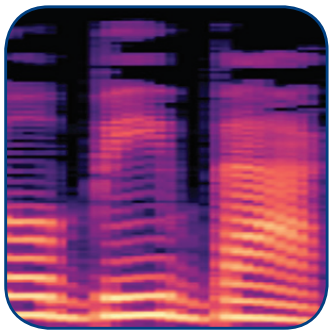
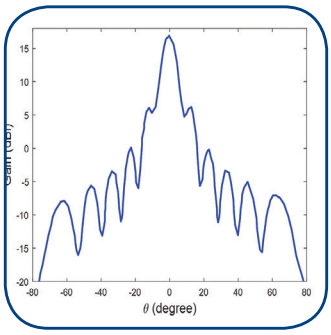
Volume 15, Number 7, 2024 ISSN 1847-6996

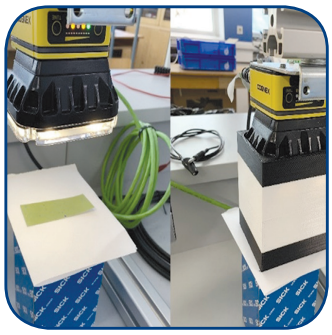
FERIT
FACULTY OF ELECTRICAL ENGINEERING, COMPUTER SCIENCE AND INFORMATION TECHNOLOGY **OSIJEK**

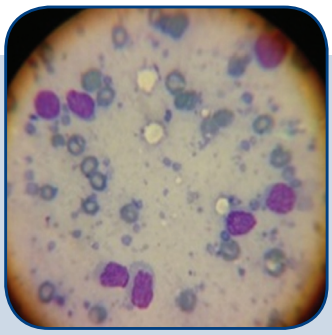
**IJECES** International Journal of Electrical and Computer Engineering Systems

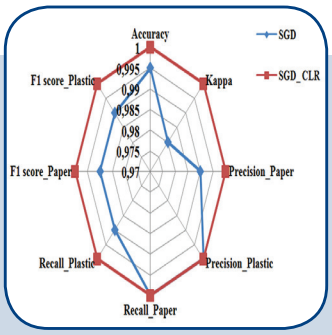
# International Journal of Electrical and Computer Engineering Systems

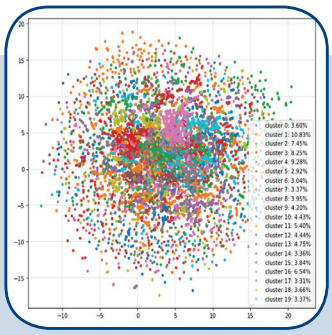








ISSN: 1847-6996

# INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

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

Osijek, Croatia | Volume 15, Number 7, 2024 | Pages 553 - 629

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

**CONTACT**

**International Journal of Electrical and Computer Engineering Systems (IJECES)**

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

**Subscription Information**

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



**Journal is referred in:**

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

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

**IJECES online**
https://ijeces.ferit.hr

# TABLE OF CONTENTS

# Quantitative Assessment of UAV Assisted Particle Spraying Distribution in Agriculture: An Image Analysis Approach Using Water-Sensitive Papers

Original Scientific Paper

**László Gogolák**
University of Szeged, Faculty of Engineering,
Department of Mechatronics and Automation,
Mars tér 7, 6720 Szeged, Hungary
gogolak@mk.u-szeged.hu

**János Simon***
University of Szeged, Faculty of Engineering,
Department of Mechatronics and Automation,
Mars tér 7, 6720 Szeged, Hungary
simon@mk.u-szeged.hu

*Corresponding author

**Árpád Pletikosity**
Subotica Tech – College of Applied Sciences,
Marka Oreškovića 16, Subotica, Serbia
arpad@vts.su.ac.rs

**Igor Fürstner**
Óbuda University, Donát Bánki Faculty of Mechanical and Safety Engineering,
Népszínház utca 8, 1034 Budapest, Hungary
furstner.igor@bgk.uni-obuda.hu

***Abstract*** *– The overall well-being and productivity of crops rely on a series of interconnected processes throughout their entire growth cycle. Among these processes, the quality of spraying plays a pivotal role in maintaining crop health and ensuring increased productivity. An Unmanned Aerial Vehicles assisted particle spraying system in agriculture involves the use of Unmanned Aerial Vehicles equipped with specialized equipment to distribute particles such as plant protection products, fertilizers, or other agricultural inputs over crops. This technology offers several advantages over traditional ground-based methods, including increased efficiency, precision, and reduced environmental impact. The effectiveness of spraying, in turn, hinges on various factors, one of which is the even distribution of spraying droplets. Consequently, there exists a need for a reliable, consistent, precise, and accurate automated method to assess the parameters governing this distribution. In this study, a methodology is introduced for evaluating the quality of plant spraying, and the results of this method's testing are presented. Data is gathered by employing water-sensitive papers positioned on the crops, which are then scanned using an industrial-grade camera. Subsequently, this data undergoes processing through image analysis algorithms using Matlab. The outcomes of the research demonstrate the robustness of the proposed methodology in obtaining the essential data required for determining spraying distribution. Compared to existing solutions, the presented approach offers increased reliability, consistency, precision, and automation, thereby addressing the need for a more reliable and accurate method of assessing spraying quality in agriculture.*

***Keywords***: *Droplet segmentation, Spray quality, Computer simulation, UAV (Unmanned Aerial Vehicle), Water-sensitive papers (WSP)*

*Received: March 4, 2024; Received in revised form: May 18, 2024; Accepted: May 18, 2024*

## 1. INTRODUCTION

In this paper, the examination and confirmation of a method for inspecting the quality of plant spraying was conducted. More specifically, the research focused on determining the occurrence of droplets on liquid-reactive papers during spraying. Based on this, methods for measuring droplets, mapping the dispersion of spray and the study of the density, size, and distribution of droplets on paper were investigated and established.

Numerous methods have been devised for plant spraying to enhance the growth of specific plant types and eradicate unwanted weeds. Also, different investigations were conducted with the aim of gaining insights into the effectiveness and characteristics of plant spraying techniques. Also, in recent years, Unmanned Aerial Vehicle (UAV) path planning and spraying droplet analysis were introduced that are essential components of modern precision agriculture practices. By optimizing coverage, minimizing environmental impact, and providing valu-

able data for decision-making, these processes have contributed to sustainable and efficient plant protection products (PPP) application in agriculture. Accurate path planning and droplet analysis has helped to ensure compliance with regulatory requirements for PPP application. By demonstrating uniform coverage and minimal drift, farmers can provide evidence of responsible PPP use and mitigate potential regulatory risks. Efficient path planning helps optimize resource usage, including PPP, fuel, and time. By following a predetermined flight path, UAVs can avoid unnecessary overlap and reduce the amount of PPP required to achieve adequate coverage. This not only saves costs but also minimizes environmental impact by reducing chemical usage. Properly measuring the quality of spraying helps maintain crop health and maximize productivity. By ensuring that the spray is evenly distributed and properly targeted, it effectively controls pests and diseases, resulting in healthier plants and improved yields. Due to the rapid development of our world and the continuous growth of the population, there is a need to increase food and grain production. Building upon this, it is essential to engage in the spraying of plants and continually monitor and enhance the process. The use of PPP has to be channeled in such a way that there is no waste, thereby avoiding environmental pollution. When using the sprays, the outside wind can act as a disturbance, so the given spray is not sprayed on the desired area. Based on this, more PPP is needed [1]. Agricultural spraying methods encompass various techniques and equipment used to apply agricultural chemicals, including PPP, herbicides, fungicides, and fertilizers, to crops for protection against pests, diseases, and weeds, as well as to provide essential nutrients. The most used spraying methods are:

- Backpack Sprayers: Backpack sprayers are manually carried/transported, and they are suitable for smaller areas or limited access [2].
- Boom Sprayers: Mounted on tractors or specialized sprayer vehicles, equipped with a tank, pump, and a boom structure with multiple spray nozzles for efficient coverage over larger areas [3].
- Airblast Sprayers: Utilize high-velocity air streams to propel spray droplets for precise targeting mainly in orchards and vineyards [4]. By harnessing powerful air streams, airblast sprayers can penetrate deep into the crop canopy, ensuring thorough coverage and effective pest and disease control. This high-rate application capability makes airblast sprayers a popular choice for growers seeking to maximize spraying efficiency and optimize crop protection.
- Aerial Spraying: Involves using aircraft, such as helicopters or fixed-wing planes, to apply sprays over large areas or challenging terrains. These days UAVs are also taking part in aerial spraying [5, 6].
- Drip Irrigation and Fertigation: Systems delivering water, fertilizers, and chemicals directly to the root zone of plants through a network of pipes with emitters [7].

