponic setup (FAHS) and the semi-automated hydroponic setup (SAHS), to conduct the experimental analysis. To monitor the progress of plant growth, a variety of metrics were periodically recorded during the lettuce's growth phase. On days 5, 10, 15, 20, 25, and 30 of the growth cycle, three parameters—current plant elevation, current plant breadth, and current leaf length—were recorded [Table 2]. On the day of harvest, or day 30, other parameters—such as the aerial weight of dry lettuce, root weight of dry lettuce, and total number of leaves in the final harvest—were recorded [Table 3].

Using a weighing instrument, the plants were immediately weighed after being picked from both setups to examine the properties of fresh plants. The total number of lettuce leaves gathered in each setup was counted to get the total Leaf Count (TLC) metric. To conduct additional research, the aerial parts of each plant, or the leaves and stems, were separated from their roots and weighed individually once again using a weighing apparatus to determine the following parameters: Aerial Fresh lettuce weight (AW-FL) and fresh lettuce root weight (RW-FL). The aerial and root parts of the lettuce gathered from both setups were dried in an oven set at 500 °C for five hours to examine the weight parameters of dry lettuce. The aerial weight of dry lettuce (AW-DL) and root weight of dry lettuce (RW-DL) were measured following the drying process.

**Table 2.** Measured Values of Plant Growth Parameters for FAHS and SAHS for 10 plants taken on the days 5, 10, 15, 20, 25 and 30 of the growth cycle

Days	Mean Plant Elevation (Cm)		Mean Plant Width (Cm)	
	FAHS	SAHS	FAHS	SAHS
Day 5	2.5	2.3	3.4	3.2
Day 10	7.8	7.2	4.5	3.3
Day 15	12.5	11.3	7.3	4.8
Day 20	19.3	15.7	10.8	6.7
Day 25	24.8	20.7	13.9	9.8
Day 30	28.9	22.3	16.7	11.9

##### 4.1. RESULT ANALYSIS AND DISCUSSION

Calculating the plants' agronomic parameters was the first step in the experimental investigation. The average plant height and average plant breadth for ten plants growing in automated (FAHS) and semi-automated (SAHS) farms were measured on days 5, 10, 15, 20, 25, and 30 of the growth cycle, as shown in Table 2. Following the completion of the growth cycle, the plants were harvested, and measurements were made of the following parameters for ten plants each from FAHS and SAHS [Table 3]: Fresh lettuce (AW-FL), dry lettuce (AW-DL), fresh lettuce (RW-DL), and total leaf

count (TLC) are the weights measured at the air, on the root, and on the surface.

The mean plant height and mean plant width in FAHS were 28.9 cm and 16.7 cm, respectively, and in SAHS they were 22.3 cm and 11.9 cm, according to the results shown in Table 2. It follows that for plants cultivated in the automated farm, there was an average improvement of 6.6 cm in plant height and 4.8 cm in plant breadth.

Similarly, Table 3 results show that when comparing the automated setup to the semi-automated configuration, there was an improvement of 10.4 gm in AW-FL, 0.7 gm in AW-DL, 0.11 gm in RW-DL, 0.7 gm in RW-FL,

and 3 in Total Leaf Count. The ideal development environment that the plants were given for their growth was the root cause of the superior growth quality in FAHS compared to SAHS. In the semi-automated configuration, all the growth parameters were managed manually, while in the automated arrangement, the IoT system carried out the entire process of temperature, humidity, light monitoring, irrigation, and nutrient regulation automatically. This causes a variation in the efficiency and precision of the growing conditions given to the plants, which enhances the growth quality in FAHS. Fig. 4 shows the comparative analysis of agromonic characteristics between FAHS and SAHS.

**Table 3.** Sensor Values of Ambient Parameters for FAHS and SAHS for days 20 to 30 of the plant growth cycle

DAY	AVG. TEMP (CELCIUS)		AVG. HUMIDITY(%)		AVG. pH		AVG. TDS (PPM)	
	FA	SA	FA	SA	FA	SA	FA	SA
20	26.5	28.6	42.5	45.7	5.9	5.8	810	823
21	25.7	21.0	45.4	50.9	5.7	6.0	850	783
22	25.9	26.8	47.8	39.8	5.6	5.7	834	793
23	24.4	24.9	41.8	36.9	6.0	5.2	856	910
24	26.2	24.7	42.4	58.2	5.9	5.5	879	854
25	24.9	20.4	44.2	39.1	6.0	6.1	894	869
26	25.7	19.0	43.4	55.8	5.3	4.9	872	813
27	24.9	21.4	41.8	59.5	6.0	5.4	852	784
28	25.8	19.9	46.2	44.9	5.9	5.5	885	790
29	25.5	20.3	46.3	35.9	5.6	5.8	823	834
Mean	24.54	22.71	44.18	46.97	5.7	5.9	855.5	825.3

**Table 4.** Measured Values of Plant Growth Parameters for Fully automated(FA) and Semi automated(SA) for 10 plants taken on the 30th day of growth cycle

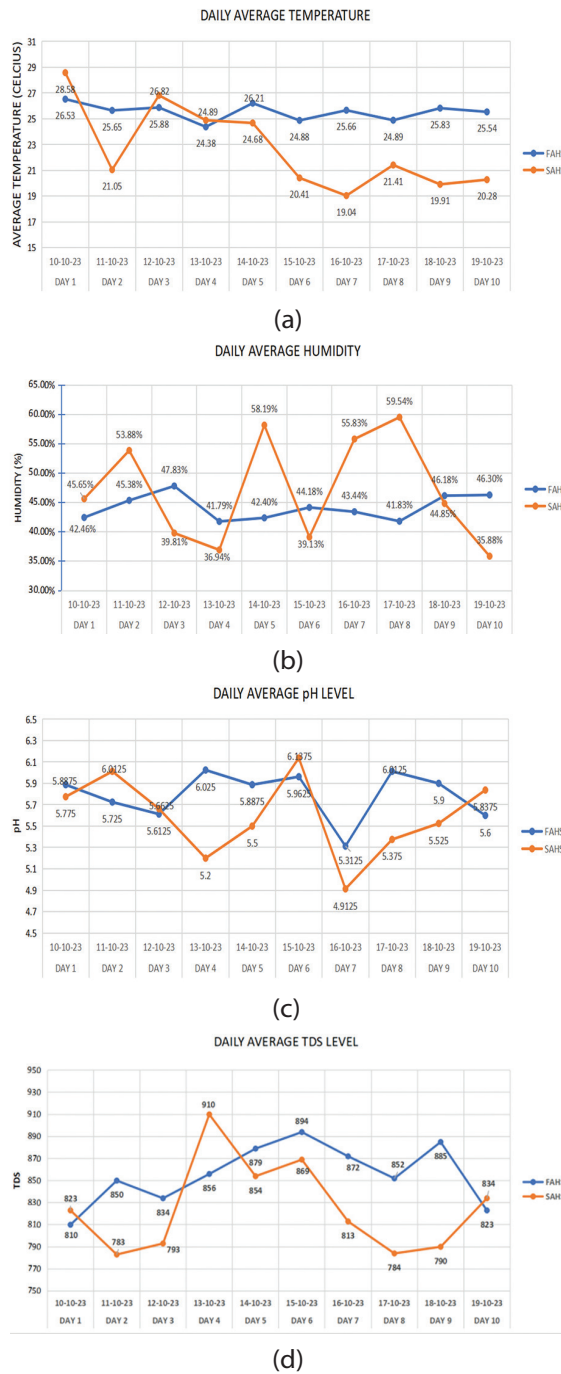
Plant No.	AW - DL (grams)		AW - FL (grams)		RW - DL (grams)		RW - FL (grams)		TLC (leaves)	
	FA	SA	FA	SA	FA	SA	FA	SA	FA	SA
1	3.0	1.9	58.4	46.1	0.35	0.13	3.6	2.3	24	19
2	3.2	1.8	57.9	47.9	0.43	0.18	3.7	2.5	23	22
3	3.1	1.9	57.3	47.4	0.42	0.21	3.9	2.7	25	24
4	2.9	2.2	58.3	47.8	0.29	0.19	2.9	2.4	22	23
5	2.8	2.4	56.9	47.5	0.24	0.17	4.2	2.9	22	20
6	2.6	2.0	57.4	46.7	0.26	0.26	3.5	2.5	24	18
7	3.0	2.3	56.6	46.2	0.31	0.23	3.1	2.5	21	18
8	2.6	1.9	58.2	45.9	0.29	0.21	3.6	2.1	27	19
9	2.9	2.3	55.3	47.3	0.31	0.19	3.3	2.6	26	21
10	3.1	2.9	57.5	47.3	0.33	0.21	2.9	2.3	21	22
Mean	2.9	2.2	57.4	47.0	0.3	0.19	3.5	2.8	23.5	20.6

Every three hours, sensors collected data on the temperature, humidity, pH, and TDS of the nutrient water on the farms over the thirty-day growth cycle of the lettuce plant. This data was then uploaded to the ThinkSpeak platform for analysis. The sensor data obtained from both farms during days 20 to 30 of the growth cycle is shown in Table 3, and Fig. 3 displays the comparative analysis. It is evident from the sensor data in Tables 3 and Fig. 3 as well as the optimal ranges in Table 1 that the automated setup was successful in keeping the temperature, humidity, pH, and TDS within the required optimal ranges of 20 to 25 °C, 40 to 70%, and 800 to 900. With FAHS, there was likewise minor difference between the ambient parameters and the ideal values.

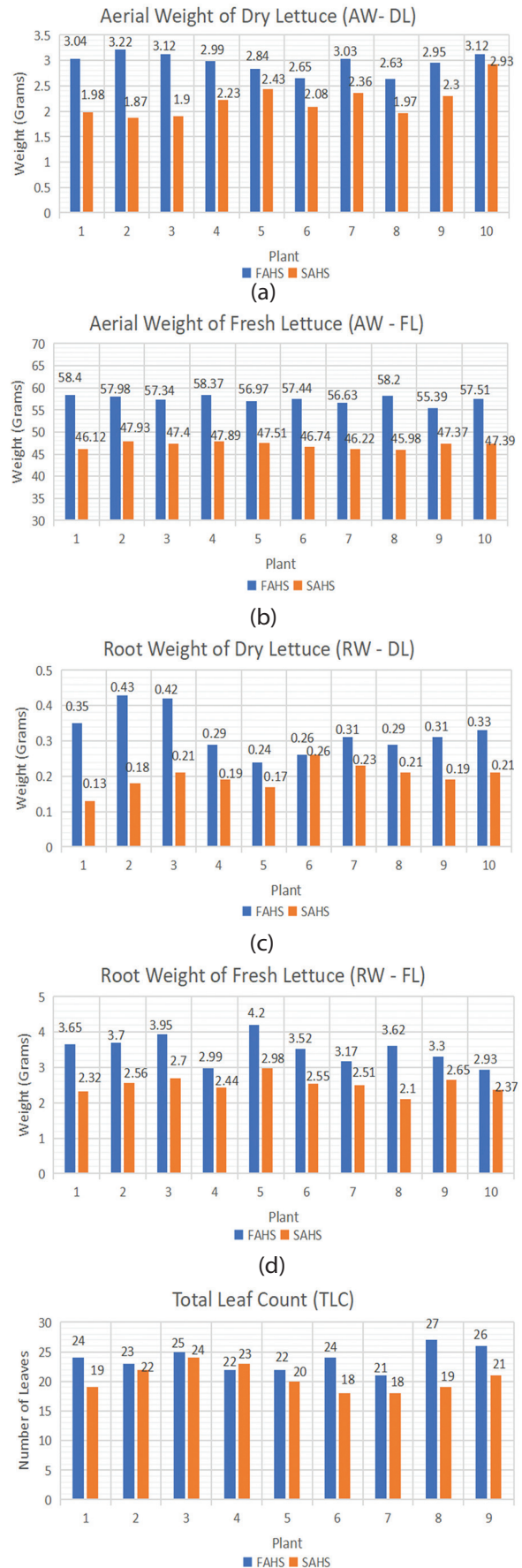
Conversely, the SAHS sensor data shows that there were times when the parameter was outside of the ideal range. Considering the scenarios for day 22, the average humidity of the SA farm was 39.8% which was not in the ideal range, and the next day the average humidity reduced to 36.9%. This clearly shows that the farm owner took no corrective actions to rectify the humidity levels to their ideal range. For example, SAHS's average temperature was 24.68 °C on day 5, 20.41 °C on day 6, and 19.41 °C on day 7. From day 5 to day 7, the farm's mean ambient temperature consistently dropped. Parameters in SAHS frequently and noticeably deviate from the optimal range. For instance, the farm's average humidity was below 40% on days 3, 4, 6,



and 10, which is outside of the recommended range. In contrast, no day in FAHS saw any agricultural parameter fall outside of its ideal range. The reason for this was that FAHS was self-sufficient for maintaining its parameters in the intended range; whenever temperature, humidity, pH, or TDS deviated from the required range, the appropriate actuators would quickly activate and fix the errors. Conversely, in SAHS, human resources had to turn on the relevant equipment manually to fix deviations. The quality of plant growth for FAHS and SAHS, as examined in Fig. 4, showed similar outcomes.



**Fig. 3.** Comparative analysis of the ambient conditions maintained automatically for FAHS and manually for SAHS setup with respect to (A) Temperature, (B) Humidity, (C) pH level and (D) TDS for days 20 to 30 of the plant growth



**Fig. 4.** Comparative Analysis of the plant growth parameters for 10 plants taken from FAHS and SAHS each in terms of (A) AW - DL, (B) AW - FL, (C) RW - DL, (D) RW - FL and (E) TLC

Hence, we can say that fully automated hydroponic systems surpass semi-automated ones in terms of precision, efficiency, and scalability. They continuously monitor and adjust environmental conditions, ensuring optimal plant growth and higher yields while minimizing resource waste. Automation reduces labour costs by minimizing manual intervention and allows for easy scalability of operations. Additionally, these systems provide valuable data for fine-tuning processes, leading to improved long-term productivity and efficiency.

Thus, the practical relevance of automated IoT hydroponic farms is that they offer precise control over environmental variables like temperature, humidity, and nutrient levels, optimizing plant growth and yield. This precision reduces resource wastage and labour costs while increasing productivity and crop consistency. Real-time monitoring and remote control capabilities enhance efficiency and allow for proactive management, minimizing risks of crop failure.

Compared to semi-automated systems, which require more manual intervention and oversight, fully automated IoT farms provide greater scalability, reliability, and potential for integration with data analytics and AI for predictive maintenance and decision-making. Thus, they represent a leap forward in sustainable, high-yield agriculture with lower operational overheads.

## 5. CONCLUSION

New and smart agricultural practices are a key priority in view of population growth, reduced land and water resources available to agriculture and growing demand for food. Hydroponics, a technology used in vertical farms, allows plants to be grown without any soil nutrients. This will decrease the reliance on arable land and water resources for agriculture. As hydroponics creates an artificial growth environment, they require automation technology to keep and preserve it in the correct range. In this experimental study, researchers compare the effects of automation on plant growth.

The practical implications of this research are significant for the agricultural industry. By comparing plant growth in fully automated IoT-based hydroponic farms with semi-automated manual hydroponic farms, we gain insights into the impact of automation on crop productivity and resource efficiency. This can inform farmers, agricultural businesses, and policymakers about the potential benefits of adopting advanced automation technologies in hydroponic farming.

Moreover, the findings may guide investment decisions in agricultural technology, helping stakeholders prioritize resources for implementing automation solutions that maximize yield and sustainability. Additionally, this research opens avenues for future studies to explore the optimization of automation parameters, such as fine-tuning environmental controls and nutrient delivery systems, in order to further enhance plant

growth and resource utilization efficiency in hydroponic farming. We compared the growth outcomes of the two farms using agronomic criteria related to both quantity and quality. By examining the agronomic factors, it was determined that the cases of AW-FL, AW-DL, RW-DL, RW-FL, and TLC had average improvements of 10.4 gm, 0.7 gm, 0.11 gm, 0.7 gm, and 3, respectively. Additionally, they compared the efficiency of the automated technologies in maintaining the artificial growth conditions for both farms. The increased quality of lettuce crops in FAHS resulted from the continuous monitoring, management, and maintenance of all critical environmental factors within optimal limits by automated farms.