Air-assisted spraying is a spraying technique that combines the use of liquid sprays and high-velocity air to improve the coverage and penetration of the spray droplets. In this method, a spray solution is atomized into fine droplets, and then a powerful air stream is used to carry and disperse the droplets effectively. Air-assisted spraying offers several advantages in agricultural applications:

- Enhanced Coverage: The high-velocity air assists in breaking down the spray solution into smaller droplets, which can result in improved coverage of the target surface, such as plant foliage. This is particularly beneficial for achieving better coverage in dense crops or complex plant structures.
- Increased Penetration: The air stream created by the sprayer helps propel the droplets deeper into the crop canopy, ensuring better penetration into the target area. This can be advantageous when targeting pests or diseases residing within the foliage.
- Reduced Drift: The use of air assistance in spraying can help minimize drift, which refers to the movement of spray particles away from the target area. By generating larger droplets and directing them with controlled airflow, air-assisted spraying reduces the risk of off-target deposition.
- Improved Efficiency: The combination of atomization and air assistance allows for more precise application of agricultural chemicals, resulting in reduced chemical usage and potential cost savings. Additionally, the improved coverage and penetration can enhance the effectiveness of the applied products [8, 9].
- Optimized Droplet Size: UAVs can be equipped with nozzles that produce droplets of an optimal size for the specific application, reducing evaporation and drift. Smaller, uniformly sized droplets are more likely to adhere to plant surfaces, improving the effectiveness of the pesticide.
- Lower Fuel Consumption: UAVs are generally more fuel-efficient compared to traditional ground-based vehicles like tractors or self-propelled sprayers. This efficiency translates to lower carbon emissions and a reduced environmental footprint.
- Resource Efficiency: By using precise application techniques, UAVs help conserve resources such as water and pesticides. This efficiency not only lowers costs but also minimizes the environmental impact associated with the production and transport of these resources.
- Environmental Monitoring: UAVs can be equipped with sensors to monitor environmental conditions in real-time, adjusting spraying parameters as needed to account for wind speed, humidity, and temperature. This adaptive approach further reduces waste and environmental impact by ensuring that spraying is conducted under optimal conditions.

Air-assisted spraying is commonly used in orchards, vineyards, and other specialty crops where precise targeting and thorough coverage are important. It requires specialized sprayer equipment that includes air blowers or fans to create the necessary air stream and atomizers to generate fine droplets [10-12]. The specific recommendations for air-assisted spraying techniques may vary depending on the target crop, spray solution, and local regulations. It's essential to follow manufacturer guidelines and consult relevant agricultural resources or experts for proper application techniques and equipment calibration [13]. On the other hand, the conventional methods are promising too, even if some other methods, like air-assisted methods, can provide several advantages. Inappropriate spraying methods, or equipment used for application can cause health issues for people. When working with PPP-s, or other chemicals, the missing calibrations or regulations can affect the quality of spraying and can result in off-target spraying that is unfortunate [14]. UAVs equipped with advanced GPS and sensors can follow precise flight paths, ensuring that PPP-s are applied only where needed. This precision reduces the amount of pesticide used by avoiding unnecessary overlap and targeting specific areas that require treatment. UAVs can be equipped with nozzles that produce droplets of an optimal size for the specific application, reducing evaporation and drift. Smaller, uniformly sized droplets are more likely to adhere to the plant surfaces, improving the effectiveness of the PPP-s.

The need for quality check and way to collect data is briefly described in the paper [15]. That work provided a system that required several portable elements that were working simultaneously. The new method presented in this paper is a novelty and the main difference between other solutions lies in the use of a more compact data collecting unit and an industry-ready high-quality camera. The proposed method used in this research is presented in Fig. 1.

Fig. 1 presents the steps during the process. The proposed methodology and the equipment that was used, as well as the results are described in the following chapters. The current methods for measuring the properties of sprayed droplets suffer from limitations such as being manual, imprecise, or expensive. Traditional approaches involve spraying colored water onto a white sheet or using a "patternator" and then analyzing the resulting patterns to estimate droplet sizes, but these methods provide only rough estimations. Alternatively, a more refined technique involves replacing the white sheet with a glass plate coated with silicone oil, known as the immersion sampling method. Nevertheless, accurately determining of the droplets' characteristics remains a significant challenge in precision agriculture. Addressing this aspect is crucial for a comprehensive understanding of droplet behavior during spraying operations.

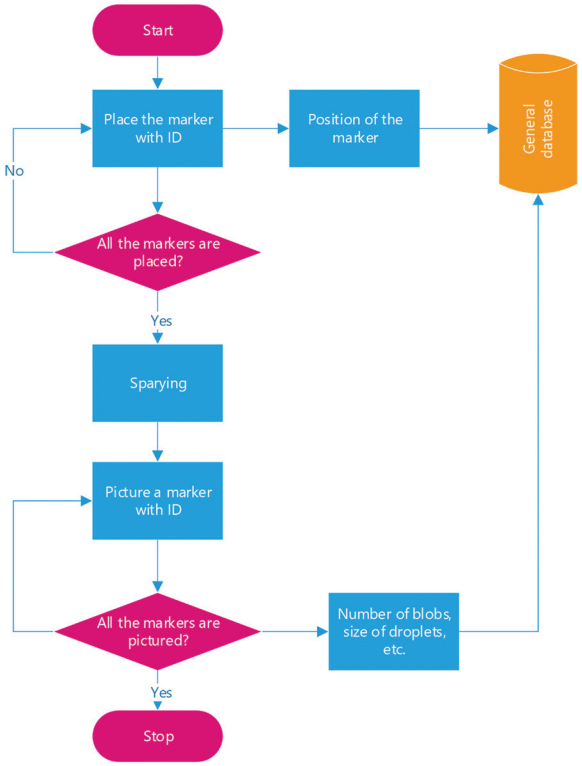



**Fig. 1.** Method of gathering on-site data

Future advancements in measurement techniques, such as the proposed method, are needed to overcome these limitations and enhance the precision and efficiency of agricultural spraying practices [16]. In this paper, significant contributions are made to the field of agricultural spraying by addressing the critical need for reliable and precise methods of assessing spray quality. Specifically, a methodology for evaluating the quality of plant spraying is examined and confirmed, with particular emphasis placed on detecting droplets on liquid-reactive papers during spraying. The contributions of this paper can be summarized as follows:

- A comprehensive methodology is introduced, encompassing various aspects of assessing spray quality, including methods for measuring droplets, mapping spray dispersion, understanding the importance of quality assessment, and studying droplet density, size, and distribution on paper.
- Integration of Advanced Technologies such as UAV path planning and spraying droplet analysis are leveraged to enhance precision, efficiency, and environmental sustainability in agricultural spraying practices.
- A novel approach for data collection is introduced, utilizing water-sensitive papers positioned on crops and scanned using an industrial-grade camera. This approach enables automated data collection and processing, enhancing accuracy and reliability.
- By providing insights into the effectiveness and characteristics of plant spraying techniques, this study contributes to the advancement of preci-

sion agriculture practices, ultimately leading to improved crop health and productivity.

By addressing the limitations of existing techniques and leveraging cutting-edge technologies, the methodology provided in this paper serves as a valuable tool for farmers, researchers, and practitioners seeking to optimize spraying operations and minimize environmental impact.