The semi-automated farm was neglected multiple times during the night, causing temperature, humidity, pH, or TDS levels to deviate from ideal ranges and impact the quality of lettuce growth. The comparison research has clearly shown that the automated process has a favorable effect on the quality of plant growth while making efficient use of the resources at hand and reducing waste. For this reason, the answer to every issue with conventional agricultural methods is a hydroponic farm outfitted with sensors and actuators to automate farm management.

The results of a study comparing automation's effects on lettuce growth in automated versus semi-automated hydroponic farms have broad implications for agriculture beyond lettuce and hydroponics. Insights gained from the study can be applied to various crops with similar growth patterns and environmental requirements, facilitating crop adaptability across different systems. Additionally, understanding how automation influences growth parameters like plant weight and height can inform practices in both hydroponic and soil-based farming systems.

The findings can guide the adoption of automation strategies in precision agriculture, enhancing efficiency and sustainability across diverse farming methods. By optimizing resource use and minimizing environmental impact of automation, farmers can work towards more sustainable food production practices. Ultimately, the study contributes to advancing agricultural technology and practices, providing valuable insights that can be translated into improved crop outcomes and resource management strategies in agricultural systems worldwide.

Future research could focus on optimizing automation parameters and assessing long-term effects on plant health, nutritional content, and resistance to pests and diseases. Exploring the economic viability and scalability of automated systems across different regions and crop varieties is also essential. Ultimately, this research contributes to advancing agricultural practices towards more efficient, sustainable, and resilient food production systems, addressing pressing challenges in food security and environmental sustainability.

## 6. REFERENCES:

- [1] Nandhini J, Jinu A, Sathian K K, "An IoT - based automated nutrient management in vertical hydroponics", *Pharma Innovation*, Vol. 12, No. 1, 2023, pp. 1954-1957.
- [2] N. C. Eli-Chukwu, "Applications of Artificial Intelligence in Agriculture: A Review", *Engineering, Technology & Applied Science Research*, Vol. 9, No. 4, 2019, pp. 4377-4383.
- [3] L. Lucero, D. Lucero, E. Ormeno-Mejia, G. Collaguazo, "Automated aeroponics vegetable growing system. Case study Lettuce", *Proceedings of IEEE ANDESCON*, Quito, Ecuador, 13-16 October 2020.
- [4] A. Shrivastava, C. K. Nayak, R. Dilip, S. R. Samal, S. Rout, S. M. Ashfaque, "Automatic robotic system design and development for vertical hydroponic farming using IoT and big data analysis", *Materials Today: Proceedings*, Vol. 80, 2023, pp. 3546-3553.
- [5] G. Kaur, P. Upadhyaya, P. Chawla, "Comparative analysis of IoT-based controlled environment and uncontrolled environment plant growth monitoring system for hydroponic indoor vertical farm", *Environmental Research*, Vol. 222, 2023, p. 115313.
- [6] S. Jain, M. Kaur, "Design and Implementation of an IoT-based automated EC and pH Control System in an NFT-based Hydroponic Farm", *Engineering, Technology & Applied Science Research*, Vol. 14, No. 1, 2024, pp. 13078-13081.
- [7] M. Dutta et al. "Evaluation of Growth Responses of Lettuce and Energy Efficiency of the Substrate and Smart Hydroponics Cropping System", *Sensors*, Vol. 23, No. 4, 2023, p. 1875.
- [8] L. Wang et al. "Performance analysis of two typical greenhouse lettuce production systems: commercial hydroponic production and traditional soil cultivation", *Frontiers in Plant Science*, Vol. 14, 2023, p. 1165856.
- [9] M. Mehra, S. Saxena, S. Sankaranarayanan, R. J. Tom, M. Veeramanikandan, "IoT based hydroponics system using Deep Neural Networks", *Computers and Electronics in Agriculture*, Vol. 155, 2018, pp. 473-486.
- [10] Y. L. Teo, Y. I. Go, "Techno-economic-environmental analysis of solar/hybrid/storage for vertical farming system: A case study, Malaysia", *Renewable Energy Focus*, Vol. 37, 2021, pp. 50-67.
- [11] K. K. Y. Shin, T. P. Ping, M. G. B. Ling, C. Chee Jiun, N. A. B. Bolhassan, "SMART GROW - Low-cost automated hydroponic system for urban farming", *HardwareX*, Vol. 17, 2024, p. e00498.
- [12] S. A. B. Anas, R. S. S. Singh, N. A. B. Kamarudin, "Designing an IoT Agriculture Monitoring System for Improving Farmer's Acceptance of Using IoT Technology", *Engineering, Technology & Applied Science Research*, Vol. 12, No. 1, 2022, pp. 8157-8163.
- [13] G. Marques, D. Aleixo, R. Pitarma, "Enhanced Hydroponic Agriculture Environmental Monitoring: An Internet of Things Approach", *Proceedings of the 19<sup>th</sup> International Conference on Computational Science*, Faro, Portugal, 12-14 June 2019, pp. 658-669.
- [14] S. Jain, M. Kaur, "Design and Implementation of an IoT-Based Indoor Hydroponics Farm with Automated Climate and Light Control", *Proceedings of the Fourth International Conference on Computing, Communications, and Cyber-Security*, Ghaziabad, India, 16-17 December 2022, pp. 3-15.

# Improved Parameter Estimation of Three-Phase Squirrel-Cage Induction Motors Using the Nelder-Mead Simplex Algorithm

Original Scientific Paper

## Son T. Nguyen\*

Hanoi University of Science and Technology,  
School of Electrical and Electronic Engineering  
Dai Co Viet Street, Hanoi, Vietnam  
son.nguyenthanh@hust.edu.vn

## Linh V. Trieu

Hanoi University of Science and Technology,  
School of Electrical and Electronic Engineering  
Dai Co Viet Street, Hanoi, Vietnam  
linh.trieuviet@hust.edu.vn

\*Corresponding author

## Tu M. Pham

Hanoi University of Science and Technology,  
School of Electrical and Electronic Engineering  
Dai Co Viet Street, Hanoi, Vietnam  
tu.phamminh@hust.edu.vn

## Anh Hoang

Hanoi University of Science and Technology,  
School of Electrical and Electronic Engineering  
Dai Co Viet Street, Hanoi, Vietnam  
anh.hoang@hust.edu.vn

**Abstract** – This work presents a technique for precisely determining the characteristics of a squirrel-cage three-phase induction motor using the Nelder-Mead simplex algorithm. This approach is a frequently employed numerical optimization technique for determining the minimal value of a multi-dimensional objective function. An advantageous feature of the Nelder-Mead simplex algorithm is its independence from the need to calculate partial derivatives of the objective function. Nevertheless, similar to several optimization techniques, the Nelder-Mead simplex approach can also exhibit sensitivity to the initial conditions. Thus, the initial estimation of the parameters of the approximated equivalent circuit of the induction motor was used as the starting point for the Nelder-Mead optimization approach. The experiment's results are compared to those obtained using the polynomial regression approach to demonstrate the efficacy of the proposed method.

---

**Keywords:** three-phase induction motor, parameter estimation, the Nelder-Mead simplex algorithm

---

Received: January 22, 2024; Received in revised form: June 14, 2024; Accepted: June 26, 2024

## 1. INTRODUCTION

Three-phase induction motors remain prevalent in many industrial applications due to their affordability, exceptional dependability, and uncomplicated design. The operating point of a three-phase induction motor can be determined by analyzing its equivalent circuit. Engineers can utilize the equivalent circuit to forecast the motor's behavior in different operating conditions, without taking into account the specific design characteristics of the motor. Nevertheless, manufacturers do not furnish the corresponding circuit parameters apart from the motor nameplate information, which encompasses the rated voltage, rated output power, efficiency, power factor, and rotor speed.

The three-phase induction motor's equivalent circuit has the following parameters: stator resistance, stator leakage inductance, core loss resistance, magnetizing

inductance, rotor resistance with respect to the stator, and rotor leakage inductance with respect to the stator. The prevalent technique for determining the characteristics of a squirrel-cage induction motor relies on conducting no-load and locked rotor tests. Nevertheless, the locked rotor test poses certain difficulties due to the necessity of accurate voltage regulation and the use of appropriate equipment to securely immobilize the rotor in the locked position.

Manufacturers' nameplate data is valuable basic information for determining the equivalent circuit parameters of three-phase induction motors [1-3]. These efforts focused on developing non-iterative and iterative analytical approaches for solving simultaneous equations based on the motor parameters. Nevertheless, a significant limitation of these approaches is the absence of comprehensive nameplate data for all real



three-phase induction motors. Furthermore, the motor characteristics can only be approximated in the absence of unpredictable variables, such as the alteration of stator and rotor winding resistances caused by an increase in temperature.

Typically, approaches for estimating the parameters of a three-phase induction motor can be categorized into two groups: offline estimation methods [4-9] and online estimation methods [10-12]. Typically, online estimations involve including power converters and additional controllers, followed by integrating the parameter estimation method into the control algorithm. Offline estimating methods necessitate fewer devices or operate without a controller, in contrast to online estimation methods.

The parameters of the three-phase motor can be calculated by utilizing data from various operational states of the motor. A phase-to-phase standstill variable frequency test was conducted in [4] to determine various operating modes of the motor. The test involved minimizing an error function by utilizing the measured resistance and reactance at different frequencies. Only the motors of the double-cage variant are appropriate for this technique. Different values of the load torque can be used to create various working conditions for a three-phase induction motor. A polynomial regression was then employed to solve the matrix equation formed by the partial derivatives of a cost function that is dependent on the input impedance of the motor's equivalent circuit [5, 6].

The motor parameters were computed using optimization algorithms such as the genetic algorithm (GA) [7, 8] and particle swarm optimization (PSO) [9]. The optimization approaches employed in the parameter estimation of three-phase induction motors are very responsive to initial circumstances. Without properly initialized parameters for the motor, the algorithm may produce undesirable outcomes. Currently, the field-oriented control (FOC) and direct-torque control (DTC) approaches are widely used for vector control of three-phase induction motors. If the parameters utilized for the motor do not align with the real motor specifications, the motor may not achieve the desired performance. The d-q model, which is commonly employed in vector induction motor control applications, is extensively utilized in the frequency domain [10]. A Luenberger observer [11] can be used to estimate the state variables and parameters of a three-phase induction motor concurrently. To achieve high-performance speed sensorless control of three-phase induction motors, it is necessary to make online estimations of the precise stator and rotor resistances. This is because these values can experience considerable increases because of the motor's temperature [12].

This paper introduces a comprehensive offline estimation procedure for determining the equivalent circuit parameters of a squirrel-cage three-phase induction motor in a laboratory setting. To determine the

equivalent circuit parameters, it is necessary to measure the voltage, current, and active power of the motor input at various slip rates. The measurement system is straightforward, consisting solely of a single-phase electric meter and a digital rotational speed meter. The Nelder-Mead simplex algorithm [13, 14] is employed to minimize an objective function that is defined based on the discrepancy between the actual and estimated values of the input resistance and inductance of the precise equivalent circuit of the motor. To establish an appropriate beginning condition for the Nelder-Mead method, the initial motor parameters are derived by considering the approximate equivalent circuit of the motor. Subsequently, these characteristics are employed to determine the least value of the goal function associated with the precise equivalent circuit of the motor. This paper also includes a concise explanation of polynomial regression for motor parameter estimation, which serves to validate the efficacy of the proposed method.

The paper is structured in the following manner: Section 2 provides a concise overview of several equivalent circuits for three-phase induction motors and explains the process of determining the motor parameters using the estimated equivalent circuit. Section 3 provides a concise overview of how polynomial regression is employed to calculate the parameters of three-phase induction motors. Section 4 provides a thorough description of the Nelder-Mead simplex algorithm used for estimating the parameters of the three-phase induction motor. To verify the effectiveness of the proposed approach, an experimental setup is created and comprehensively explained in Section 5. The research findings are presented in Section 6.

## 2. EQUIVALENT CIRCUITS OF THREE-PHASE INDUCTION MOTORS

The equivalent circuit of a three-phase induction motor provides vital insights on the motor's performance. The comparable circuit of three-phase induction motors enables engineers to make critical decisions regarding design, operation, and efficiency enhancement. The equivalent circuit of an induction motor closely resembles that of a transformer since both devices facilitate the flow of energy from the primary to the secondary side. In the case of an induction motor, this energy transfer occurs from the stator side to the rotor side. Per-phase equivalent circuits of three-phase induction motors can be classified as follows:

- The exact per-phase equivalent circuit.
- The exact per-phase equivalent circuit without considering the magnetizing resistance.
- The approximate per-phase equivalent circuit.

## 2.1. THE EXACT PER-PHASE EQUIVALENT CIRCUIT

Fig. 1 shows the exact per-phase equivalent circuit of three-phase induction motors.

In which

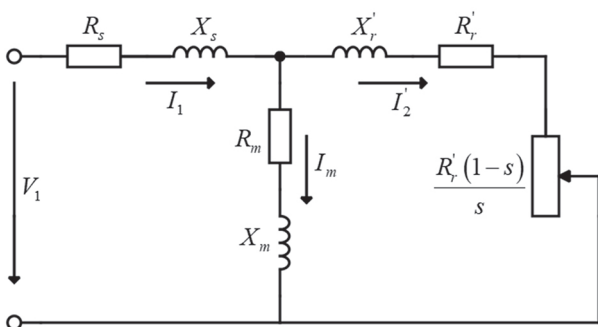
- $R_s$ : The stator winding resistance.
- $X_s$ : The stator leakage reactance.
- $R_r'$ : The rotor winding resistance referred to the stator side.
- $X_r'$ : The rotor leakage reactance referred to the stator side.
- $R_m$ : The magnetising resistance.
- $X_m$ : The magnetising reactance.
- $V_1$ : The stator phase voltage.
- $I_1$ : The stator phase current.
- $I_m$ : The magnetising current.
- $I_2$ : The rotor current referred to the stator side.

To convert the rotor side parameters to the stator side, the rotor side parameters are divided by the square of a value  $K$  known as the effective rotor to stator turns per phase ratio.

$$R_r' = \frac{R_r}{K^2} \quad (1)$$

$$X_r' = \frac{X_r}{K^2} \quad (2)$$

The variables  $R_r$  and  $X_r$  represent the actual rotor resistance and the actual rotor leakage reactance, respectively.  $R_r'((1-s)/s)$  is a resistance that measures the amount of power transferred into mechanical power or the power that is useful.



**Fig. 1.** The exact per-phase equivalent circuit of three-phase induction motors

## 2.2. THE EXACT PER-PHASE EQUIVALENT CIRCUIT IGNORING MAGNETISING RESISTANCE

Often, the magnetizing reactance is significantly more than the magnetizing resistance. The circuit depicted in

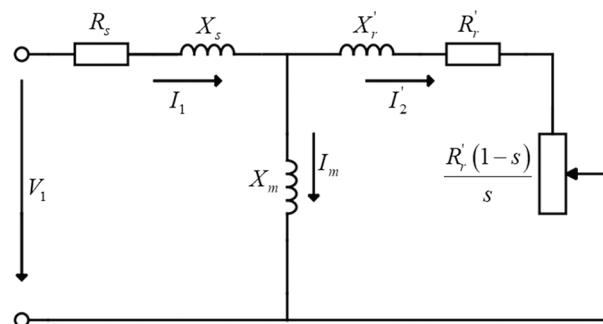
Fig. 2 is obtained from the identical equivalent circuit shown in Fig. 1. The circuit depicted in Fig. 2 offers the benefit of facilitating the estimation of motor parameters through the utilization of the polynomial regression method, which will be discussed in Section 3.

## 2.3. THE APPROXIMATED PER-PHASE EQUIVALENT CIRCUIT

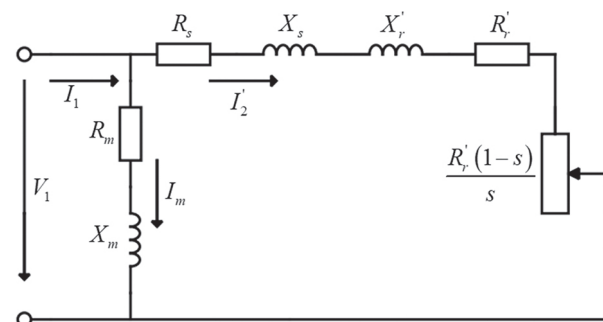
Fig. 3 depicts the estimated per-phase equivalent circuit, which is created by assuming a slight disparity between the supply voltage and the induced voltage. The approximate equivalent circuit is frequently employed to analyse the operational characteristics of three-phase induction motors in electric propulsion systems.

The parameters of the approximate equivalent circuit can be easily determined by conducting a direct current (DC) test, a no-load test, and a load test. The essential procedures for determining the values of the parameters in the approximate per-phase circuit are as follows:

- A DC test is conducted to directly quantify the resistance of the stator winding, denoted as  $R_s$ .
- A no-load test is conducted to calculate  $R_m$  and  $X_m$ .
- A load test is conducted using the data obtained from the no-load test in order to derive the values of  $R_r'$ ,  $X_r'$  and  $X_r'$ .



**Fig. 2.** The exact per-phase equivalent circuit of three-phase induction motors ignoring the magnetising resistance



**Fig. 3.** The approximate per-phase equivalent circuit of three-phase induction motors

In the no-load test, the effects of the winding resistances and inductances are disregarded. Thus, just the component of the magnetizing resistance and induc-

tance has persisted. The input impedance is expressed in the following manner:

$$Z_{in} = R_m + jX_{in} \quad (3)$$

$$R_m \approx R_r \quad (4)$$

$$X_m \approx X_r \quad (5)$$

The magnetising current is calculated as follows:

$$I_m = \frac{V_1}{R_m + jX_m} \quad (6)$$

The current in the winding branch is determined as follows:

$$I_2 = I_1 - I_m \quad (7)$$

The impedance of the winding branch is computed as follows:

$$Z_2 = \frac{V_1}{I_2} \quad (8)$$

The series connection of the resistors and inductors in the winding branch gives:

$$\text{Re}(Z_2) = R_s + \frac{R_r'}{s} \quad (9)$$

$$\text{Im}(Z_2) = X_s + X_r' \quad (10)$$

Using (9),  $R_r'$  is determined as follows:

$$R_r' = s(\text{Re}(Z_2) - R_s) \quad (11)$$

Assuming  $X_s \approx X_r'$  gives:

$$X_s = X_r' = \frac{\text{Im}(Z_2)}{2} \quad (12)$$

### 3. THREE-PHASE MOTOR PARAMETER ESTIMATION USING THE POLYNOMIAL REGRESSION

Polynomial regression can be employed to estimate the motor parameters by establishing the formula for the input impedance of the exact equivalent circuit, without considering the magnetizing resistance illustrated in Fig. 2.

$$Z = R_s + jX_s + \frac{jX_m \left( jX_r' + \frac{R_r'}{s} \right)}{jX_m + \left( jX_r' + \frac{R_r'}{s} \right)} \quad (13)$$

Equation (13) enables us to define an objective function by calculating the discrepancy between the  $n$ -th measured values and the  $n$ -th polynomial values in a straightforward manner.

$$E_n = R_n + jX_n - \frac{(a_0 + a_1s_n + a_2s_n^2) + j(a_3 + a_4s_n^2)}{1 + b_2s_n^2} \quad (14)$$

The  $n$ -th measured value of the slip rate is denoted by  $s_n$ . The coefficients  $a_0, a_1, a_3, a_4$ , and  $b_2$  in equation

(14) can be regarded as intermediate coefficients and are determined by:

$$a_0 = R_s; a_1 = \frac{X_m^2}{R_r'}; a_2 = R_s \frac{(X_m + X_r')^2}{R_r'^2}; a_3 = X_s + X_m \quad (15)$$

$$a_4 = \frac{(X_s + X_m)(X_m + X_r')^2 - X_m^2(X_m + X_r')}{R_r'^2} \quad (16)$$

$$b_2 = \frac{(X_m + X_r')^2}{R_r'^2} \quad (17)$$

Re-arranging (14) gives:

$$E_n = \left[ \frac{R_n(1 + b_2s_n^2) - (a_0 + a_1s_n + a_2s_n^2)}{1 + b_2s_n^2} \right] + j \left[ \frac{X_n(1 + b_2s_n^2) - (a_3 + a_4s_n^2)}{1 + b_2s_n^2} \right] \quad (18)$$

Equation (18) represents a combination of complex numbers. To minimize equation (18), it is necessary to establish an objective function with the following form:

$$S_E = \sum_{n=1}^N \left[ R_n(1 + b_2s_n^2) - (a_0 + a_1s_n + a_2s_n^2) \right]^2 + \left[ X_n(1 + b_2s_n^2) - (a_3 + a_4s_n^2) \right]^2 \quad (19)$$

The concept of reactance distribution is presented by the utilization of the following ratio:

$$\eta = \frac{X_m + X_r'}{X_m + X_s} \quad (20)$$

The range of values for  $\eta$  is between 0.95 and 1.05, and it may take multiple attempts to determine the appropriate ratio. By computing the partial derivatives of with  $S_E$  regard to  $b_2, a_0, a_1, a_2, a_3,$  and  $a_4$ , a matrix equation is derived, as depicted in equation (21). By solving this matrix equation, the values of  $b_2, a_0, a_1, a_2, a_3,$  and  $a_4$  can be calculated. The motor parameters can then be found using equations (22) and (23).

### 4. THREE-PHASE MOTOR PARAMETER ESTIMATION USING THE NELDER-MEAD SIMPLEX ALGORITHM

#### 4.1. OPTIMISATION PROBLEM FOR THREE-PHASE INDUCTION MOTOR PARAMETER ESTIMATION

The variables  $R_s, R_r', X_m, R_r, X_s,$  and  $X_r'$  can be determined by three measurement tests. These variables will then serve as the initial conditions for an optimization procedure aimed at enhancing the estimation of motor parameters. In the context of optimizing induction motor parameter estimates, it is necessary to create an objective function, as depicted in equation (24).

$$\begin{bmatrix} -\sum_{n=1}^N (R_n^2 s_n^4 + X_n^2 s_n^4) & \sum_{n=1}^N (R_n s_n^2) & \sum_{n=1}^N (R_n s_n^3) & \sum_{n=1}^N (R_n s_n^4) & \sum_{n=1}^N (X_n s_n^2) & \sum_{n=1}^N (X_n s_n^4) \\ -\sum_{n=1}^N R_n s_n^2 & N & \sum_{n=1}^N s_n & \sum_{n=1}^N s_n^2 & 0 & 0 \\ -\sum_{n=1}^N R_n s_n^3 & \sum_{n=1}^N s_n & \sum_{n=1}^N s_n^2 & \sum_{n=1}^N s_n^3 & 0 & 0 \\ -\sum_{n=1}^N R_n s_n^4 & \sum_{n=1}^N s_n^2 & \sum_{n=1}^N s_n^3 & \sum_{n=1}^N s_n^4 & 0 & 0 \\ -\sum_{n=1}^N X_n s_n^2 & 0 & 0 & 0 & N & \sum_{n=1}^N s_n^2 \\ -\sum_{n=1}^N X_n s_n^4 & 0 & 0 & 0 & \sum_{n=1}^N s_n^2 & \sum_{n=1}^N s_n^4 \end{bmatrix} \begin{bmatrix} b_2 \\ a_0 \\ a_1 \\ a_2 \\ a_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} \sum_{n=1}^N [R_n^2 s_n^2 + X_n^2 s_n^2] \\ \sum_{i=1}^N R_n \\ \sum_{i=1}^N R_n s_n \\ \sum_{i=1}^N R_n s_n^2 \\ \sum_{i=1}^N X_n \\ \sum_{i=1}^N X_n s_n^2 \end{bmatrix} \quad (21)$$

$$X_m = \sqrt{\frac{\eta a_3 (b_2 a_3 - a_4)}{b_2}}; X_s = a_3 - X_m \quad (22)$$

$$X_r = \eta a_3 - X_m; R_r = \frac{X_m^2}{a_1} \quad (23)$$

$$S = |R_m - R|^2 + |X_m - X|^2 = f(R_s, R_m, X_m, X_s, X_r, R_r) \quad (24)$$

$R_{in}$  and  $X_{in}$  represent the measured input resistance and inductance, respectively, of the exact equivalent circuit of the motor. The input resistance and inductance of the exact equivalent circuit of the motor determine its functionalities. These functions are contingent upon the variables of the motor parameters.

The input impedance of the exact per-phase equivalent circuit can be expressed as a function in the following manner:

$$Z = R_s + jX_s + \frac{(jX_m) \left( \frac{R_r'}{s} + jX_r' \right)}{(jX_m) + \left( \frac{R_r'}{s} + jX_r' \right)} \quad (25)$$

The input resistance and inductance are determined by the functions:

$$R = \text{Re}(Z) \quad (26)$$

$$Z = \text{Im}(Z) \quad (27)$$

#### 4.2. THE NELDER-MEAD SIMPLEX ALGORITHM

The Nelder-Mead simplex algorithm can be elucidated by referencing Fig. 4 and performing the subsequent steps [13]:

1. A simplex has a set of points  $x(i)$ ,  $i=1, \dots, n+1$ .
2. The point in the simplex is ordered from the lowest function value  $f(x(1))$  to the highest function value  $f(x(n+1))$ . In every step of the iteration, the current worst point  $x(n+1)$  are discarded to accepts another point in the simplex.
3. Generate the reflect point:

$$r = 2m - x(n+1) \quad (28)$$

where  $m = \sum_{i=1}^n \frac{x_i}{n}$  and calculate  $f(r)$ .

4. If  $f(x(1)) \leq f(r) < f(x(n))$ ,  $r$  is accepted and this iteration is stopped (**Reflect**).

5. If  $f(r) < f(x(1))$ , the expand point  $s$  is calculated as

$$s = m + 2(m - x(n+1)) \quad (29)$$

and  $f(s)$  is calculated.

- a. If  $f(s) \leq f(r)$ ,  $s$  is accepted and the iteration is stopped (**Expand**).
  - b. Otherwise,  $r$  is accepted and the iteration is stopped (**Reflect**).
6. If  $f(r) \geq f(x(n))$ , a contraction is performed between  $m$  and the better of  $x(n+1)$  and  $r$ .

- a. If  $f(r) < f(x(n+1))$  (i.e.,  $r$  is the better than  $x(n+1)$ ),  $c$  is calculated as

$$c = m + (r - m) / 2 \quad (30)$$

and  $f(c)$  is calculated.

If  $f(c) \leq f(r)$ ,  $c$  is accepted and the iteration is stopped (**Contract outside**).

Otherwise, Step 7 is continued (**Shrink**).

- b. If  $f(r) \geq f(x(n+1))$ ,  $cc$  is calculate as

$$cc = m + (x(n+1) - m) / 2 \quad (31)$$

$f(cc)$  is calculated. If  $f(cc) < f(x(n+1))$ ,  $cc$  is accepted and the iteration is stopped (**Contract inside**).

Otherwise, Step 7 is continued (**Shrink**).

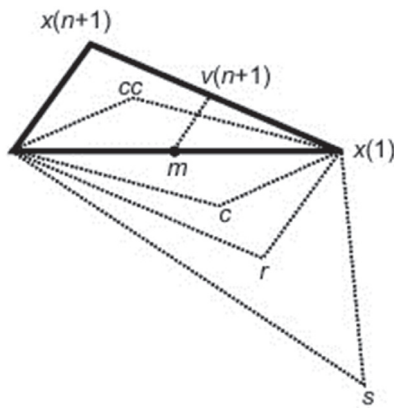
7. The  $n$  points are computed.

$$v(i) = x(1) + (x(i) - x(1)) / 2 \quad (32)$$

$f(v(i))$  is computed,  $i=1, \dots, n+1$ . At the next iteration, the simplex consists of  $x(1), v(2), \dots, v(n+1)$  (**Shrink**).

The figure below illustrates the points that are computed using the Nelder-Mead algorithm during the procedure.



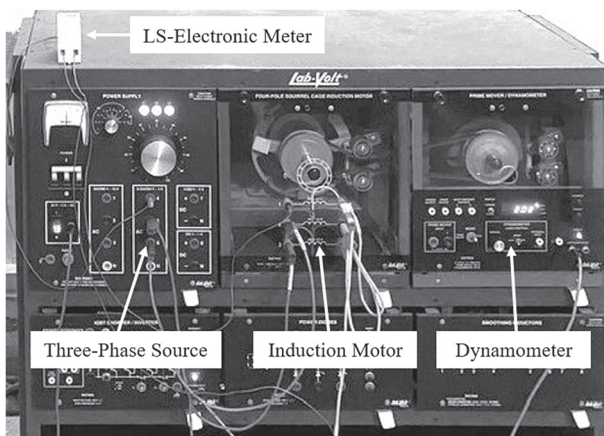


**Fig. 4.** The principle of generating points in a simplex

### 5. EXPERIMENTAL VALIDATION

Fig. 5 displays the experimental setup. The system comprises the following components:

- A 175W squirrel-cage three-phase induction motor manufactured by Lab-Volt company.
- A three-phase voltage source, 50Hz-220/380V, manufactured by Lab-Volt company.
- A dynamometer working as an adjustable load manufactured by Lab-Volt company.
- A single-phase electronic meter manufactured by LS company used to measure parameters of single-phase loads including the voltage, current and active power.



**Fig. 5.** The experimental setup.

#### 5.1. THE POLYNOMIAL REGRESSION METHOD

Table 1 presents the recorded data for the speed, voltage, current, and active power of the motor at various load torque values. The dynamometer can be manually adjusted or modified by an external signal to achieve the appropriate load torque, which can vary between 0 and 1.5 (N.m). The per-phase voltage, current, and real power were recorded for a certain load torque value. Furthermore, the slip rate can also be determined by employing a digital speed meter.

**Table 1.** The collected data includes the rotor speed, per-phase voltage, current, and active power of the motor measured for various load torque values

TL (N.m)	Speed (rpm)	Voltage (Vrms)	Current (Arms)	Active Power (W)
0	1470	227.3	0.32	28
0.1	1457	227.3	0.34	31
0.2	1452	227.2	0.35	37
0.3	1445	226.9	0.37	43
0.4	1436	226.7	0.38	49
0.5	1427	226.6	0.4	55
0.6	1411	226.4	0.42	61
0.7	1404	226.5	0.44	68
0.8	1397	227.3	0.46	73
0.9	1385	227.4	0.49	79
1	1371	227.3	0.51	86
1.1	1360	227.1	0.54	92
1.2	1347	227.2	0.56	99
1.3	1338	227.2	0.59	106
1.4	1324	227.5	0.62	114
1.5	1311	227.5	0.66	125

The motor's speed, voltage, current, actual power, and slip rate were measured and utilized to estimate its electrical properties using the polynomial regression approach. Table 1 indicates that there are a total of 16 distinct instances of the operational states of the motor. The slip rate and input impedance of the motor are calculated using the following formulas:

$$s_n = \frac{n_1 - n_n}{n_1} \quad n = 1, \dots, 16 \quad (33)$$

$$|Z_n| = \frac{V_n}{I_n} \quad n = 1, \dots, 16 \quad (34)$$

$$R_n = \frac{P_n}{|I_n|^2} \quad n = 1, \dots, 16 \quad (35)$$

$$X_n = \sqrt{|Z_n|^2 - R_n^2} \quad n = 1, \dots, 16 \quad (36)$$

The slip rate, resistance, and reactance values of the motor, derived from equations (33), (35), and (36), are utilized in matrix equation (21) to obtain the motor parameters, as shown in Table 2. In this situation, the resistance caused by magnetization is ignored.

**Table 2.** Motor parameters estimated using the polynomial regression approach

$R_s$	$X_s$	$X_m$	$R_r'$	$X_r'$
47.8	57.5390	656.8813	52.9246	57.5390

#### 5.2. THE NELDER-MEAD METHOD

To utilize the Nelder-Mead method for motor parameter estimation, it is necessary to possess an appropriate initial condition. To ensure ease, it is advisable to establish the beginning state by estimating the motor parameters from the approximate equivalent circuit. Under these conditions, it is necessary to do a DC test, a no-load test, and a load test.