## 2. RELATED WORK

In [1], an analysis of the spray drift mechanism and the primary factors influencing aerial spraying was conducted, and previous research on Drift Reducing Technologies (DRT) in aerial spraying was reviewed and summarized. It was found that DRT-s in aerial spraying, such as aerial electrostatic spray technology, aerial spray adjuvant, aerial air-assisted spray technology, drift reducing nozzles, and aerial variable-rate spray technology, can effectively reduce environmental pollution caused by PPP drift by decreasing the spraying amount and improving the control effect of PPP. The exact methodology and equipment are required in every spraying application, whether it is from the air or from the ground. From [2], it can be concluded that calibration is vital, and there are rarely overall results that can be used with complete confidence for every sprayer.

When time and energy play a major role in the process, optimal methods should be applied. Some methods are faster than others and are easier to replicate, just like the DEIAFA method from [3]. This work shows that every workflow should be overviewed before deciding to use one. The study presented in [4] proposes an alternative indirect methodology for the comparative measurements of drift reduction potential generated by airblast sprayers, aimed at addressing the practical inconveniences and drawbacks. In [5], precision aerial application for site-specific crop management is discussed, with an examination of several current trends and propositions for future development. Research paper [6] highlights that domestic research on the aerial spraying application of plant protection UAVs primarily focuses on the impact of aerial spraying operation parameters on the distribution of droplet deposition. However, it tends to overlook the evaluation and testing of the effective spraying width of aerial spraying by plant protection UAVs. Even though the exact method, equipment, and setting are crucial, the time of different interventions is also important. From irrigation systems we can learn that the dose, and the way of application can bring savings in financials, and in water usage too [7]. Reducing, or better called optimizing deposit volumes can be achieved by using finer nozzles, changing travel speeds, or using air assistance. Air assistance can increase the possibility of aiming vertical targets [8]. In [9], a study was conducted to ascertain the effects of different parameters on spraying. It was observed that the operating speed of the sprayer had a significant impact on spray droplets deposition and distribution, being notably greater at lower operating speeds. In air-assisted sprayers, an increase in blower speed significantly enhanced drift force and expanded the tree canopy area covered in spraying. In [10], an integrated computational fluid dynamics (CFD) model was developed. This model predicts the displacement of PPP spray droplets discharged from an air-assisted sprayer, their depositions onto tree canopies, and off-target deposition and airborne drift in an apple orchard. In order to maximize target deposition, the outcomes made it evident how useful CFD models are for examining various sprayer configurations. As a result of the research presented in [11], CFD may be utilized as a tool for numerical prototyping, which cuts down on the number of tests. Also, the model's coupling with the droplet deposition and canopy airflow models already in use help the design and operation of sprayers with a lower risk of environmental contamination.

Environmental diversification also affects the quality of spraying. Therefore, in the paper [12] the authors anticipated the need for a smart spray analytical system that helped the calibration of the air-assisted sprayer continuously. Several materials can be used in the droplet tracking workflow, like water-soluble food dye. Thanks to the approach from [13], it's possible to boost the amount of total deposition on the canopy from 48.4% to 65.6% of the applied dose, without significantly raising the amount of spray that is lost to the ground. Researchers in [14] were engaged in experimental studies on the design, development, and testing of precision spraying technologies for crops and orchards. Numerous new alternative methods were published. For example, research in [17] shows that UAVs can serve as a potential machine that can work as an alternative for spraying applications. With water-sensitive paper, the nature of airflow is discussed. Coverage was measured in several levels of the canopy, in the aim to have a better understanding of the behavior of the deposit. Even though the unit can be used as an air-assisted sprayer, the volume of mixed chemicals is reduced. The droplet deposit is lower than by using conventional methods, but it doesn't necessarily mean that the volume of active ingredient is also reduced. According to the results of research presented in [18] the way of measurement can be improved by other methods. AI-based algorithms and machine learning can provide precise, and fast decision making in real time droplet management. The review presented in [19] reveals that drones can redefine agriculture in a way that drives this industry on a new path. Implementation of modern technologies is able, for example, to perform electromechanical flow control and cutting-edge nozzles, and transformative AI.

## 3. MATERIALS AND METHODS

### 3.1. PATH PLANNING FOR OPTIMAL AREA COVERAGE

Ardupilot Mission Planner is a comprehensive software application used for configuring and controlling

unmanned vehicles that run on the Ardupilot open-source autopilot firmware. Ardupilot is one of the most advanced, full-featured, and reliable open-source autopilot software available, supporting a wide range of UAVs [20, 21]. It is necessary to provide the key GPS coordinates of the area for the software to calculate the overflight path of the given surface as shown in Fig. 2.

**Fig. 2.** Mission planner environment

A mission planner is capable of planning complex missions with multiple waypoints and conditional commands and offers real-time data streaming from the vehicle, enabling live monitoring of various metrics such as altitude, speed, battery status, and GPS data [22].

**Fig. 3.** Planning a mission with waypoints and events

Mission Planner is particularly popular among hobbyists, researchers, and professionals in the field of unmanned systems, as it provides a user-friendly interface and a rich set of features for effectively managing and controlling unmanned missions. It runs on Windows and is integrated to many Ardupilot-based UAV operations. Defining the drone's flight path is solved by the Ardupilot software package, which calculates the coordinates that the aircraft will follow in order to optimally cover the defined area as shown in Fig. 3. After planning the optimal path of the drone over the defined area, the flight plan is executed, and the area is dusted.

### 3.2. PROPOSED METHODOLOGY FOR THE SPRAYING QUALITY EVALUATION

To be able to determine the quality of dusting, water-sensitive paper indicators are placed in appropriate places in the investigated area. After the spraying the paper slips are collected and digitized.

Algorithm 1 outlines the steps for measuring droplet deposition from an image containing sprayed droplets. It involves preprocessing the image, identifying droplet spots, extracting relevant features, and visualizing the results for analysis and interpretation. The search for pixels consisted of the following steps:

1. Color Space Transformation;
2. Thresholding;
3. Morphological Operations;
4. Region Analysis;
5. Statistical Analysis;
6. Visualization.

To implement these steps, a MATLAB script and applied MATLAB's built-in image processing functions were used.

The method presented in this article analyzes the dusting quality, i.e. the size of the droplets. The pseudocode serves as a blueprint for implementing the algorithm in software. It outlines the steps and logic required to analyze the quality of dusting, providing guidance for programmers tasked with writing the actual code.

**Algorithm 1** The droplet deposition measurement

```
% Looking for blue spots from image
img = imread('Original source image.jpg');
% Color sharpening (Identification of blue spots)
blueMask =(img(:,:,1) >0 & img(:,:,1) < 80) & ...
    (img(:,:,2) > 0 & img(:,:,2) < 80) & ...
    (img(:,:,3) > 0 & img(:,:,3) < 200);
% Morphological Operations (Noise Reduction)
% Remove spots smaller than 50 pixels
blueMaskCleaned = bwareaopen(blueMask, 5);
% Labeling (Spot Labeling)
labeledBlueObjects = bwlabel(blueMask);
% Feature Extraction (Definition of Properties)
stats = regionprops(labeledBlueObjects, 'Area',
'Centroid', 'Perimeter');
% Histogram (Histogram of spots by area)
areas = [stats.Area];
histogram(areas);
% Display spots on the image
imshow(img);
hold on;
for i = 1:numel(stats)
    plot(stats(i).Centroid(1), stats(i).Centroid(2), 'r+');
% Display of spot centers
plot(stats(i).Centroid(1), stats(i).Centroid(2), 'r*');
end
hold off;
% Show results
figure;
imshow(blueMask);
```

The algorithm facilitates the digitization of data collected from the paper slips. This process is essential for transforming raw observational data into a format that can be analyzed and interpreted by computational methods, enabling quantitative assessment of dusting quality. By analyzing the droplet distribution on the water-sensitive or substance-sensitive papers collected from certain areas, the algorithm helps evaluate the effectiveness of the spraying process. This assessment is vital for ensuring optimal coverage and efficacy in agricultural spraying operations.