## A. DC Test

The DC test was conducted to measure the resistance of the stator winding. To conduct the test, a multi-meter can be used to measure the direct current (DC) resistance of the stator winding  $R_s=49(\Omega)$ .

## B. No-Load Test

The no-load test was conducted to determine the values of the magnetizing resistance  $R_m$  and the magnetizing inductance  $X_m$ . Based on Table 1, the first case can be selected for the no-load test, using the following parameters:

- $V_1=227.3(V)$
- $I_1=0.32(A)$
- $P_1=28(W)$

The magnitude of the input impedance of the analogous circuit is determined by:

$$|Z_m| = \frac{V_1}{I_1} = \frac{227.3}{0.32} = 710.3125(\Omega) \quad (37)$$

The determination of the input resistance of the equivalent circuit is conducted in the following manner:

$$R_m = \frac{P_1}{I_1^2} = \frac{28}{0.32^2} = 273.4375(\Omega) \quad (38)$$

The input inductance of the equivalent circuit is determined by:

$$X_m = \sqrt{|Z_m|^2 - R_m^2} = \sqrt{710.3125^2 - 273.4375^2} = 655.5729(\Omega) \quad (39)$$

The magnetizing resistance and inductance can be determined in the following manner:

$$R_m = R_m = 273.4375(\Omega) \quad (40)$$

$$X_m = X_m = 655.5729(\Omega) \quad (41)$$

The calculation of the magnetizing current is as follows:

$$I_m = \frac{V_1}{R_m + jX_m} = \frac{227.3}{273.4375 + j655.5729} = 0.1232 - j0.2953(A) \quad (42)$$

## C. Load Test

A load test was conducted to measure the rotor winding resistance referred to the stator side  $R_r'$ , as well as the stator leakage inductance  $X_s$  and the rotor leakage inductance referred to the stator side  $X_r'$ . Table 1 indicates that the load test can be one of the scenarios where the load torque is not equal to zero. Under these conditions, the operational state of the motor was evaluated in relation to a load torque of 1.5 (N.m) using the following parameters:

- $V_1=227.5(V)$
- $I_1=0.66(A)$
- $P_1=125(W)$
- $n=1311(\text{rpm})$

The magnitude of the input impedance of the equivalent circuit is calculated as follows:

$$|Z_m| = \frac{V_1}{I_1} = \frac{227.5}{0.66} = 344.6970(\Omega) \quad (43)$$

The circuit's input resistance is determined by:

$$R_m = \frac{P_1}{I_1^2} = \frac{125}{0.66^2} = 286.9605(\Omega) \quad (44)$$

The circuit's input inductance is determined by:

$$X_m = \sqrt{|Z_m|^2 - R_m^2} = \sqrt{344.6970^2 - 286.9605^2} = 190.9703(\Omega) \quad (45)$$

The input current of the equivalent circuit is calculated as follows:

$$I_1 = \frac{V_1}{Z_m} = \frac{227.5}{286.9605 + j190.9703} = 0.5495 - j0.3657(A) \quad (46)$$

The current flowing through the branch of the stator winding and rotor winding referred to the stator side is determined using the following calculation:

$$I_2 = I_1 - I_m = 0.4263 - j0.0703(A) \quad (47)$$

The impedance of the branch of the stator winding and rotor winding referred to the stator side is calculated in the following manner:

$$Z_2 = \frac{V_1}{I_2} = \frac{227.3}{0.4263 - j0.0703} = 519.57 + j85.707(\Omega) \quad (48)$$

The slip rate is calculated as follows:

$$s = \frac{n_1 - n}{n_1} = \frac{1500 - 1311}{1500} = 0.126 \quad (49)$$

The calculation for the rotor winding resistance referred to the stator side is as follows:

$$R_r' = (\text{Re}(Z_2) - R_s) s = (519.57 - 47.8) \times 0.126 = 59.4426(\Omega) \quad (50)$$

The stator leakage reactance and rotor leakage reactance referred to the stator side are calculated using the following formulas:

$$X_s = X_r' = \frac{\text{Im}(Z_2)}{2} = \frac{85.707}{2} = 42.8536(\Omega) \quad (51)$$

The estimated values of  $R_s$ ,  $R_r'$ ,  $X_m$ ,  $X_s$ ,  $X_r'$ , and  $R_r'$  obtained from the approximate equivalent circuit were subsequently employed as the initial parameters for the optimization problem, aiming to directly determine the final values of the motor parameters using the exact equivalent circuit. Fig. 6 depicts the step-by-step process of the Nelder-Mead algorithm during the initial twenty rounds.

Table 3 displays the starting and end values of the motor parameters. Table 4 includes a statement of the comparison between the polynomial approach and the Nelder-Mead method. The Nelder-Mead approach offers the benefit of being able to determine all the parameters of the exact equivalent circuit.

### 5.3. MECHANICAL MOTOR PARAMETER ESTIMATION

The preceding sections provide a detailed explanation of the process for determining the electrical characteristics of an induction motor. The inertia moment  $J$  and the friction factor  $F$  of the motor can be estimated by a dynamic simulation, using pre-defined electrical parameters.

Fig. 7 depicts the simulation of the three-phase induction motor using Simulink. The moment of inertia  $J$  and the coefficient of friction of the motor  $F$  can be approximated by analysing the motor's start-up and the stable condition of the rotor's speed. The parameters of and can be iteratively adjusted multiple times until the simulated and actual motor speeds converge. The ultimate values of the mechanical parameters of the motor are  $J=0.001(\text{kg}\cdot\text{m}^2)$  and  $F=0.0032(\text{N}\cdot\text{m}\cdot\text{s})$ .

Fig. 8 displays the rotor speed and electromagnetic torque during the motor's start-up. Table 5 presents a comparison of the simulated and measured values for the motor's voltage, current, and real power. The simulated values of the motor voltage, current, and real power closely match the measured values of the motor voltage, current, and real power. Table 6 displays the rated motor voltage, current, power factor, and efficiency for a motor with a mechanical power rating of 175 W.

Iteration	Func-count	min f(x)	Procedure
0	1	9348.63	
1	6	9321.64	initial simplex
2	8	4981.55	expand
3	9	4981.55	reflect
4	10	4981.55	reflect
5	11	4981.55	reflect
6	13	2477.1	expand
7	15	393.079	expand
8	16	393.079	reflect
9	18	265.274	reflect
10	20	8.69005	reflect
11	21	8.69005	reflect
12	23	8.69005	contract outside
13	25	8.69005	contract inside
14	27	8.69005	contract inside
15	29	8.69005	contract outside
16	31	8.69005	contract inside
17	32	8.69005	reflect
18	34	8.69005	contract inside
19	36	8.69005	contract inside
20	38	8.69005	contract inside

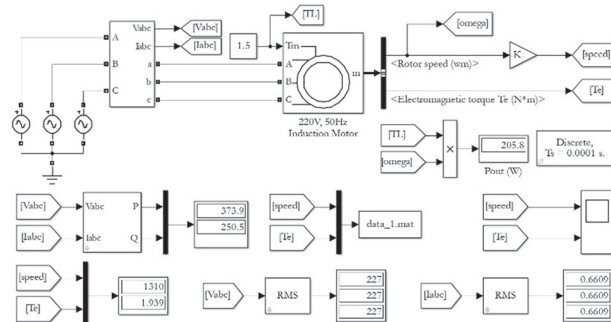
**Fig. 6.** The Nelder-Mead algorithm executed for the first twenty iterations

**Table 3.** Parameters of the exact equivalent circuit estimated using the Nelder-Mead method

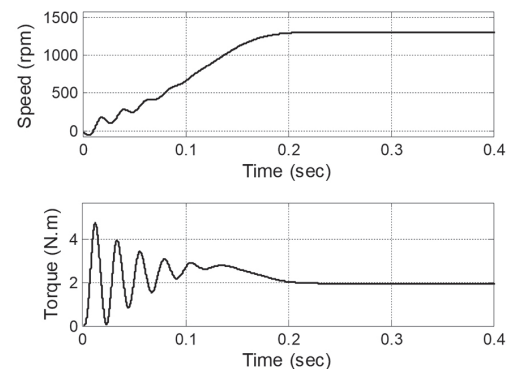
Parameters	Initial Values	Final Values
$R_s (\Omega)$	47.8	53.6589
$X_m (\Omega)$	655.7729	685.8604
$X_s (\Omega)$	42.8536	45.7919
$X_r' (\Omega)$	42.8536	47.5990
$R_r' (\Omega)$	59.4426	39.8770

**Table 4.** Parameters of the exact equivalent circuit estimated using the polynomial regression and the Nelder-Mead method

Parameters	Polynomial Regression Method	The Nelder-Mead Method
$R_s (\Omega)$	47.8	53.6589
$X_m (\Omega)$	656.8813	685.8604
$X_s (\Omega)$	57.5390	45.7919
$X_r' (\Omega)$	57.5390	47.5990
$R_r' (\Omega)$	52.9246	39.8770



**Fig. 7.** Dynamic simulation of the three-phase induction in Simulink



**Fig. 8.** Rotor speed and electromagnetic torque during the motor start-up

**Table 5.** Simulated and measured values of the motor voltage and current

TL (N.m)	Estimated Values		Measured Values	
	V (V)	I (A)	V (V)	I (A)
0	227	0.3287	227.3	0.32
0.1	227	0.339	227.3	0.34
0.2	227	0.3514	227.2	0.35
0.3	227	0.3657	226.9	0.37
0.4	227	0.3818	226.7	0.38
0.5	227	0.3996	226.6	0.40
0.6	227	0.4191	226.4	0.42
0.7	227	0.4401	226.5	0.44
0.8	227	0.4625	227.3	0.46
0.9	227	0.4865	227.4	0.49
1.0	227	0.5118	227.3	0.51
1.1	227	0.5386	227.1	0.54
1.2	227	0.5669	227.2	0.56
1.3	227	0.5966	227.2	0.59
1.4	227	0.6279	227.5	0.62
1.5	227	0.6609	227.5	0.66

**Table 6.** The rated motor voltage, current, real power, power factor, and efficiency

<b>Phase Voltage (V)</b>	<b>220</b>
Phase Current (A)	0.59
Output Power (W)	175
Power Factor	0.78
Efficiency (%)	56.69

## 6. CONCLUSIONS

The study outlines a comprehensive methodology for accurately determining the specifications of a three-phase induction motor that is not already in use. An efficient approach can be employed by utilizing a basic experimental setup and applying the Nelder-Mead simplex algorithm, a derivative-free optimization method. The Nelder-Mead method is initialized using the approximate equivalent circuit of the motor. Afterwards, the parameters of the exact equivalent circuit are estimated by minimizing a cost function that represents the discrepancy between the measured and estimated input impedance of the equivalent circuit. The Nelder-Mead simplex algorithm enables the calculation of all parameters of the exact equivalent circuit of the motor, without relying on assumptions like the reactance distribution ratio. This contrasts with the polynomial regression technique. The method suggested in this paper can also be applied to estimate parameters of single-phase induction motors. In terms of future research, the study intends to investigate the influence of temperature increase in the stator and rotor windings on the estimate of motor parameters required for the efficient control of three-phase induction motors using high-performance vector control.

## 7. ACKNOWLEDGEMENTS

This research is funded by Hanoi University of Science and Technology (HUST) under project number T2023-PC-041.

## 8. REFERENCES

- [1] G. F. V. Amaral, J. M. R. Baccarini, F. C. R. Coelho, L. M. Rabelo, "A High Precision Method for Induction Machine Parameters Estimation From Manufacturer Data", *IEEE Transactions on Energy Conversion*, Vol. 36, No. 2, 2021, pp. 1226-1233.
- [2] K. Lee, S. Frank, P. K. Sen, L. G. Polese, M. Alahmad, "Estimation of induction motor equivalent circuit parameters from nameplate data", *Proceedings of the North American Power Symposium*, Champaign, IL, USA, 9-11 September 2012.
- [3] M. H. Haque, "Determination of NEMA Design Induction Motor Parameters From Manufacturer Data", *IEEE Transactions on Energy Conversion*, Vol. 23, No. 4, 2008, pp. 997-1004.
- [4] L. Monjo, H. Kojooyan-Jafari, F. Córcoles, J. Pedra, "Squirrel-Cage Induction Motor Parameter Estimation Using a Variable Frequency Test", *IEEE Transactions on Energy Conversion*, Vol. 30, No. 2, 2015, pp. 550-557.
- [5] R.-C. Wu, Y.-W. Tseng, C.-Y. Chen, "Estimating Parameters of the Induction Machine by the Polynomial Regression", *Applied Sciences*, Vol. 8, No. 7, 2018, p. 1073.
- [6] S. T. Nguyen, S. Goetz, T. M. Pham, A. Hoang, T. V. Pham, "A Detailed Procedure of Squirrel-Cage Three-Phase Induction Motor Parameter Estimation Using Polynomial Regression", *Proceedings of the 5<sup>th</sup> International Conference on Engineering Research and Applications*, Thai Nguyen, Vietnam, 1-2 December 2022.
- [7] K. S. Huang, Q. H. Wu, D. R. Turner, "Effective identification of induction motor parameters based on fewer measurements", *IEEE Transactions on Energy Conversion*, Vol. 7, No. 1, 2002, pp. 55-60.
- [8] M. Z. Fortes, V. H. Ferreira, A. P. F. Coelho, "The Induction Motor Parameter Estimation Using Genetic Algorithm", *IEEE Latin America Transactions*, Vol. 11, No. 5, 2013, pp. 1273-1278.
- [9] W.-M. Lin, T.-J. Su, R.-C. Wu, "Parameter Identification of Induction Machine With a Starting No-Load Low-Voltage Test", *IEEE Transactions on Industrial Electronics*, Vol. 59, No. 1, 2012, pp. 352-360.
- [10] H. A. Toliyat, E. Levi, M. Raina, "A review of RFO induction motor parameter estimation techniques", *IEEE Transactions on Energy Conversion*, Vol. 18, No. 2, 2003, pp. 271-283.
- [11] S. M. N. Hasan, I. Husain, "A Luenberger-Sliding Mode Observer for Online Parameter Estimation and Adaptation in High-Performance Induction Motor Drives", *IEEE Transactions on Industry Applications*, Vol. 45, No. 2, 2009, pp. 772-781.
- [12] B. Karanayil, M. F. Rahman, C. Grantham, "Online Stator and Rotor Resistance Estimation Scheme Using Artificial Neural Networks for Vector Controlled Speed Sensorless Induction Motor Drive", *IEEE Transactions on Industrial Electronics*, Vol. 54, No. 1, 2007, pp. 167-176.
- [13] J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, "Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions", *SIAM Journal of Optimization*, Vol. 9, No. 1, 1998, pp. 112-147.
- [14] F. Gao, L. Han, "Implementing the Nelder-Mead simplex algorithm with adaptive parameters", *Computational Optimization and Applications*, Vol. 51, 2012, pp. 259-277.