### 3.3. MEASUREMENT SYSTEM

The device used to assist us in completing the task was a Cognex In-Sight 7000 industrial camera. It is a high-performance and reliable tool that greatly facilitates and accelerates the counting process of droplets. The specific camera used has a resolution of 640 x 480 / 800 x 600 pixels. It provides high-speed and fast processing capabilities, enabling quick image processing and data analysis. It features built-in image processing capabilities, such as contour detection, shape recognition, and other functionalities. It allows communication via Ethernet, RS-232, and USB interfaces. The camera is housed in a robust aluminum casing with IP67 protection. Fig. 4 shows the physical dimensions of the Cognex In-Sight industrial camera.

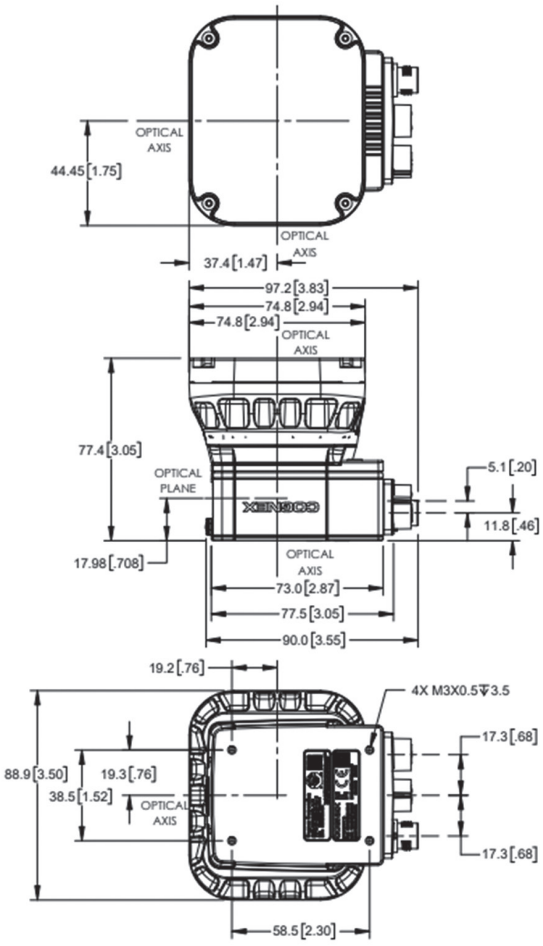
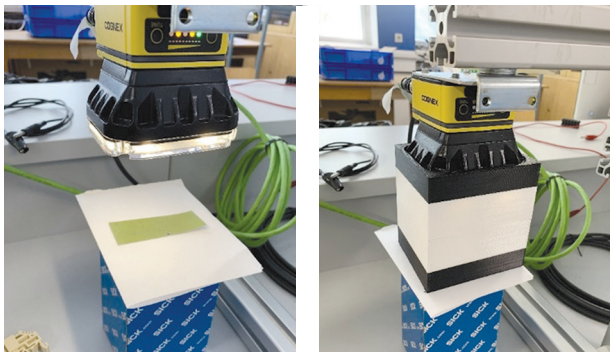



**Fig. 4.** The physical dimensions of the Cognex In-Sight industrial camera

The camera communicates with multiple software, and a program called In-Sight Explorer. For this task, the settings and filters to easily locate and count the droplets on the water-sensitive papers were adjusted. The camera has its own built-in LED lights, which illuminate the paper placed in the focal point due to their high brightness. External lighting for the task was not used, but there were external ambient lights present [23]. These external lights disrupted the measurement process as the camera could not accurately detect the positions of the droplets. The multiple light sources caused the image to become blurry. The camera uses an M12 lens type, which is equipped with a chain of circularly arranged LED lights. There are a total of 8 built-in LEDs. These lights can be toggled on and off to aid in focusing as depicted in Fig 5.



(a) (b)

**Fig. 5.** (**a**) Digitalization without constant light chamber and (**b**) Digitalization with constant light chamber

Since the LED light is directional, it caused significant reflection due to the paper's surface properties, resulting in focusing issues. To address this, a solution was implemented where a diffusing film directly after the LEDs was placed. This achieved light diffusion, preventing the reflection from the paper surface. Table 1 shows the droplet classification system ASAE standard S-572.

**Table 1.** Droplet classification system ASAE standard S-572

| Nozzle category | Symbol | Colour code | VMD |
|---|---|---|---|
| Very fine | VF | Red | <150 |
| Fine | F | Orange | 150-250 |
| Medium | M | Yellow | 250-350 |
| Coarse | C | Blue | 350-450 |
| Very coarse | VC | Green | 450-550 |
| Extremely coarse | XC | White | >550 |

The methodology takes into account factors such as spray rate, UAV speed, climatic conditions, and wind speed, ensuring a holistic approach to spray quality evaluation as shown in Table 2.

By providing insights into the effectiveness and characteristics of plant spraying techniques, this study contributes to the advancement of precision agriculture

practices, ultimately leading to improved crop health and productivity. By considering factors such as spray rate, UAV speed, climatic conditions, and wind speed, the methodology offers a comprehensive approach to spray quality evaluation, enabling farmers to make informed decisions and optimize spraying operations.

**Table 2.** Agricultural spraying operation factors

| Parameter | Values |
|---|---|
| Spray Rate | 12.5 liters/hectare |
| UAV Speed | 3 meters/second |
| Climatic Conditions | Temperature: 25°C <br> Humidity: 60% Atmospheric Stability: Stable |
| Wind Speed | 5 meters/second |
| Droplet Size | Reference ASABE S572 |

## 4. RESULTS

It has been shown in studies that among various factors, droplet size is the most significant in causing spray drift. The smaller the droplet, the longer it remains airborne and the more susceptible it becomes to drifting with the wind.

**Table 3.** Effects of droplet sizes on drift potential

| Droplet diameter /µm | Suspension time required for droplets to drop by 3 m/s |
|---|---|
| 5 | 3960 |
| 20 | 252 |
| 100 | 10 |
| 240 | 6 |
| 400 | 2 |
| 1000 | 1 |

Table 3 contains the effects of droplet sizes on drift potential. Meteorological parameters are essential factors that cannot be overlooked in the examination of droplet deposition and drift. During the process of droplet deposition from the nozzle to the ground, droplets are influenced by temperature and relative humidity, and evaporation results in the reduction of droplet size, making them more prone to drifting extensively in the natural wind [23, 24]. Therefore, the primary meteorological parameters that affect droplet deposition and drift include natural wind, temperature, humidity, and atmospheric stability.

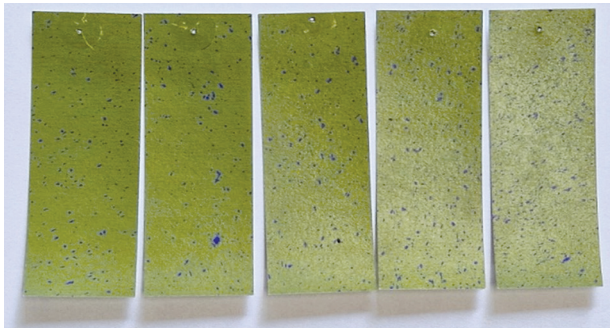


**Fig. 6.** Water-sensitive papers after the UAV-assisted particle spraying process

For the analysis, water-sensitive papers were placed under the leaves of the plants to ensure the accuracy of spraying i.e. the droplet deposition measurement as depicted in fig. 6. When the spray comes into contact with these papers, they undergo color changes. The droplets, varying in size depending on the amount of liquid, help in determining the proper air-to-water ratio during the spraying process. Different results were provided by the collected papers due to modifications in the spraying process. The task involved gathering these papers and studying the droplets formed on them. To expedite the counting process, industrial cameras are being used.

### IMAGE PROCESSING OF THE SAMPLES AND DROPLET SEGMENTATION

Image processing on the computer could begin after scanning the samples. Each sample was individually saved and processed separately. In this work, the processing of only two samples is presented, but the same algorithm has been applied for all samples as well.

### COLOR SPACE TRANSFORMATION

The MATLAB code represents a classic approach in image processing to segment and analyze specific regions or features of interest, commonly referred to as "blobs" or "regions". In this instance, the regions of interest are potentially blue-colored in the image. The image in MATLAB will be in a matrix form where each pixel is represented by its RGB (Red, Green, Blue) values. Digital images are typically represented in the RGB color space, where each pixel's color information is encapsulated by three intensity values corresponding to the red, green, and blue channels. However, the RGB representation is not always ideal for segmentation tasks, primarily due to the tight coupling between color and intensity. Consequently, many segmentation tasks leverage the HSV (Hue, Saturation, Value) color space. In HSV, colors are described based on their hue (type of color), saturation (vibrancy), and value (brightness). This transformation provides a more intuitive paradigm for many image analysis tasks, as hue decouples color information from luminance, often simplifying the segmentation process. Fig. 7 presents the original RGB picture, the HSV color space picture and every component of the HSV picture.

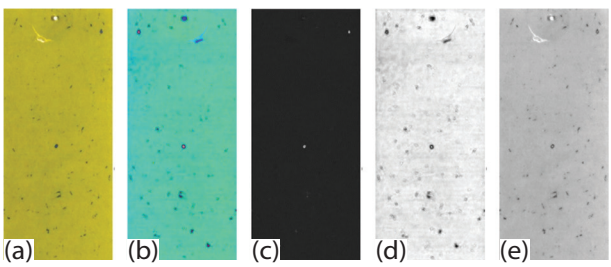



**Fig. 7.** Water-sensitive paper in the image processing: (**a**) Original RGB picture, (**b**) HSV color space picture represented in MATLAB, (**c**) Hue component, (**d**) Saturation component, (**e**) Value component

From the images, it is clear that useful information can only be obtained from the Saturation and Value components. This is where the droplets can be best distinguished. In the remainder of the paper, the Saturation component was used.

## THRESHOLDING

The histogram of the Saturation component shows which values the pixels take and what distribution they follow (see Fig. 8). This information will be crucial for the subsequent segmentation of the image.

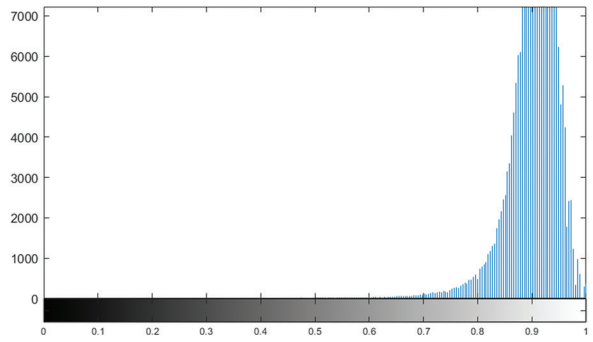



**Fig. 8.** Histogram of the Saturation component

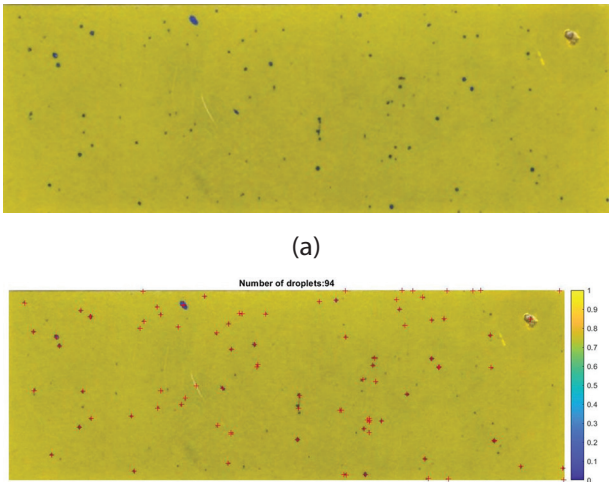
An image thresholding technique was used to segment the blue spots on the water-sensitive papers. Specifically, a color-based thresholding method was applied to isolate the blue regions from the background. This method ensures accurate identification and measurement of spray droplets. A widely utilized technique in image segmentation is thresholding. It involves categorizing pixel values based on specific thresholds and creating a binary mask where pixels within the desired range are highlighted. In this instance, the Saturation channel of the HSV image is subjected to thresholding to identify potential dark blue droplets. The rationale behind targeting the Value channel is based on the understanding that blue regions might possess a specific brightness range, which can be isolated via thresholding. By setting lower and upper bounds on this channel, a mask that potentially contains blue spots is obtained. The best result can be obtained with the usage of lower bounds of 0.3 and the upper bound is set to 0.6.

## MORPHOLOGICAL OPERATIONS AND REGION ANALYSIS

Following thresholding, images often contain noise or small unwanted artifacts. Morphological operations offer tools for refining segmented regions, enhancing the accuracy of the segmentation. Two such operations are:

- Area opening: This operation removes small connected components or blobs based on a specified pixel threshold, ensuring the elimination of noise and tiny undesired regions;
- Hole filling: Sometimes, segmented regions may have small gaps or holes. Hole filling ensures that these regions are filled, providing a more coherent segmented region.

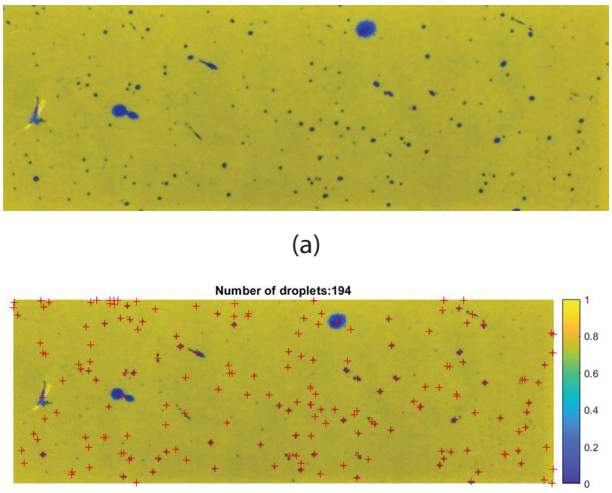
Once the regions of interest are segmented, one may desire to extract properties from these regions for further analysis. The region-props function in MATLAB facilitates the extraction of numerous properties from labeled regions. In this context, the area, centroid, and filled area of the detected spots are extracted. These parameters offer insights into the size, distribution, and morphology of the detected blue spots. Using the aforementioned functions, some results can be seen in Figs. 9 and 10.




(a)

(b)

**Fig. 9.** (**a**) Original sample 1 and (**b**) processed sample 1




(a)

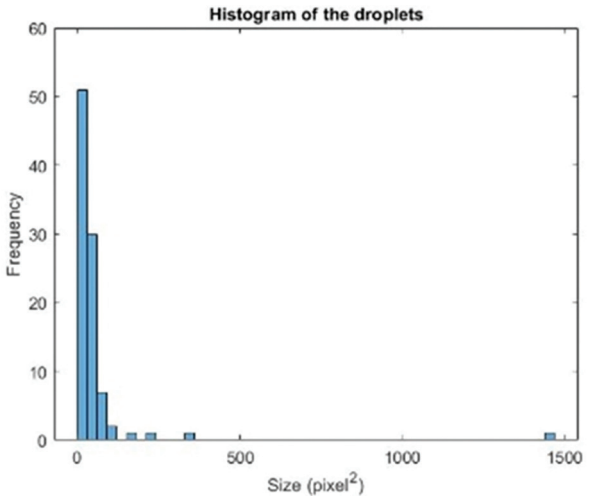
(b)

**Fig. 10.** (**a**) Original sample 2 and (**b**) processed sample 2

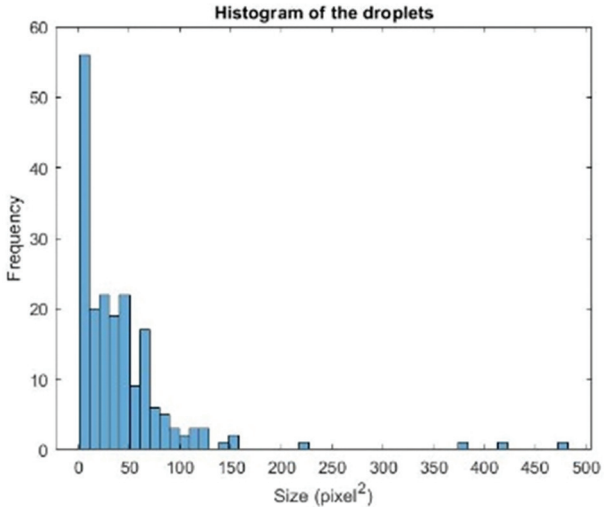
## STATISTICAL ANALYSIS AND VISUALIZATION

Subsequent to the extraction of properties, statistical analyses, such as plotting histograms, can be employed to understand the data's distribution. Histograms represent data distribution by organizing a series of data points into user-defined ranges.