# CRITIC-TOPSIS Method: Design of hybrid renewable energy systems based on multi-criteria decision-making

Original Scientific Paper

## Ade Gafar Abdullah\*

Universitas Pendidikan Indonesia,  
Fakultas Pendidikan Teknologi dan Kejuruan,  
Program Studi Pendidikan Teknik Elektro  
Jl. Dr. Setiabudhi No. 207 Bandung, Indonesia  
ade\_gaffar@upi.edu

## Yusril Nurhikam

Universitas Pendidikan Indonesia,  
Fakultas Pendidikan Teknologi dan Kejuruan,  
Program Studi Pendidikan Teknik Elektro  
Jl. Dr. Setiabudhi No. 207 Bandung, Indonesia  
yusrilnurhikam@upi.edu

## Dadang Lukman Hakim

Universitas Pendidikan Indonesia,  
Fakultas Pendidikan Teknologi dan Kejuruan,  
Program Studi Pendidikan Teknik Elektro  
Jl. Dr. Setiabudhi No. 207 Bandung, Indonesia  
dadanglh@upi.edu

\*Corresponding author

## Nanin Trianawati Sugito

Universitas Pendidikan Indonesia,  
Fakultas Pendidikan Ilmu Pengetahuan Sosial, Program  
Studi Survei Pemetaan dan Informasi Geografis  
Jl. Dr. Setiabudhi No. 207 Bandung, Indonesia  
nanintrianawati@upi.edu

## Diky Zakaria

Universitas Pendidikan Indonesia,  
Program Studi Mekatronika dan Kecerdasan Buatan  
Jl. Dr. Setiabudhi No. 207 Bandung, Indonesia  
dikyzak@upi.edu

**Abstract** – The increasing demand for electricity has driven the development of hybrid renewable energy systems (HRES) in developing countries and remote areas. HRES combines various renewable energy sources to overcome the limitations of each source by utilizing their respective strengths. Although HRES has advantages such as flexibility and low pollution, the development of renewable energy needs to consider environmental impacts and principles of sustainable development. This study aims to find the optimal configuration of a hybrid renewable energy system in the new capital region of Indonesia, namely IKN Nusantara in East Kalimantan province. We conducted a geospatial approach and literature review to gather relevant information, such as solar radiation data, hydropower potential, and biomass potential. We utilized the hybrid optimization model for electric renewables (HOMER) software to design the HRES, and we optimized the results using the multi-criteria decision-making (MCDM) method. The experimental results show that the combination of PV and hydro is the optimal configuration in terms of energy and environmental aspects, with a relative closeness value of 0.675 compared to the ideal solution. Therefore, the implementation of MCDM to evaluate and determine the best configuration of HRES in the IKN Nusantara region is worth considering for further research and development.

---

**Keywords:** hybrid renewable energy system, energy, environment, HOMER, CRITIC, TOPSIS

---

Received: October 23, 2023; Received in revised form: May 2, 2024; Accepted: May 3, 2024

## 1. INTRODUCTION

The growing demand for power, particularly in developing nations and isolated locations with limited electrical access, has pushed the development of hybrid renewable energy systems (HRES) over the past few

decades [1, 2]. The issue of energy scarcity in emerging nations has emerged as a significant global concern in contemporary times. The depletion of fossil fuels and the threat of global warming have compelled the global community to transition towards renewable energy sources [3]. However, the inherent stochasticity of re-

renewable energy hinders it, making it unpredictable and less dependable. Hence, the inclusion of HRES is very important because it can combine numerous renewable energy sources [4], thereby reducing the problems connected with individual sources while taking advantage of the benefits of others [5]. The utilization of renewable energy resources for electricity generation offers a multitude of benefits, including its widespread accessibility, long-term viability, and contribution to environmental conservation. HRES offers numerous advantages, such as versatility, dependability, minimal environmental impact, and the capacity to provide power to distant settlements, major urban areas, and tiny municipalities [6]. Despite the potential for energy creation and their emission-free status, renewable energy sources such as solar power, wind power, and hydropower development still have environmental repercussions. The non-compliance of renewable energy technologies with sustainable development principles can give rise to a range of consequences, encompassing both advantageous and detrimental impacts on human health and climate change. Considering the principles of sustainable development, it is imperative to prioritize the development of renewable energy technologies [7, 8].

The field of HRES has garnered significant interest from researchers, leading to the conduct of numerous investigations [9]. Several prominent studies evaluating HRES using the HOMER software have been carried out. For instance, a study in Iran identified the optimal off-grid HRES model from environmental and economic perspectives using the hybrid optimization model for electric renewables (HOMER) [10]. The researchers conducted a separate study to assess the technical, load, environmental, and economic factors in residential regions, yielding favorable outcomes [11]. The Nigerian Space Research Laboratory has utilized HOMER to evaluate environmental and economic factors [12]. The findings of a study indicate that off-grid HRES have the potential to provide the whole energy requirements of six isolated settlements while also having minimal environmental consequences [13]. In a separate study, the University of Lethbridge employed the particle swarm optimization (PSO) algorithm to optimize the design of an on-grid HRES. This approach yielded a significant reduction in carbon emissions, exceeding 75%, and a power supply outage probability of 5% [14]. Researchers in India found that using the multi-objective genetic algorithm (NSGA-II) in MATLAB software can develop a hybrid system that can meet the electrical needs of rural areas near the Sundarbans region at a cost of 0.1967 \$/kWh. Additionally, this system has the capability to reduce CO<sub>2</sub> emissions by 75,832 kg/year [15]. A separate investigation conducted a comparison between the NSGA-II and Interactive Decision-Making Approach (IDEA) methodologies, revealing that the NSGA-II performed superior optimization in terms of both environmental and economic aspects [16]. Researchers have conducted numerous studies to evaluate the economic and energy

aspects of HRES. However, the energy and environmental aspects of HRES in emerging urban or settlement regions continue to receive insufficient attention.

Numerous scholarly investigations have demonstrated the wide-ranging utilization of HRES in various geographic regions, applying multi-criteria decision making (MCDM) techniques to achieve optimal design and implementation. The energy requirements of an academic township in Sikkim, India, are addressed by an autonomous HRES that incorporates solar, wind, biogas, syngas, and hydrokinetic energy sources. We also utilize batteries as a backup power source. The system successfully attains a competitive levelized cost of energy (LCOE) of 0.095 kWh/kWh by utilizing the analytical hierarchy process [3]. In Kenya, a methodical decision-making framework employing geographic information systems (GIS) and the best worst method (BWM) assesses solar, wind, and hybrid systems for energy access, with a preference for solar/wind/diesel/battery HRES due to its efficiency and cost-effectiveness [17]. Optimized HRES planning is advantageous for airports since it involves the use of MCDM models to estimate the appropriate size of renewable energy and storage devices. This planning takes into account both environmental and social benefits simultaneously [18]. Moreover, Cameroon is implementing grid-connected HRES in residential buildings across various climatic zones. By employing sophisticated optimization techniques, this study aims to identify configurations that exhibit reduced costs and greenhouse gas emissions. These findings have significant potential for addressing energy deficiencies in developing regions. These studies highlight the adaptability and effectiveness of MCDM-based methods in creating HRES customized to particular situations, hence supporting sustainable energy transitions and tackling global energy concerns [19].

The utilization of MCDM techniques plays a crucial role in the development of HRES in diverse settings, as demonstrated by numerous research initiatives. To address the issue of gas flaring in oil and gas fields, a novel approach called combinational MCDM has been developed. This approach incorporates various techniques, such as equal distribution, intuitive distribution, and the Analytical Hierarchy Process (AHP), to determine the weights assigned to different criteria. Additionally, a modified technique for ordering preference by similarity to the ideal solution (TOPSIS) is employed to accurately rank appropriate locations for gas flares [20]. In the context of off-grid power solutions in Malaysia and South Africa, the implementation of a complete approach that incorporates numerical approaches, multi-objective principles, and group decision-making processes is crucial to achieve optimal designs. This approach aims to minimize the probability of load loss and minimize life cycle costs [21]. The ideal configuration for a HRES in Faro-Poli, Cameroon, is determined by MCDM. This process takes into account performance ratings while also considering social and environmental restrictions.

MCDM is advantageous for hospitals as it allows for the identification of a cost-effective mix of PV, WT, and DG systems using the TOPSIS approach [22]. Furthermore, the evaluation of the appropriateness of renewable energy systems in the southern region of the Philippines is conducted through the utilization of fuzzy-AHP and GIS methodologies, with a particular focus on the significance of energy production. The utilization of a multi-objective framework in the operation of on-grid PV-PHS hybrid energy systems exhibits notable dependability and cost-effectiveness, as evidenced by the evaluation of MCDM [23]. This framework facilitates optimal decision-making inside intricate energy systems. These studies jointly emphasize the importance of MCDM approaches in effectively resolving various issues and enhancing the performance of HRES designs in different geographical and operational settings.

The utilization of MCDM techniques has been demonstrated to yield significant advantages in the optimization of decentralized energy systems, with the identification of 1945 best designs for ensuring a dependable, cost-effective, and environmentally sustainable energy supply [24]. MCDM approaches such as TOPSIS, combined with weighing methods, were employed to select cost-effective and ecologically sustainable designs, emphasizing the significance of MCDM in enabling efficient decision-making processes for energy systems [25]. Favourable prospects for remote regions without grid connectivity are presented by decentralized HRES. The study showcases how MCDM assists in identifying the most cost-effective and environmentally friendly solutions [26]. The integration of photovoltaic, wind generators, and lithium-ion technologies proves to be both technically and economically feasible. The prioritization of dependability over economic efficiency in the integration of PV and WT (wind turbine) systems for rural energy distribution leads to enhanced reliability and decreased costs for standalone hybrid renewable energy systems [27]. Moreover, the significance of providing clean energy to underserved locations is emphasized, demonstrating that the combination of PV, WT, bio, and batteries is the optimal solution for areas in Iran. [28] Furthermore, a thorough examination of energy system designs for a geographically isolated rural community in India underscores the significance of MCDM in the process of determining the most advantageous configuration, considering technical, economic, and environmental considerations [29]. In general, the aforementioned findings highlight the significant importance of MCDM in the progression of sustainable energy solutions and the resolution of energy requirements in rural and underprivileged populations [30].

In determining the appropriate renewable energy sources for use in HRES, the decisions made are often not directly related to the core issues. This is due to the complex criteria involved in determining HRES. Currently, there is limited research on the development of HRES design for new development areas, such as the

new capital city of Indonesia located in IKN Nusantara, East Kalimantan province. This is because the current focus of HRES development is mainly on providing energy for remote, off-grid areas. However, there has been relatively little research conducted on the development of grid connected HRES designs with a focus on energy and environmental aspects. Some studies tend to be partial and focus on only one aspect, while others are more centered on economic analysis. Therefore, it is necessary to adopt an approach that combines suitable methods and software when designing grid connected HRES. This is especially important for selecting hybrid renewable energy options suitable for the development of new residential areas, such as the IKN Nusantara region, which is planned to become the new capital city of the Republic of Indonesia. In this research, the contribution lies in the design of HRES using the proposed HOMER software combined with the MCDM CRITIC-TOPSIS (Criteria Importance Through Intercriteria Correlation-Technique for Order of Preference by Similarity to Ideal Solution) method in the grid-connected area of IKN Nusantara, with a specific focus on energy and environmental aspects.

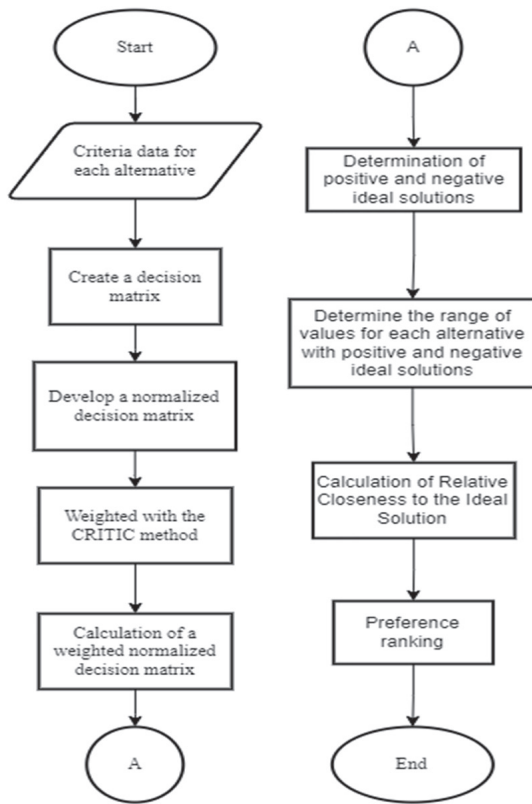
## 2. METHODS

### 2.1. RESEARCH PROCEDURE

The present study encompasses two unique phases, both of which contribute to a full analysis of the adoption of renewable energy. The primary stage of the process is the processing of data using the HOMER tool. During this phase, the procedure begins with a thorough identification of the optimal location for the deployment of the proposed renewable energy system. Following this, a thorough compilation of pertinent data from renowned international journals ensues. The provided data not only serves to inform the project, but also acts as a fundamental resource for making well-informed decisions. Subsequently, a comprehensive evaluation is undertaken to determine the extent of renewable energy capacity within the designated geographical area. Simultaneously, a comprehensive assessment of the electricity demand in the IKN region is being conducted to guarantee that the suggested renewable energy solutions are customized to adequately address the unique requirements of the local community. In addition, robust data validation protocols are implemented to guarantee the precision and dependability of the collected information. The research progresses to the subsequent level alone once the validation of the data has been confirmed. During the second phase, a thorough design process is undertaken to create various combinations of on-grid renewable energy sources. Each combination is carefully tuned to maximize energy production and delivery within the specific region. The phase encompasses the novel element of the study, as it aims to establish the optimal and environmentally friendly approaches to meet the energy requirements of the IKN region.



Within the scope of this study, the evaluation of alternative designs for renewable energy systems (HRES) involves a series of key steps in the application of the TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) algorithm process, as seen in Fig. 1.



**Fig. 1.** TOPSIS procedure

The initial stage of the TOPSIS algorithm involves the identification of criteria that will be utilized for the evaluation of HRES alternatives. These criteria often encompass performance factors, including energy efficiency, environmental effect, cost, and additional considerations. The criteria frequently rely on optimization outcomes obtained by software applications like HOMER, which is utilized for the purpose of designing renewable energy systems. After the establishment of the criteria, the relevant data pertaining to each design choice for the Hybrid Renewable Energy System (HRES) is inputted into the decision matrix. The matrix provides a comprehensive representation of the performance of each alternative across all criteria. To enhance the accuracy of decision-making, the matrix then undergoes a process of assigning weights. During the weighting stage, the CRITIC approach is employed to ascertain the relative weight assigned to each criterion. The methodology employed in this approach is founded upon the intercorrelation among criteria, resulting in the derivation of weights that accurately represent the relative significance of each criterion in relation to the desired objective. The outcome of the weighting phase yields a decision matrix that has undergone both normalization and weighting processes. The matrix presented herein provides a more accurate

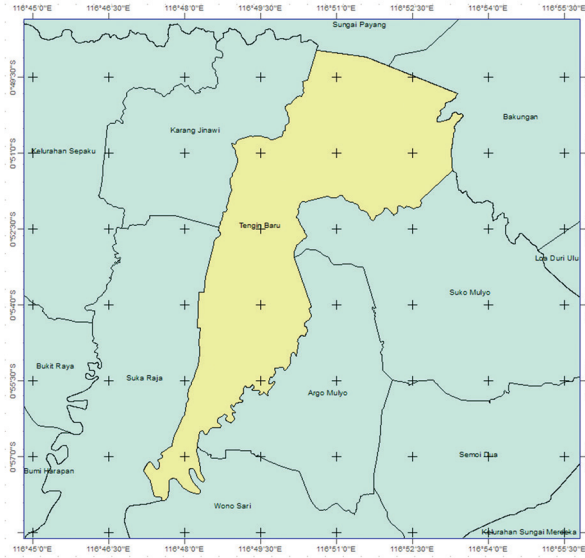
evaluation of the performance of each alternative in relation to pre-established criteria. The matrix also serves as a reflection of the researcher's choices about each criterion. Subsequently, the normalized and weighted matrix is examined to ascertain the positive and negative ideal solutions. The positive ideal solution refers to the optimal combination of values for each criterion, whereas the negative ideal solution represents the converse, namely the least favorable combination of values. By utilizing positive and negative ideal solutions, it is possible to determine the proximity of each alternative to the positive ideal solution and the distance it deviates from the negative ideal solution. These values offer a more comprehensive perspective on the degree to which each alternative aligns with the researcher's preferences. In general, the performance of an alternative tends to improve as it approaches the positive ideal solution and moves away from the negative ideal solution.