In this context, the area of the detected blue spots is visualized using a histogram, providing insights into the size distribution of these spots. Such visual statistical tools are pivotal in understanding the underlying patterns and characteristics of segmented features. Visual validation is an integral aspect of image processing, allowing researchers to confirm the efficacy of their methodologies as can be seen in Fig. 11.




(a)




(b)

**Fig. 11.** (**a**) Histogram of the sample 1 and (**b**) sample 2

In the code, the original image is overlaid with markers indicating the centroids of the detected blue spots, providing a visual affirmation of the segmented regions. The accuracy of image processing can be improved by enhancing the scanning quality and by fine-tuning the threshold parameters. The quality of the water-resistant paper also greatly affects the quality of image processing. If the system has not been calibrated properly for the lighting conditions or the specific camera being used, it can result in inaccuracies in the HSV color space conversion, subsequently affecting the segmentation process. Factors like uneven lighting, shadows, or reflections can affect the appearance of objects in the image, potentially leading to inaccuracies in detection.

## 5. CONCLUSION

In conclusion, the presented methodology exemplifies a structured approach to image segmentation and analysis. By leveraging the unique properties of the HSV color space, combined with thresholding, morphological refinements, and statistical analyses, one can accurately segment and interpret specific features in an image. Such techniques, rooted in the fundamental principles of image processing, offer a scientific approach to extract patterns and features from complex image data. As the ecological environment continues to deteriorate and people's expectations for modern life quality increase, increasing attention is being paid to the issues of PPP pollution and residues. Consequently, the trend in PPP usage is inevitably moving towards achieving a higher utilization rate of PPP and minimizing environmental pollution. An analysis of the spray drift mechanism and a review of prior studies have revealed that factors such as droplet size, meteorological parameters, nozzles, operating parameters, and the physicochemical properties of liquid medicine all influence the droplet deposition and drift in aerial spraying. Therefore, employing an effective combination of low- or ultra-low-volume spraying operation modes in aerial spraying is an important strategy to attain this goal. By optimizing spray quality and distribution, our methodology helps reduce PPP use, lower environmental pollution, and promote sustainable farming practices.

## 6. REFERENCES:

[1] S. Chen, Y. Lan, Z. Zhou, X. Deng, J. Wang, "Research advances of the drift reducing technologies in application of agricultural aviation spraying", International Journal of Agricultural and Biological Engineering, Vol. 14, No. 5, 2021, pp. 1-10.

[2] C. G. Landgren, "Calibrating and using backpack sprayers", A Pacific Northwest Extension Publication, 1996.

[3] P. Balsari, P. Marucco, M. Tamagnone, "A test bench for the classification of boom sprayers according to drift risk", Crop Protection, Vol. 26, No. 10, 2007, pp. 1482-1489.

[4] M. Grella, P. Marucco, P. Balsari, "Toward a new method to classify the airblast sprayers according to their potential drift reduction: comparison of direct and new indirect measurement methods", Pest Management Science, Vol. 75, No. 8, 2019, pp. 2219-2235.

[5] Y. Lan, S. J. Thomson, Y. Huang, W. C. Hoffmann, H. Zhang, "Current status and future directions of precision aerial application for site-specific crop management in the USA", Computers and Electronics in Agriculture, Vol. 74, No. 1, 2010, pp. 34-38.

[6] S. Chen, Y. Lan, J. Li, X. Xu, Z. Wang, B. Peng, "Evaluation and test of effective spraying width of aerial spraying on plant protection UAV", Transactions of the Chinese Society of Agricultural Engineering, Vol. 33, No. 7, 2017, pp. 82-90.

[7] M. M. Ibrahim, A. A. El-Baroudy, A. M. Taha, "Irrigation and fertigation scheduling under drip irrigation for maize crop in sandy soil", International Agrophysics, Vol. 30, No. 1, 2016, pp. 47-55.

[8] E. Nordbo, "Effects of nozzle size, travel speed and air assistance on deposition on artificial vertical and horizontal targets in laboratory experiments", Crop Protection, Vol. 11, No. 3, 1992, pp. 272-278.

[9] C. V. Jadav, K. K. Jain, B. C. Khodifad, "Spray of Chemicals as Affected by Different Parameters of Air Assisted Sprayer: A Review", Current Agriculture Research Journal, Vol. 7, No. 3, 2019, pp. 289-295.

[10] S. W. Hong, L. Zhao, H. Zhu, "CFD simulation of pesticide spray from air-assisted sprayers in an apple orchard: Tree deposition and off-target losses", Atmospheric Environment, Vol. 175 No. 1, 2018, pp. 109-119.

[11] M. A. Delele, P. Jaeken, C. Debaer, K. Baetens, A. M. Endalew, H. Ramon, P. Verboven, "CFD prototyping of an air-assisted orchard sprayer aimed at drift reduction", Computers and Electronics in Agriculture, Vol. 55, No. 1, 2007, pp. 16-27.

[12] H. Y. Bahlol, A. K. Chandel, G. A. Hoheisel, L. R. Khot, "The smart spray analytical system: Developing understanding of output air-assist and spray patterns from orchard sprayers", Crop Protection, Vol. 127, 2020, pp. 1-10.

[13] G. Pergher, N. Zucchiatti, R. Gubiani, "Influence of spray application parameters on deposition in an asparagus crop", Journal of Agricultural Engineering Research, Vol. 73 No. 1, 1999, pp. 19-28.

[14] F. Ahmad, M. Sultan, "Advancement in spraying technology in agriculture", Technology in Agriculture, IntechOpen, 2021, pp. 33-51.

[15] H. Zhu, M. Salyani, R. D. Fox, "A portable scanning system for evaluation of spray deposit distribution", Computers and Electronics in Agriculture, Vol. 76, No. 1, 2011, pp. 38-43.

[16] P. Acharya, T. Burgers, K. D. Nguyen, "Ai-enabled droplet detection and tracking for agricultural spraying systems", Computers and Electronics in Agriculture, Vol. 202, No. 1, 2022, p. 107325.

[17] Y. Rashid, M. D. M. Nasir, A. A. M. Noh, S. A. Bakar, W. N. Wan, "Effectiveness of Drone Spraying to Control Bagworm Outbreak", The Planter, Vol. 99, 2023, pp. 723-735.

[18] S. Guo, C. Chen, G. Du, F. Yu, W. Yao, L. Yubin, "Evaluating the use of unmanned aerial vehicles for spray applications in mountain Nanguo pear orchards", Pest Management Science, 2024. (in press)

[19] A. Taseer, X. Han, "Advancements in variable rate spraying for precise spray requirements in precision agriculture using Unmanned aerial spraying Systems: A review", Computers and Electronics in Agriculture, Vol. 219, 2024, p. 108841.

[20] D. Csik, Á. Odry, P. Sarcevic, "Fingerprinting-Based Indoor Positioning Using Data Fusion of Different Radiocommunication - Based Technologies", Machines, Vol. 11, No. 2, 2023, p. 302.

[21] P. A. Hobson, P. C. H. Miller, P. J. Walklate, C. R. Tuck, N. M. Western, "Spray drift from hydraulic spray nozzles: the use of a computer simulation model to examine factors influencing drift", Journal of Agricultural Engineering Research, Vol. 54, No. 4, 1993, pp. 293-305.