## 2.2. STUDY AREA

IKN Nusantara, situated in the North Penajam Paser and Kutai Kartanegara Regencies of East Kalimantan, is designated as the future National Capital. This designated area encompasses a total expanse of 256,142.74 hectares, with the central city covering 56,180.87 hectares, and the administrative center occupying 5,664 hectares [31, 32]. The development of IKN Nusantara is expected to pose environmental challenges, which the government plans to address by implementing the concepts of a smart city and forest city. These concepts aim to create a sustainable and eco-friendly urban environment by harmonizing nature and technology, thereby preventing ecological degradation resulting from development [33]. Furthermore, there's a strong emphasis on utilizing renewable energy sources to meet the energy demands of IKN Nusantara. The East Kalimantan Province boasts a significant renewable energy potential, estimated at around 23,841 MW [34]. To cater to the power needs of the research site, a priority is placed on harnessing locally available renewable energy sources. This approach aligns with the core principle of IKN Nusantara's development - the integration of nature and technology. Solar energy, hydropower, wind energy, and biomass represent a few examples of renewable energy sources that can be effectively deployed within IKN Nusantara. Leveraging locally sourced renewable energy not only reduces reliance on nonrenewable resources but also promotes a path toward sustainable development.

The researchers opted for a site situated near the Sepaku Semoi Dam, with precise coordinates at 0°54'32.01" Southern Latitude and 116°50'25.72" East Longitude, located in Tengin Baru Village, Sepaku District, North Penajam Paser Regency. As seen in Fig.2, this location was chosen due to its substantial renewable energy potential and a land area spanning 6,230.87 hectares, rendering it suitable for the establishment

of a hybrid power plant integrating hydro power with other renewable energy sources. Furthermore, the selection of this site was influenced by its proximity to the zero point of IKN Nusantara, located approximately 16.3 kilometers away.



**Fig. 2.** The research site map

### 2.3. CRITIC METHODS

The Criteria's Significance the through inter criteria correlation (CRITIC) technique is an analytical method that employs correlation analysis to investigate significant decision criterion specifications. This method can determine the appropriate weight by employing contrast strength and conflicting criteria [35]. The initial step in applying the CRITIC method is to normalize the initial decision matrix using the following equation:

$$r_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \quad (1)$$

Using the standard deviation and the correlation between one criterion and other criteria, determine the weight of the  $w_j$  criteria using the following equation:

$$w_j = \frac{C_j}{\sum_{i=1}^m c_i} \quad (2)$$

Where,  $C_j$  is the quantity of data contained in the  $j$ -th criterion established in:

$$C_j = \sigma_j \sum_{i=1}^m (1 - r_{ij}) \quad (3)$$

Where,  $\sigma_j$  the standard deviation of the  $j$ -th criteria and  $r_{ij}$  is the correlation coefficient of the  $j$ -th and  $i$ -th criteria. A higher  $C_j$  value indicates that more information is obtained from a given criterion, thereby indicating that the criterion has a higher level of significance in the decision-making process [36].

### 2.4. TOPSIS METHODS

TOPSIS was chosen as the MCDM method due to its efficacy in decision-making processes for energy systems.

The TOPSIS method is well-suited for scenarios in which it is necessary to simultaneously consider numerous criteria to assess and prioritize different solutions. TOPSIS enables the optimization of decentralized energy systems by considering multiple criteria, including cost-effectiveness, environmental sustainability, and technological feasibility. The study seeks to determine the most advantageous designs that achieve a balance between these requirements by utilizing TOPSIS in conjunction with weighing methodologies. Therefore, the selection of TOPSIS was based on its capacity to offer thorough assessments and enable well-informed decision-making in intricate energy system optimization endeavors.

TOPSIS is a practical and valuable technique for ranking and selecting alternative options based on Euclidean distance that arises from various multicriteria decision making techniques. It is a straightforward, conceptually, and practically simple ranking method. This methodology's objective is to identify ideal solutions and negative ideal solutions to compare the relative performance of alternatives and to identify alternatives that are closer to the ideal solution. The first step in implementing this strategy is to create a normalized decision matrix. [37]:

$$r_{ij} = \frac{x_{ij}}{\sqrt{\sum_i^m x_{ij}^2}} \quad (4)$$

Then calculate the normalized and weighted matrices:

$$v_{ij} = w_j r_{ij} \quad (5)$$

The positive ideal solution is derived from the optimal performance of the normalized and weighted matrix to formulate ideal and anti-ideal solutions:

$$A^* = \{v_1^*, v_2^*, \dots, v_n^*\} = \begin{cases} \max_j & v_{ij} | i \in I' \\ \min_j & v_{ij} | i \in I'' \end{cases} \quad (6)$$

$$A^- = \{v_1^-, v_2^-, \dots, v_n^-\} = \begin{cases} \min_j & v_{ij} | i \in I' \\ \max_j & v_{ij} | i \in I'' \end{cases} \quad (7)$$

Determine the  $D_i^*$  and  $D_i^-$  value values for each alternative using the formulas:

$$D_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2} \quad (8)$$

$$D_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (9)$$

Lastly, the Relative Closeness to the Ideal Solution calculation:

$$C_i^* = \frac{D_i^-}{D_i^* + D_i^-} \quad (10)$$

$C_i^*$  is always between 0 and 1 and an alternative is best if it is close to 1. This is calculated for each alternative and is defined as the above equation [38].

### 3. DATA AND HOMER DESIGN INPUT

#### 3.1. DATA INPUT

Before starting to optimize the design of the HRES, there is an important first step that needs to be taken: gathering and analyzing data about the potential of solar radiation, the viability of water sources, the availability of biomass, and the needs of the electrical load. This information is very important for making sure that the HRES design can meet the location's energy needs effectively. If we know how much potential solar radiation there is, we can guess how much solar energy can be turned into electricity. In the same way, looking at how easy it is to get to water sources and biomass, along with electrical load data that accurately shows energy needs, gives a full picture of the resources available and the energy needs.

Indonesia boasts significant potential for the advancement of solar power as a viable alternative energy source, owing to its equatorial location and the abundant solar radiation it receives. The East Kalimantan region, in particular, benefits from an exceptional level of solar radiation, rendering it ideally suited for the efficient harnessing of solar energy. Within this geographical area, solar energy stands out as the predominant renewable energy source, constituting a substantial 74.89 percent of East Kalimantan's overall renewable energy potential. [39, 40].

The data provided in Table 1 applies to the solar energy capacity in the village of Tengin Baru Sepaku, which is one of the chosen locations within the IKN Nusantara region. The data utilized in this study was obtained from NASA POWER and encompasses information regarding the mean solar radiation levels seen during the months of January to December during a duration of 21 years, specifically spanning from 2000 to 2021. The average solar radiation possesses the capability to generate a monthly electricity output of 4,661 kWh/m<sup>2</sup>/day.

**Table 1.** Average solar radiation

Month	Daily Radiation (kWh/m <sup>2</sup> /day)	Month	Daily Radiation (kWh/m <sup>2</sup> /day)
January	4.555	July	4.481
February	4.914	August	4.699
March	4.898	September	4.775
April	4.800	October	4.608
May	4.633	November	4.609
June	4.406	December	4.563
<b>Average</b>	<b>4.661</b>		

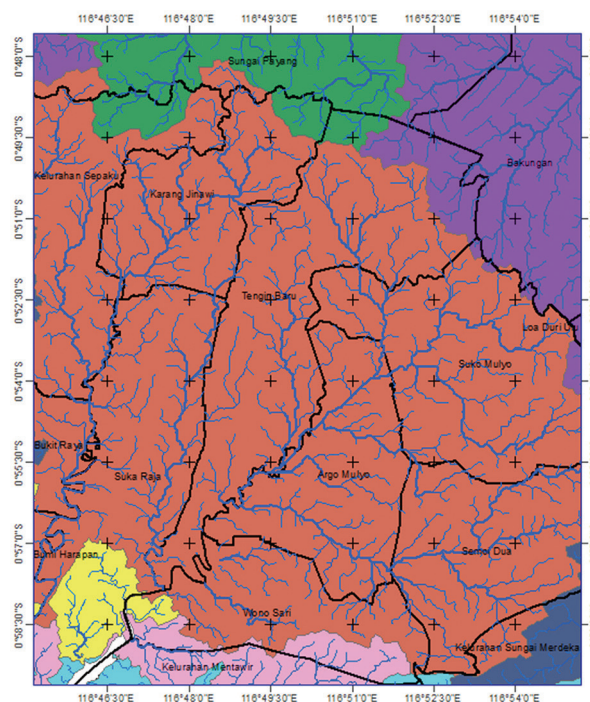
East Kalimantan, characterized by its abundant river flow and substantial water discharge, presents a highly auspicious prospect for the development of hydroelectric power plants. Additionally, the government's National Energy General Plan (NEGP) supports this claim, saying that East Kalimantan has a large energy potential of 13,479 MW thanks to its many water and solar

resources [41, 42]. Within the scope of this study, we leverage the region's substantial water resources, with a particular focus on harnessing the water discharge from the Sepaku Semoi Dam for the establishment of a hydroelectric power plant.

**Table 2.** Water discharge of the Sepaku Semoi dam

Month	Discharge m <sup>3</sup> /s	Month	Discharge m <sup>3</sup> /s
January	2.21	July	3.07
February	2.06	August	2.35
March	1.75	September	2.17
April	2.78	October	1.96
May	3.26	November	1.99
June	3.58	December	2.21
<b>Average</b>	<b>2.45 m<sup>3</sup>/s</b>		

Table 2 provides data indicating that the maximum water discharge occurred in June at 3.58 m<sup>3</sup>/s, while the minimum discharge was observed in March at 1.75 m<sup>3</sup>/s. The annual mean discharge stands at 2.45 m<sup>3</sup>/s. Based on precipitation calculations spanning the previous two decades, it is ascertained that water availability remains consistently adequate. Given a hydraulic head of 15.85 meters, the total electricity generation potential amounts to 2,593,140 kWh annually [43].



**Fig. 3.** Tengin Baru village river flow

Fig. 3 presents a comprehensive visualization of the river flow distribution within Tengin Baru Village. This visualization was meticulously analyzed using the robust ArcGIS software. The primary objective of this analysis was to precisely evaluate and quantify the hydropower potential inherent to the specific study site under investigation. The utilization of ArcGIS software allowed for a thorough examination of the geographic



and hydrological factors influencing the hydropower prospects in the area, enhancing the accuracy and reliability of the results obtained.

Indonesia, recognized as one of the world's leading palm oil producers, grapples with a significant volume of palm oil waste. Based on the results of the thorough survey, there is a huge potential for making electricity from palm oil waste, which adds up to an impressive 12,654 megawatts (MW). Among the regions assessed, Sumatra exhibited the highest potential, with a capacity of 8,812 MW, while Kalimantan demonstrated the lowest potential at 3,384 MW. This remarkable revelation underscores the immense opportunity to harness palm oil residue as a sustainable and renewable energy source within the region [44]. Moreover, Table 3, which is presented herewith, offers a detailed breakdown of palm oil waste production in the East Kalimantan Province, a province within the Kalimantan region. This dataset was collected in 2019, revealing that ten regencies and local communities in East Kalimantan collectively generate approximately 800 million tons of palm oil solid waste annually [45].

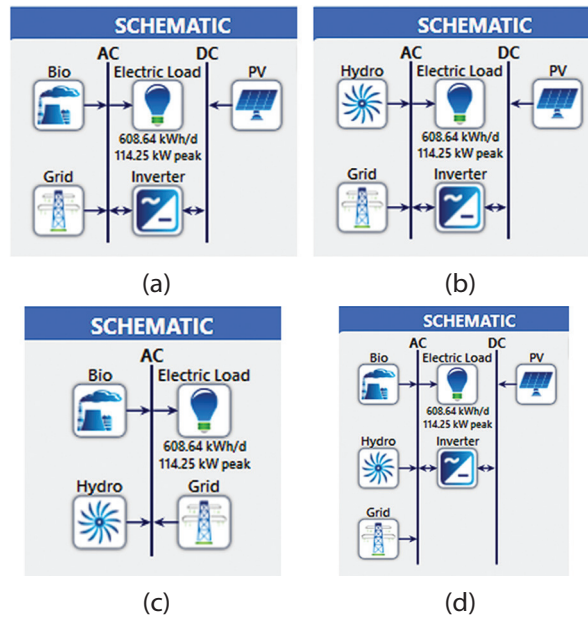
**Table 3.** Production of palm oil waste

No	Districts	TKKS (ton)	Serat (ton)	Cangkang (ton)
1	Kutai Kartanegara	808265	464323	171971
2	Kutai Timur	1866382	1072177	397103
3	Kutai barat	263456	151347	56.055
4	Mahakam Hulu	29871	1716	6356
5	Penajam paser utara	219584	126144	46720
6	Paser	563094	323479	119807
7	Berau	557463	320245	118609
8	Samarinda	2501	1437	532
9	Balikpapan	110	63	23
10	Bontang	79	45	17

### 3.1. HRES DESIGN USING HOMER

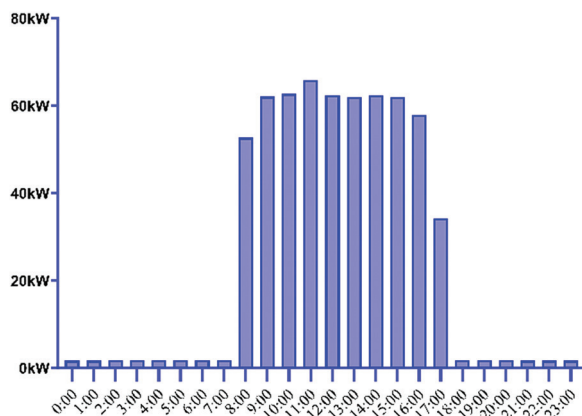
The research site is where the modeling and simulation of the HRES takes place. This is done by gathering and analyzing relevant data. The visual representation in the Fig.4. showcases four distinct HRES design models, each meticulously developed using HOMER software. While these design schemes share common load requirements and component ratings, their uniqueness lies in the combination of renewable energy sources employed within each design. Design 1 embrace a blend of biomass and photovoltaic (PV) sources, whereas the design 2 integrates hydropower with PV. In the Design 3, biomass and hydropower come together, and the design 4 stands out by combining hydropower, biomass, and PV sources. These diverse combinations of renewable energy sources have been strategically chosen to optimize the simulation outcomes, with a primary focus on enhancing energy efficiency, promoting sustainability, and mitigating the environmental impact of the energy generation process. By running these simulations, the researchers hope to find

the best HRES configuration—one that not only meets the site's energy needs but also fits with the main goals of reducing damage to the environment and promoting long-term energy solutions. Fig. 4 shows the four HOMER designs in this study.



**Fig. 4.** HOMER Design (a) PV-Biomass (b) PV-Hydro (c) Biomass-Hydro (d) PV-Biomass-Hydro

The load profile implemented at the research site is initially assumed to be representative of the Bali Provincial DPRD (Regional People's Representative Council) office building [46] (see Fig.5).



**Fig. 5.** Daily load profile for the research location

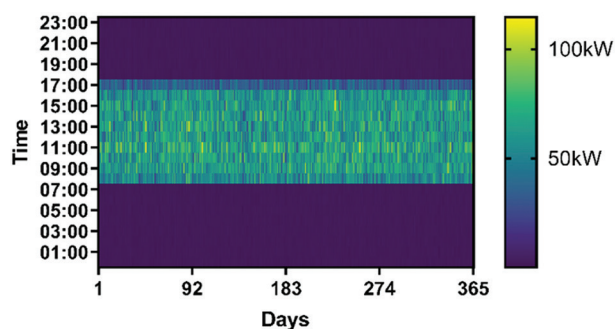
This assumption is based on the premise that such an office facility may offer a reasonable approximation for the energy usage patterns encountered in the area under study. However, to enhance the accuracy and relevance of the load profile, further adjustments are made to align it with the specific daily load characteristics observed at the Office of Highways and Irrigation located in Badung Regency [47]. This meticulous calibration process ensures that the load profile accurately reflects the real-world energy consumption patterns



observed in the study location. This considers the Office of Highways and Irrigation's unique operational needs and scheduling quirks, which may be different from those of the Bali Provincial DPRD Office building. This fine-tuning of the load profile is essential for the reliable modeling and simulation of the hybrid renewable energy system (HRES) under consideration, facilitating more precise and context-appropriate results.