[22] C. Gong, F. Chen, B. Cui, A. Wang, Z. Zhang, Z. Zhou, Y. Liu, "Droplet spatial distribution of oil-based emulsion spray", Frontiers in Plant Science, Vol. 14, 2023, p. 1183387.

[23] P. Yang, K. Tang, J. A. Lozano, X. Cao, "Path planning for single unmanned aerial vehicle by separately evolving waypoints", IEEE Transactions on Robotics, Vol. 31, No. 5, 2015, pp. 1130-1146.

[24] Z. Qadir, M. H. Zafar, S. K. R. Moosavi, K. N. Le, M. P. Mahmud, "Autonomous UAV path-planning optimization using metaheuristic approach for predisaster assessment", IEEE Internet of Things Journal, Vol. 9, No. 14, 2021, pp. 12505-12514.

# Acute Leukemia Subtype Recognition in Blood Smear Images with Machine Learning

Original Scientific Paper

**Ashwini P. Patil***
Department of Computer Science, CHRIST (Deemed to be University),
Bengaluru, India
patil.ashwini.p@gmail.com

**Manjunatha Hiremath**
Department of Computer Science, CHRIST (Deemed to be University),
Bengaluru, India
manju.gmtl@gmail.com

*Corresponding author

***Abstract*** *– Acute leukemia is a swiftly progressing blood cancer affecting white blood cells which poses a significant threat to the immune system and often leads to fatal outcomes if not detected and treated promptly. The current manual diagnostic method, being time-consuming and prone to errors, necessitates an urgent shift toward a comprehensive automated system. This paper presents an innovative approach to automatically identify acute leukemia cells and their subtypes by analyzing microscopic blood smear images. The proposed methodology involves the segmentation of clustered lymphocytes, isolation of nuclei, and extraction of diverse features from each nucleus. A random forest classifier is then trained to categorize nuclei into healthy or cancerous, with further precision in classifying cancerous nuclei into specific subtypes. The method achieves an impressive 97% accuracy across all evaluations, holding profound implications for pathologists and medical practitioners in their decision-making processes*

***Keywords***: *Acute Leukemia; Segmentation; Image processing; cell analysis; Leukemia Classification*

*Received: March 6, 2024; Received in revised form: May 26, 2024; Accepted: May 28, 2024*

## 1. INTRODUCTION

Leukemia is anticipated to witness a 35.1% surge in cases across Asia between 2020 and 2040, as projected by the World Health Organization's International Agency for Research on Cancer [1]. This cancer targets white blood cells, integral to our immune system. It triggers the excessive production of abnormal white blood cells within the blood and bone marrow, potentially spreading to other organs and compromising immunity. The disease is broadly classified into four main types: Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Chronic Lymphocytic Leukemia (CLL), and Chronic Myeloid Leukemia (CML). The classification of leukemia depends on its progression speed and affected cell lineage. Acute leukemia, with its swift progression, features immature blasts multiplying rapidly in the bone marrow, causing sudden symptoms. In contrast, chronic leukemia advances more slowly, with mature abnormal cells resembling normal ones, resulting in a gradual onset of symptoms. Lymphoid or myeloid categorization depends on the origin of affected white blood cells.

According to FAB classification within each primary category multiple subtypes exist. The subtypes of Acute Lymphoblastic Leukemia (ALL) are categorized as L1, L2, and L3. L1 is characterized by small and uniform lymphoblasts while L2 involves larger and more diverse lymphoblasts. L3 also identified as Burkitt's leukemia is distinguished by large lymphoblasts with intense staining. The classification of Acute Myeloid Leukemia (AML) spans from M0 to M7. M0 denotes undifferentiated AML with undefined characteristics, while M1 through M5 signify increasingly mature myeloid cells. M6 signifies acute erythroleukemia, involving both erythroid and myeloid precursors, and M7 is linked to acute megakaryoblastic leukemia, characterized by abnormal megakaryocytes.

Advancements in genetic research have significantly enhanced our understanding of leukemia, leading to ongoing changes in classification systems, such as the World Health Organization's (WHO) classification. The complexity of leukemia arises from a combination of genetic, molecular, and cytogenetic factors, making each case unique [2, 3]. Factors like leukemia type, age,

overall health, genetics, and exposure to risk elements contribute to varying risks associated with the disease.

Current manual diagnosis methods, involving the analysis of blood smears and bone marrow smears under a microscope, have limitations due to subjective interpretations and lack of quantitative precision, particularly in distinguishing visually similar leukemia subtypes. In response to these challenges, machine learning has emerged as a promising solution for leukemia diagnosis [4]. Utilizing sophisticated algorithms, machine learning provides a quantitative assessment of cell morphology, detects intricate patterns, and objectively evaluates anomalies, reducing human subjectivity. In summary, machine learning revolutionizes our approach to leukemia, promising improved patient care and enriching research insights by elevating accuracy and adapting to evolving disease characteristics.

The focal point of this research paper is the utilization of machine learning methodologies for the detection and classification of the main types and subtypes of acute leukemia. This model stands poised to assist pathologists and medical practitioners in decision-making processes. The paper's structure unfolds as follows: Section 1 presents the study's scope; Section 2 delves into the realm of literature review; Section 3 furnishes dataset particulars; Section 4 elaborates on the proposed methodology; Section 5 discusses experimental outcomes; and Section 6 culminates in a conclusive wrap-up.

## 2. BACKGROUND AND RELATED WORKS

The critical nature of leukemia necessitates early diagnosis for effective treatment. To achieve this, various automated systems have been developed, leveraging affordable microscopic image analysis, to facilitate timely and accurate diagnoses. Putzu et al. [5] introduced a fully automated method for Acute Lymphoblastic Leukemia (ALL) classification achieving 93% accuracy using SVM-based machine learning on the ALL-IDB database. Mohapatra et al. [6] achieved a notable 99% classification accuracy with an ensemble-based technique on their private dataset. Rawat et al. [7] distinguished acute lymphoblast and acute myeloblast subtypes, attaining an exceptional 99.5% overall accuracy with a publicly available dataset.

Inbarani et al. [8] applied an innovative hybrid histogram-based soft covering rough k-means clustering (HSCRKM) algorithm for segmenting leukemia nucleus images. Devi et al.[9] introduced an innovative approach incorporating Gaussian Blurring, Hue Saturation Value (HSV) and morphological operations, achieving 96.30% accuracy using private dataset whereas 95.41% using ALL-IDB1 public dataset. M. Ashok et al. [10] proposed an automated machine learning approach (Chabot) achieving accurate classification of infected and healthy cells in blood smear images, aiming to detect ALL by utilizing CMYK color space and K-means clustering.

While interpretable and resource-efficient, traditional machine learning models may fall short in capturing intricate patterns within medical images compared to the effectiveness of deep learning, particularly transfer learning, known for its high accuracy in leukemia detection and classification [11, 12]. Ansari et al. [13] suggested CNN model which uses Tversky loss function, designed for the classification of acute leukemia images, attained an impressive accuracy rate of 99%. Boldú et al. [14] constructed with two sequentially connected convolutional networks (ALNET), was trained using a dataset comprising over 16,000 blood cell images collected from clinical practice. Abass et al. [15] proposed CAD3 system utilizing YOLO v2 and CNN for leukocyte detection and classification.

One of the challenges in deep learning is the availability of large datasets, and in many cases, these datasets may not be properly labeled [16]. Manescu et al. [17] used patient diagnostic labels to train weakly-supervised models for the detection of various acute leukemia types. Depto et al. [18] explore various deep learning approaches designed to address imbalanced classification issues. AI techniques were used to assess the interpretability of these often considered "black box" models [19].

Certain researchers leverage the advantages of both machine learning and deep learning. They often use deep techniques for feature extraction and traditional algorithms for tasks like segmentation and classification. Elhassan et al. [20] utilized CMYK-moment localization and CNN-based feature extraction, while, Jha et al. [21] integrated segmentation from active contour and fuzzy C means, employing a Chronological SCA-based Deep CNN classifier for classification.