Figure 5 shows the daily load data used in the IKN region in a visual way, which helps us understand how energy use changes over time. On average, the daily energy consumption stands at a noteworthy 608.64 kWh. Notably, a discernible spike in electricity usage is observed during the hours spanning from 08:00 to 17:00. This time frame corresponds to the conventional working hours, and as a direct consequence, it leads to a substantial surge in energy demand. This data reveals the prominent role of daily work routines in shaping the energy consumption pattern within this region.

The annual load profile at the study site is depicted in Fig. 6. By analyzing this load profile, trends, and patterns of energy consumption over a certain period can be identified, forming the premise for designing and optimizing renewable energy systems that can efficiently meet energy demands.



**Fig. 6.** Annual load profile for research site

The components used in this study have been adapted to the potential and environment of the research site. Where the following components are used:

- The Tirani Solar Vertex 500 WP is a flat panel with a DC current output and a capacity of 0.5 kW. The initial investment and replacement costs for this component are USD 1726,74. The operational and maintenance expenses are 20% of the initial investment, or USD 50 [48]. This PV component has a depreciation factor of 80% and a useful life of 25 years.
- Francis's turbine model HL110-WJ-30 This turbine has a minimum head requirement of 10 m and a capacity of 10 kW. The initial investment cost and replacement cost of this water turbine are USD 4759,60, and operational and maintenance costs are 12% of the initial investment, which is USD 571,15 [49, 50].

- A 15-kW biogas generator with the model number 15GFT-J. With an initial investment cost and replacement cost of USD 6712,25, this generator has operational and maintenance costs of USD 50 [51]. A generator with a lifespan of 20,000 hours.
- 50-250kW Growatt inverter with 99% efficiency and a 15-year service life. Initial investment and replacement expenses total USD 4116,51 while operational and maintenance expenses total IDR USD 82 [52].
- The grid is used to sell excess electrical energy produced by hybrid generators so that it is not wasted. With an electricity purchase price of USD 0,098 per kWh and a selling price of electricity to PLN of USD 0,075 per kWh [53].

## 4. RESULTS AND DISCUSSION

### 4.1. RESULTS

After simulating alternative designs using HOMER software, results are obtained that include various parameters related to economic, energy, and environmental aspects. However, in the context of this study, the main emphasis is placed on energy and environmental aspects. The results of the optimization of the four designs that have been made are then presented in detail in Table 4. Furthermore, these data are used as input to calculate the weight of each criterion using the CRITIC method.

**Table 4.** Homer Pro optimization results for four HRES designs

Criteria	System configuration			
	Design 1	Design 2	Design 3	Design 4
C1: Renewable Fraction (%)	85.7	93.4	79.9	94.4
C2: Max Renewable Energy Penetration (%)	111	109	100	108
C3: Production (Kwh/Y)	266613	377188	276704	345762
C4: Energy Sold (Kwh)	86965	179780	124160	143044
C5: Energy Purchased (Kwh)	44201	26385	69609	20310
C6: Carbon Dioxide (Kg/Y)	27993	16675	44022	12858
C7: Sulfur Dioxide (Kg/Y)	121	72.3	191	55.6
C8: Nitrogen Oxides (Kg/Y)	59.6	35.4	93.5	27.4

At this stage, the CRITIC method is applied to determine the weight of the criteria. The first step in this process is to create a normalized decision matrix using equation (1). Then the standard deviation values for each criterion will be obtained as follows:  $\sigma_j = (0.47201, 0.43913, 0.48453, 0.41661, 0.4486, 0.44882, 0.44848, 0.44922)$ .

After the standard deviation is obtained, determine the value of the correlation coefficient. The final step in the decision-making process, where the weights for each criterion are determined using equations (2) and (3). The weight calculation results for each criterion are then presented in Table 5.

**Table 5.** Weight of each criterion

C1	C2	C3	C4	C5	C6	C7	C8
0.0727	0.2341	0.146	0.2464	0.0752	0.0751	0.0753	0.0749

The decision matrix normalization table shows how different designs perform against the criteria. Design 4 has the highest scores in most criteria (C3, C4, C5, C6, C7, and C8), as shown in the Table 6. Due to its better energy production, energy sales, and greenhouse gas emissions, Design 4 may be a better choice. Design 2 has high energy production (C4) but low purchased energy (C5). Design 3 has a high purchased energy (C5) but lower renewable fraction and maximum renewable energy penetration. Design 1, which scores well in some criteria, has the lowest production score (C3), suggests it may not produce energy efficiently. The final decision about which design is best must take weighting criteria and decision-makers' preferences into account.

**Table 6.** Normalization of the decision matrix

	C1	C2	C3	C4	C5	C6	C7	C8
Design 1	0,484	0,518	0,417	0,316	0,497	0,498	0,496	0,498
Design 2	0,527	0,509	0,589	0,688	0,342	0,342	0,342	0,341
Design 3	0,451	0,467	0,432	0,655	0,960	0,960	0,960	0,960
Design 4	0,533	0,504	0,540	1,000	1,000	1,000	1,000	1,000

Table 7 displays the weighted normalization matrix, in which the corresponding criteria weight multiplies each value in the decision matrix. The criteria weights indicate the importance of each criterion in the decision-making process. By examining this table, we can observe the relative contribution of each criterion to the overall assessment of each design. From the table, criterion C4 has the highest weight among all criteria for all designs. This indicates that C4 has a significant influence on the final assessment. Furthermore, criterion C2 also holds a relatively high weight for all designs, suggesting its importance in the decision-making process.

**Table 7.** Construction of weighted normalization matrix

	C1	C2	C3	C4	C5	C6	C7	C8
Design 1	0,035	0,121	0,061	0,078	0,037	0,037	0,037	0,037
Design 2	0,038	0,119	0,086	0,169	0,026	0,026	0,026	0,026
Design 3	0,033	0,109	0,063	0,161	0,072	0,072	0,072	0,072
Design 4	0,039	0,118	0,079	0,246	0,075	0,075	0,075	0,075

Table 8 provides the positive ideal solution (V+) and negative ideal solution (V-) for each criterion. The positive ideal solution represents the best possible values for each criterion, while the negative ideal solution

represents the worst possible values. From Table 8, we can observe that the positive ideal solution (V+) has higher values compared to the negative ideal solution (V-) for all criteria. This indicates that higher values are desirable for criteria such as C1, C2, C3, and so on, while lower values are desirable for criteria such as C4, C5, and others.

**Table 8.** Identify positive ideal solutions and negative ideal solutions

	C1	C2	C3	C4	C5	C6	C7	C8
V+	0,039	0,121	0,086	0,246	0,075	0,075	0,075	0,075
V-	0,033	0,109	0,061	0,078	0,026	0,026	0,026	0,026

Table 9 provides the Euclidean distance (S+) from each alternative to the positive ideal solution (V+), the Euclidean distance (S-) from each alternative to the negative ideal solution (V-), and the proximity score (Pi) calculated based on these distances. The proximity score represents how close each alternative is to the positive ideal solution relative to the negative ideal solution. From Table 9, we can see that Design 4 has the shortest Euclidean distance to the positive ideal solution (S+), indicating that it is closest to the ideal solution among all alternatives. On the other hand, Design 1 has the longest Euclidean distance to the positive ideal solution. Similarly, Design 4 has the shortest Euclidean distance to the negative ideal solution (S-), while Design 1 has the longest Euclidean distance.

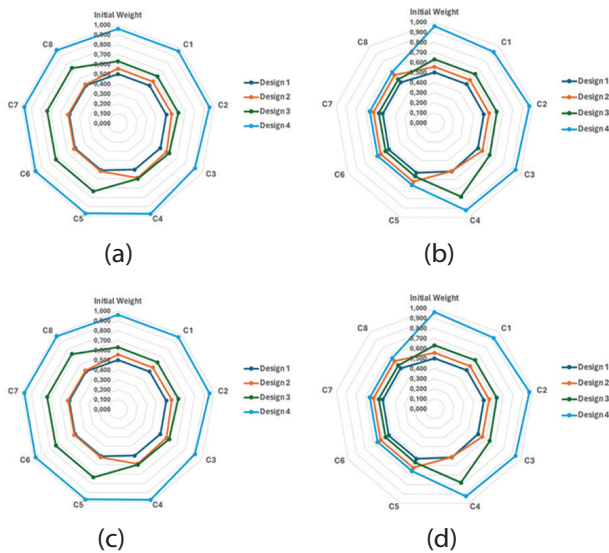
**Table 9.** Calculation of Euclidean distance, proximity score, and alternate ranking

	Design 1	Design 2	Design 3	Design 4
S+	0,1862	0,1251	0,0890	0,0079
S-	0,1880	0,1574	0,1534	0,1963
Pi	0,5025	0,5572	0,6328	0,9614
Alternative ranking	4	3	2	1

The proximity scores (Pi) are then calculated by dividing the distance to the positive ideal solution by the sum of the distances to both the positive and negative ideal solutions. Better performance relative to other alternatives is indicated by a higher proximity score. Based on the proximity scores, the highest score is attained by Design 4, indicating that it is the most preferable alternative among all designs. Conversely, the lowest proximity score is achieved by Design 1, suggesting that it is the least preferable alternative.

A sensitivity analysis was performed to verify the outcomes of decision-making using the CRITIC-TOPSIS approach. The experiment was conducted by manipulating the weight of each criterion by 25% and 50%, respectively. The outcomes of this manipulation are depicted in Fig. 7. The importance of sensitivity analysis in this work lies in its ability to provide a deeper under-

standing of how changes in the weight of evaluation criteria affect the choice of HRES design. This allows for the identification of designs that exhibit greater resilience to change, facilitates the exploration of trade-offs across criteria, and incorporates uncertainty into the decision-making process. Through the utilization of sensitivity analysis, the decision-making process can be enhanced, and multiple pertinent factors can be considered when implementing renewable energy.



**Fig. 7.** Sensitivity to changes in performance score: (a) Experiment 1: 25% weight gain; (b) Experiment 2: 25% weight reduction; (c) Experiment 3: 50% weight gain; and (d) Experiment 4: 50% weight reduction

Different responses to weighting the evaluation criteria were observed for each design. In Experiment 1, with a 25% weight gain, Design 1 was found to be performing consistently, while renewable energy penetration and production were increased in Design 2, but energy sales decreased. Design 3 was associated with higher emissions of carbon dioxide and sulfur dioxide, whereas Design 4 saw an increase in the utilization of purchased energy. Experiment 2 revealed a different response pattern with a 25% weight loss. Energy purchases increased in Design 1; sulfur dioxide emissions increased in Design 2; energy production increased but nitrogen oxide emissions decreased in Design 3; and energy production and sales decreased in Design 4. Analysis of experimental results with weight addition and reduction by 50% showed significant changes in the evaluation of HRES designs. In experiments with a 50% weight increase, it was seen that some designs experienced increased values on several evaluation criteria, such as energy production (C3) and renewable energy penetration (C2). However, the increase in weight also led to increased emissions of carbon dioxide (C6) and sulfur dioxide (C7) in some designs, suggesting that increased focus on certain criteria could have a negative impact on environmental aspects. Meanwhile, in experiments with a 50% weight reduction, it was seen

that there were designs that improved in some criteria, such as energy production (C3), but also decreased in other criteria, such as energy sales (C4). This suggests that weight reductions on certain criteria can change design priorities and result in trade-offs between different criteria. From the experimental results provided, Design 1 shows good relative consistency in performance relative to weight changes. Despite fluctuations in some criteria, such as energy purchase (C5), Design 1 tends to maintain stability in its relative performance.

## 4.2. DISCUSSION

This study aims to obtain the best configuration of HRES design in accordance with the environmental and energy potential of Indonesia's new capital region, namely the IKN Nusantara area in East Kalimantan province. The results of multi-criteria analysis using the CRITIC TOPSIS method and sensitivity analysis determined that design 4 (a combination of hydropower, biomass, and PV sources) became the main choice to be implemented in the case study area. We have proven this design to provide the optimal performance in various aspects, and the results of sensitivity testing demonstrate its robustness against variations in criterion weight. The incorporation of various renewable energy sources in Design 4 enables improved energy efficiency, sustainability, and a holistic reduction of environmental impact [54,55]. By harnessing the power of water, biomass, and solar energy, this design can provide a reliable and sustainable energy supply to meet electricity needs in the country's capital, IKN Nusantara.

However, it's important to remember that the final decision about a more robust design also depends on your preferences and specific goals. For example, if the top priority is reducing emissions of carbon dioxide and sulfur dioxide, then Design 1 may not be the best choice, as Design 1 is likely to increase emissions of both pollutants under some conditions. In this case, we may need to look at other designs that are more consistent in reducing those emissions. Increased or reduced weights on evaluation criteria can have significant consequences on the relative performance of HRES designs [56], [57], by affecting economic, environmental, and technical aspects [17]. Therefore, in decision-making related to the implementation of HRES, it is important to consider the impact of weight changes on the evaluation criteria thoroughly and choose the design that best suits the needs and desired goals.

The potential for the implementation of HRES that combine PV, hydropower, and biomass at research sites in East Kalimantan is very large. High solar radiation throughout the year, with a daily average of 4,661 kWh/m<sup>2</sup>/day, creates opportunities to build solar panel installations and expand the use of solar energy in the power sector. In addition, the existence of abundant river flows, with water discharge reaching 3.58 m<sup>3</sup>/s in several months, allows the construction of hydropower plants that can provide a stable and sustainable supply of elec-

tricity. Indonesia also has great potential to produce energy from palm oil waste, which is one of the largest producers in the world. The research site in Kalimantan, which generates significant amounts of palm oil waste, enables the development of biomass power plants to expand the renewable energy portfolio in the region [58]. By harnessing all these potentials together, the implementation of HRES, which combines solar, water, and biomass energy sources, can be an effective solution to meet the sustainable energy needs of East Kalimantan.

In this study, which introduces a new alternative combination method between the HOMER software and MCDM CRITIC-TOPSIS, this methodology functions as a highly effective, comprehensive, and dependable instrument for determining the viability of a hybrid power plant in a specific region. Among the limitations of this study is its reliance solely on expert opinions gathered through content analysis studies of international journals, as opposed to obtaining field data directly. This can result in less accurate research results compared to research involving direct participation. In addition, the limited availability of data on other renewable energy potential limits alternative solutions for hybrid generation to waterpower, biomass, and solar panels. And finally, in this study, the plan alludes only to a prototype government building.

## 5. CONCLUSION

This study has identified the best configuration for HRES design in Indonesia's new capital region, the IKN Nusantara area in East Kalimantan province. Through multi-criteria analysis using the CRITIC TOPSIS method and sensitivity analysis, Design 4 emerged as the optimal choice, combining hydropower, biomass, and PV sources. This design offers superior performance across various aspects and demonstrates robustness in sensitivity testing. By integrating diverse renewable energy sources, Design 4 enhances energy efficiency, sustainability, and environmental impact reduction. The potential for implementing HRES in East Kalimantan is substantial, leveraging high solar radiation, abundant river flows, and palm oil waste. However, decision-makers must carefully consider their priorities and goals when selecting the most suitable design. Despite limitations like reliance on expert opinions and limited data availability, this study's methodology effectively assesses hybrid power plant viability, providing a comprehensive and reliable approach for sustainable energy solutions in the region.

The implications of these findings highlight the urgency of optimizing renewable energy resources to meet sustainable energy needs in the East Kalimantan region and beyond. Design 4 stands out as an approach that has the potential to reduce dependence on fossil energy sources and reduce environmental impact. Therefore, we recommend that governments and stakeholders take these findings into account when planning and implementing energy projects in

the region. Further research may include an investigation of social, economic, and political factors that might influence HRES implementation, as well as technological development and innovation in the field of renewable energy. To ensure the achievement of long-term sustainability goals, we need to carry out continuous monitoring and evaluation of the performance of implemented energy systems. The suggestions also lead to more research using better or more advanced MCDM methods like AHP or TOPSIS, along with advanced analytical techniques like fuzzy logic or grey system theory. There is also in-depth research on the social, economic, and political factors that affect how well HRES is put into place.

## 6. REFERENCES

- [1] K. Almutairi, S. S. H. Dehshiri, S. J. H. Dehshiri, A. Mostafaeipour, A. Issakhov, K. Techato, "Use of a hybrid wind—solar—diesel—battery energy system to power buildings in remote areas: A case study", *Sustainability*, Vol. 13, No. 16, 2021, p. 8764.
- [2] S. M. Mahmoudi, A. Maleki, D. Rezaei Ochbelagh, "Optimization of a hybrid energy system with/without considering back-up system by a new technique based on fuzzy logic controller", *Energy Conversion and Management*, Vol. 229, 2020, p. 113723.
- [3] A. Baruah, M. Basu, D. Amuley, "Modeling of an autonomous hybrid renewable energy system for electrification of a township: A case study for Sikkim, India", *Renewable and Sustainable Energy Reviews*, Vol. 135, 2021, p. 110158.
- [4] A. K. S. Maisanam, A. Biswas, K. K. Sharma, "Integrated socio-environmental and techno-economic factors for designing and sizing of a sustainable hybrid renewable energy system", *Energy Conversion and Management*, Vol. 247, 2021, p. 114709.
- [5] A. Gonzalez, J. R. Riba, B. Esteban, A. Rius, "Environmental and cost optimal design of a biomass-Wind-PV electricity generation system", *Renewable Energy*, Vol. 126, 2018, pp. 420-430.
- [6] H. El-houari, A. Allouhi, T. Salameh, T. Kousksou, A. Jamil, B. El Amrani, "Energy, Economic, Environment (3E) analysis of WT-PV-Battery autonomous hybrid power plants in climatically varying regions", *Sustainable Energy Technologies and Assessments*, Vol. 43, 2020, p. 100961.
- [7] M. J. Mayer, A. Szilágyi, G. Gróf, "Environmental and economic multi-objective optimization of a



- household level hybrid renewable energy system by genetic algorithm”, *Applied Energy*, Vol. 269, 2020, pp. 2-16.
- [8] M. S. Nazir, A. J. Mahdi, M. Bilal, H. M. Sohail, N. Ali, H. M. N. Iqbal, “Environmental impact and pollution-related challenges of renewable wind energy paradigm - A review”, *Science of The Total Environment*, Vol. 683, 2019, pp. 436-444.
- [9] M. Moazzami, D. Fadaei, H. Shahinzadeh, S. H. Fathi, “Optimal sizing and technical analysis of rural electrification alternatives in Kerman province”, *Proceedings of the Smart Grid Conference*, Tehran, Iran, 20-21 December 2017, pp. 1-8.
- [10] M. Mehrpooya, M. Mohammadi, E. Ahmadi, “Techno-economic-environmental study of hybrid power supply system: A case study in Iran”, *Sustainable Energy Technologies and Assessments*, Vol. 25, 2016, pp. 1-10.
- [11] W. Ma, X. Xue, G. Liu, R. Zhou, “Techno-economic evaluation of a community-based hybrid renewable energy system considering site-specific nature”, *Energy Conversion and Management*, Vol. 171, 2018, pp. 1737-1748.
- [12] A. O. Samson, O. M. Babatunde, I. H. Denwigwe, “Powering a Space Environment Research Laboratory (SERL): Hybrid Renewable Energy System or Diesel System?”, *Journal of the Association of Energy Engineers*, Vol. 116, No. 2, 2019, pp. 41-64.
- [13] K. Aghapouramin, “Technical, Economical, Environmental Feasibility of Hybrid Renewable Electrification Systems for off-Grid Remote Rural Electrification Areas for East Azerbaijan Province, Iran”, *Technology and Economics of Smart Grids and Sustainable Energy*, Vol. 5, No. 1, 2020.
- [14] F. M. Kouhestani et al. “Multi-criteria PSO-based optimal design of grid-connected hybrid renewable energy systems”, *International Journal of Green Energy*, Vol. 17, No. 11, 2020, pp. 617-631.
- [15] D. Roy, R. Hassan, B. K. Das, “A hybrid renewable-based solution to electricity and freshwater problems in the off-grid Sundarbans region of India: Optimum sizing and socio-enviro-economic evaluation”, *Journal of Cleaner Production*, Vol. 372, 2022, p. 133761.
- [16] B. K. Das, R. Hassan, M. S. H. K. Tushar, F. Zaman, M. Hasan, P. Das, “Techno-economic and environmental assessment of a hybrid renewable energy system using multi-objective genetic algorithm: A case study for remote Island in Bangladesh”, *Energy Conversion and Management*, Vol. 230, 2021, p. 113823.
- [17] M. R. Elkadeem, A. Younes, S. W. Sharshir, P. E. Campana, S. Wang, “Sustainable siting and design optimization of hybrid renewable energy system: A geospatial multi-criteria analysis”, *Applied Energy*, Vol. 295, 2021, p. 117071.
- [18] Q. Jia, H. Q. Hu, J. L. Yu, “The model construction in airport renewable energy system scale planning based in Multi-Criteria Decision Making”, *Energy Reports*, Vol. 7, 2021, pp. 92-98.
- [19] D. T. Hermann, N. Donatien, T. K. Armel, T. René, “A Feasibility Study of an on-Grid PV/Wind/Battery/Diesel for Residential Buildings Under Various Climates in Cameroon”, *Energy Technology*, Vol. 9, No. 12, 2021.
- [20] D. Bishnoi, H. Chaturvedi, “Optimised Site Selection of Hybrid Renewable Installations for Flare Gas Reduction using Multi-Criteria Decision Making”, *Energy Conversion and Management X*, Vol. 10, 2022, p. 100181.
- [21] H. M. Ridha et al. “Multi-objective optimization and multi-criteria decision making aided by numerical method: Framework and a case study of Malaysia and South Africa”, *Energy Conversion and Management*, Vol. 274, 2022, p. 116468.
- [22] S. Peirow, F. Razi Astaraei, A. Saifoddin Asl, “Techno-Economic and Environmental Assessment of a Hybrid Renewable Energy System for a Hospital Using Multi-Criteria Decision-Making Method”, *Energies*, Vol. 16, No. 4, 2023, p. 1916.
- [23] R. Tarife, Y. Nakanishi, Y. Zhou, N. Estoperez, A. Tahud, “Integrated GIS and Fuzzy-AHP Framework for Suitability Analysis of Hybrid Renewable Energy Systems: A Case in Southern Philippines”, *Sustainability*, Vol. 15, No. 3, 2023, p. 2372.
- [24] P. Bakhshaei, A. Askarzadeh, R. Arababadi, “Demand response-based operation of a hybrid renewable energy system with energy storage by multi-objective optimization and multi-criteria

decision making”, *Environmental Progress & Sustainable Energy*, Vol. 42, No. 2, 2023, p. e13992

- [25] H. Yazdani, M. Baneshi, M. Yaghoubi, “Techno-economic and environmental design of hybrid energy systems using multi-objective optimization and multi-criteria decision making methods”, *Energy Conversion and Management*, Vol. 282, 2023, p. 116873.
- [26] S. Das, S. Pradhan, S. De, “Multi criteria decision making for the most suitable combination of energy resources for a decentralized hybrid energy solution with green hydrogen as the storage option”, *Energy Conversion and Management*, Vol. 285, 2023, p. 117028.
- [27] H. M. Ridha, H. Hizam, S. Mirjalili, M. L. Othman, M. E. Ya’acob, M. Ahmadipour, “Innovative hybridization of the two-archive and PROMETHEE-II triple-objective and multi-criterion decision making for optimum configuration of the hybrid renewable energy system”, *Applied Energy*, Vol. 341, 2023.
- [28] M. M. Tehrani, M. Akhtari, A. Kasaeian, M. A. V. Rad, A. Toopshekan, M. S. Motlagh, “Techno-economic investigation of a hybrid biomass renewable energy system to achieve the goals of SDG-17 in deprived areas of Iran”, *Energy Conversion and Management*, Vol. 291, 2023.
- [29] D. Roy, M. Bhowmik, A. P. Roskilly, “Techno-economic, environmental and multi criteria decision making investigations for optimisation of off-grid hybrid renewable energy system with green hydrogen production”, *Journal of Cleaner Production*, Vol. 443, 2024.
- [30] Z. Li, Y. Wang, J. Xie, Y. Cheng, L. Shi, “Hybrid multi-criteria decision-making evaluation of multiple renewable energy systems considering the hysteresis band principle”, *International Journal of Hydrogen Energy*, Vol. 49, 2024, pp. 450-462.
- [31] S. Pranomo, “Kementerian Komunikasi dan Informatika”, [https://www.kominfo.go.id/content/detail/40069/undang-undang-ibu-kota-negara-menandai-dimulainya-pembangunan-ikn/0/artikel\\_gpr](https://www.kominfo.go.id/content/detail/40069/undang-undang-ibu-kota-negara-menandai-dimulainya-pembangunan-ikn/0/artikel_gpr) (accessed: 2023)
- [32] T. Tukimun, “Konsep perencanaan infrastruktur transportasi smart, integrated sustainable & environment friendly di kawasan Ibu Kota Negara (IKN) Nusantara”, *Kurva S: Jurnal Keilmuan dan Aplikasi Teknik Sipil*, Vol. 1, No. 1, 2022.
- [33] M. Fajri, “Menggagas Ide Miniatur Hutan Dipterokarpa Ibu Kota Negara Nusantara”, *STANDAR: Better Standar Better Living*, Vol. 1, No. 2, 2022, pp. 5-15.
- [34] A. C. A. P. Tampubolon, “Laporan Status Energi Bersih Indonesia”, *IESR*, 2019, pp. 1-23.
- [35] S. Zafar, Z. Alamgir, M. H. Rehman, “An effective blockchain evaluation system based on entropy-CRITIC weight method and MCDM techniques”, *Peer-to-Peer Networking and Applications*, Vol. 14, No. 5, 2021, pp. 3110-3123.
- [36] M. Madic, M. Radovanovic, “Ranking of some most commonly used nontraditional machining processes using rov and critic methods”, *UPB Scientific Bulletin, Series D: Mechanical Engineering*, Vol. 77, No. 2, 2015, pp. 193-204.
- [37] T. Salameh, E. T. Sayed, M. A. Abdelkareem, A. G. Olabi, H. Rezk, “Optimal selection and management of hybrid renewable energy System: Neom city as a case study”, *Energy Conversion and Management*, Vol. 244, 2021, p. 114434.
- [38] J. Papathanasiou, N. Ploskas, “Topsis”, *Optimization and Its Applications*, Vol. 136, 2018, Springer, pp. 1-30.
- [39] P. Hasanah, S. A. Wiradinata, M. Azka, “Forecasting approach for solar power based on weather parameters (Case study: East Kalimantan)”, *Journal of Physics: Conference Series*, Vol. 2106, No. 1, 2021, pp. 1-10.
- [40] M. R. Kresnawan, I. A. Safitri, I. Darmawan, “Long term projection of electricity generation sector in east kalimantan province: LEAP model application”, *Proceedings of the 12th South East Asian Technical University Consortium, Yogyakarta, Indonesia, 12-13 March 2018*, pp. 1-5.
- [41] J. Windarta et al. “Planning for the utilization of hydro power in the Belayan river, East Kalimantan”, *Journal of Physics: Conference Series*, Vol. 1524, No. 1, 2020, pp. 1-7.
- [42] L. M. H. Zakir, J. Windarta, S. Saptadi, D. Rinaldo, “Analysis of planning for utilization of East Borneo hydro power with techno economy principle”, *AIP Conference Proceedings*, Vol. 2202, 2019, pp. 1-7.

- [43] F. Nashrulloh, M. Sulaiman, R. Budiarto, "Analysis of Potential and Feasibility of Hydropower Energy from Sepaku Semoi Dam in Penajam Paser Utara Regency", *IOP Conference Series: Earth and Environmental Science*, Vol. 927, No. 1, 2021, pp.1-8.
- [44] K. Esdm, "Kementerian ESDM RI - Media Center - Arsip Berita - Limbah Sawit Di Indonesia Berpotensi Hasilkan Listrik 12.654 MW", <https://www.esdm.go.id/id/media-center/news-archives/limbah-sawit-di-indonesia-berpotensi-hasilkan-listrik-12654-mw> (accessed: 2023)
- [45] K. Kelapa et al. "Biogas And Bioenergy Potential From Solid Wastes (Empty Fruit Bunch, Mesocarp Fiber, and Decanter Cake) of Palm Oil Industries in East Kalimantan", *Kemenperin*, Vol. 15, No. 2, 2021, pp. 487-497.
- [46] I. K. Hendy Wijaya, I. N. Satya Kumara, W. G. Aristina, "Analisis PLTSAatap 25 KWp on Grid Kantor DPRD Provinsi Bali", *Jurnal SPEKTRUM*, Vol. 9, No. 2, 2022, pp. 128-137.
- [47] E. T. Abit Duka, I. N. Setiawan, A. Ibi Weking, "Perencanaan Pembangkit Listrik Tenaga Surya Hybrid Pada Area Parkir Gedung Dinas Cipta Karya, Dinas Bina Marga Dan Pengairan Kabupaten Badung", *Jurnal SPEKTRUM*, Vol. 5, No. 2, 2018, pp. 67-73.
- [48] S. Pramonos, "Survey of Operation and Maintenance Costs of PV Plants in Switzerland", *EKP*, Vol. 13, 2015, pp. 3-6.
- [49] E. Oud, "The evolving context for hydropower development", *Energy Policy*, Vol. 30, 2002, pp. 1215-1223.
- [50] T. S. Kishore, E. R. Patro, V. S. K. V. Harish, A. T. Haghghi, "A comprehensive study on the recent progress and trends in development of small hydropower projects", *Energies*, Vol. 14, No. 10, 2021, p. 2882.
- [51] J. Li, P. Liu, Z. Li, "Optimal design and techno-economic analysis of a solar-wind-biomass off-grid hybrid power system for remote rural electrification: A case study of west China", *Energy*, Vol. 208, 2020, p. 118387.
- [52] C. Rangkuti, J. Teknik Mesin, "Perencanaan Pembangkit Listrik Tenaga Surya Di Atap Gedung Harry Hartanto Universitas Trisakti", *Prosiding Seminar Nasional Cendekiawan*, August 2016, pp. 1-22.
- [53] C. Rolos et al. "Analisis Penentuan Harga Jual Listrik Pada PT PLN (Persero) Unit Induk Wilayah Sulawesi Utara, Sulawesi Tengah dan Gorontalo Analysis of Electricity Price Determination At PT PLN (Persero) Main Units for North Sulawesi, Central Sulawesi and Gorontalo Regi", *Jurnal Riset Ekonomi, Manajemen, Bisnis dan Akuntansi*, Vol. 9, No. 3, 2021, pp. 1703-1710.
- [54] S. Abdelhady, "Techno-economic study and the optimal hybrid renewable energy system design for a hotel building with net zero energy and net zero carbon emissions", *Energy Conversion and Management*, Vol. 289, 2023, p. 117195.
- [55] M. Thirunavukkarasu, Y. Sawle, H. Lala, "A comprehensive review on optimization of hybrid renewable energy systems using various optimization techniques", *Renewable and Sustainable Energy Reviews*, Vol. 176, 2023, p. 113192.
- [56] R. Kumar, H. K. Channi, "A PV-Biomass off-grid hybrid renewable energy system (HRES) for rural electrification: Design, optimization and techno-economic-environmental analysis", *Journal of Cleaner Production*, Vol. 349, 2022, p. 131347.
- [57] Y. Sawle, S. C. Gupta, A. K. Bohre, "Review of hybrid renewable energy systems with comparative analysis of off-grid hybrid system", *Renewable and Sustainable Energy Reviews*, Vol. 81, 2018, pp. 2217-2235.
- [58] S. Kaniapan, S. Hassan, H. Ya, K. P. Nesan, M. Azeem, "The utilisation of palm oil and oil palm residues and the related challenges as a sustainable alternative in biofuel, bioenergy, and transportation sector: A review", *Sustainability*, Vol. 13, No. 6, 2021, p. 3110.

# 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](mailto: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

Authors of the International Journal of Electrical and Computer Engineering Systems must transfer copyright to the publisher in written form.

### 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