## 3. DATA DESCRIPTION

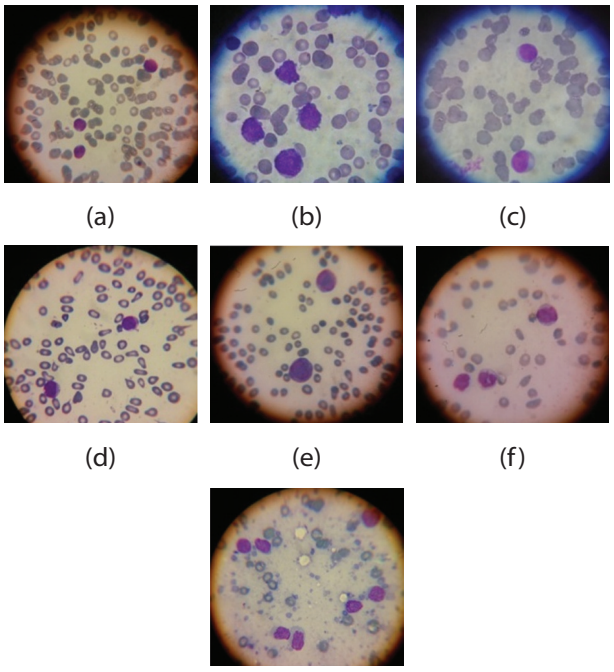
Publicly available blood smear datasets offer a valuable resource for investigating a range of blood-related disorders, like anemia, infections, and leukemia. Such datasets are instrumental in the development and validation of automated image analysis algorithms, machine learning models, and computer-aided diagnostic tools. This advancement enhances the precision and efficiency of blood smear examination and holds the potential to transform the field of hematological disease detection and classification. Additionally, these public datasets foster collaboration among researchers, encourage the establishment of standardized protocols, and facilitate the evaluation of various methodologies. As a result, patient care and diagnostic accuracy stand to benefit substantially from these collective efforts

This research utilizes the Raabin Health dataset [22], comprising 40,000 microscopic smears of WBCs from blood samples collected from patients who visited the Raabin collaborator medical laboratory in Tehran. The images are stained using the Giemsa technique and captured with a Zeiss microscope with a 100x zoom capac-

ity and LG G3 Smartphone. In this research a total of 800 blood smear images representing ALL, AML, and healthy samples have been included. Nuclei are automatically extracted from each image generating sub images. The dataset is split into 20% for testing and 80% for training. The distribution of images for each classification label is provided in Table 1, while Fig.1. illustrates sample leukemia images extracted from the database

**Table 1.** Distribution of Images

| Type | Sub Types | Images | Nucleus sub images |
|---|---|---|---|
| Acute Lymphoblastic Leukemia (ALL) | L1 | 100 | 217 |
| | L2 | 100 | 519 |
| Acute Myeloid Leukemia (AML) | M0 | 100 | 234 |
| | M1 | 100 | 204 |
| | M3 | 100 | 307 |
| | M4 | 100 | 339 |
| | M5 | 100 | 495 |
| Healthy | Healthy | 100 | 100 |
| Total | M6 | 800 | 2415 |



(a) (b) (c)

(d) (e) (f)

(g)

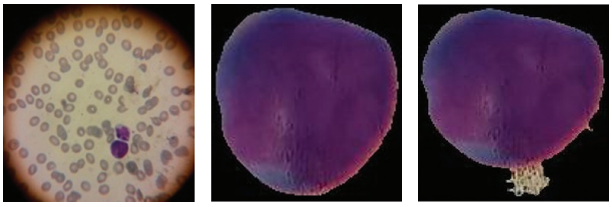
**Fig. 1.** Sample images from dataset (**a**) ALL-L1; (**b**) ALL-L2; (**c**) AML-M0; (**d**) AML-M1; (**e**)AML-M2; (**f**)AML-M3; (**g**) and AML-M5

## 4. PROPOSED WORK

### 4.1. SEGMENTATION AND NUCLEUS CROPPING

In medical image processing, resizing images is a standard preprocessing procedure aimed at ensuring uniformity in analysis and facilitating consistent feature extraction across diverse images, allowing for effective feature comparison. Hence, resizing operations are employed across all images to standardize their dimensions. The region of interest of our research is white blood cells, which need to be segmented from blood smear images. Initially, the procedure involves isolating the green channel data from the original RGB image because nucleus is more evident in the green channel [5].

A new image is created, preserving the original green channel data while substituting the red and blue channels with zeros, thereby retaining only the green channel's information. Median filter is applied to green channel image to reduce noise and improve quality of the image (see Fig 2). Subsequently, this newly formed image is converted from the YCbCr color space to the RGB color space. This transformation renders an image where the luminance data from the green channel is retained while lacking the color data from the chrominance (Cb and Cr) channels.



(a) (b) (c)

**Fig. 2.** Median Filter (**a**) Original Image; (**b**) With median filter; (**c**) Without median filter

To enhance contrast of the image, it is converted to grayscale and a histogram equalization technique is applied. The next step involves converting the image to a binary format using gray thresholding, which isolates the white blood cells while eliminating background elements like red blood cells and platelets. If white blood cells are connected or in proximity, further segmentation is necessary to differentiate them. The process includes employing the canny edge detection algorithm on the binary image to highlight edges. Morphological operations such as dilation are performed to enlarge the white regions, aiding in connecting fragmented edges and improving features. Filling operations are carried out to complete or bridge gaps in white regions, followed by erosion to decrease white region size, potentially separating interconnected regions and eliminating smaller objects.

A distance transform is calculated, and watershed segmentation is applied to segregate linked cells. The outcome of watershed segmentation undergoes a refinement process where undesired small objects are eliminated from a binary image. This is achieved by considering their area size, and the resulting refined output serves as a mask. This mask is then utilized to apply the Bounding Box technique for precisely extracting each nucleus from blood smear images. This comprehensive sequence of procedures enables effective identification and separation of white blood cells into sub-images. The results at all stages are collectively visualized in Fig 3 and fundamental stages of sub-image creation are outlined in Algorithm 1 as follows.

**Algorithm 1**

**Input**: Microscopic Blood smear images

**Output**: Nucleus sub image

1 Let $I_{resized}(x, y)$ represents the image $I(x, y)$ after resizing to a standard size

2 $G(x, y) = I_{resized}(x, y)\ green$
Where, $G(x, y)$ is isolated green channel

3 Convert the $E(x, y)$ to image while preserving green channel
$E(x, y) = \begin{cases} E(x, y) \\ 0 \end{cases}$

4 Noise reduction with median filter
$E(x, y) = Median\ E(x, y), [m,n])$
Where, $[m, n]$ defines median filter size 3×3

5 $J(x, y) = RGB\ (E(x, y))$
Where, $RGB\ (E(x, y))$ signifies transformation of $E(x, y)$ from YCbCr color space to RGB color space

6 Convert the RGB image $J(x, y)$ to grayscale image $K(x, y)$=grayscale($J(x, y)$)

7 Apply histogram equalization to $K(x, y)$

8 $B(x, y)$ - $Binary\ (K(x, y), T)$
Where, $Binary\ (K(x, y), T)$ applies binary transformation to each pixel in the grayscale image $K(x, y)$ using threshold value $T$

9 Edge detection using canny edge detector
$C(x, y) = edge\ (B(x, y), 'Canny')$

10 Performing morphological operation including dilation, hole filling, and erosion on $C(x, y)$
$R(x, y) \rightarrow Morphological\ operations\ (C(x, y))$

11 $T(x, y) = Distance\ Transform(R(x, y))$
$S(x, y) = watershed\ (T(x, y))$
Where $R(x, y)$ is binary image after morphological operation,
And $T(x, y)$ represent distance transform image, $S(x, y)$ is segmented image obtain by applying watershed segmentation algorithm

12 $M = Removesmallobjects\ (I, min\_area)$
Where $M$ is the refined binary image used as a mask to indicate region of interest

13 Boundary box is applied using generated mask $M$
$Nuclei = BoundingBox(I, M)$
Where $I$ is original image

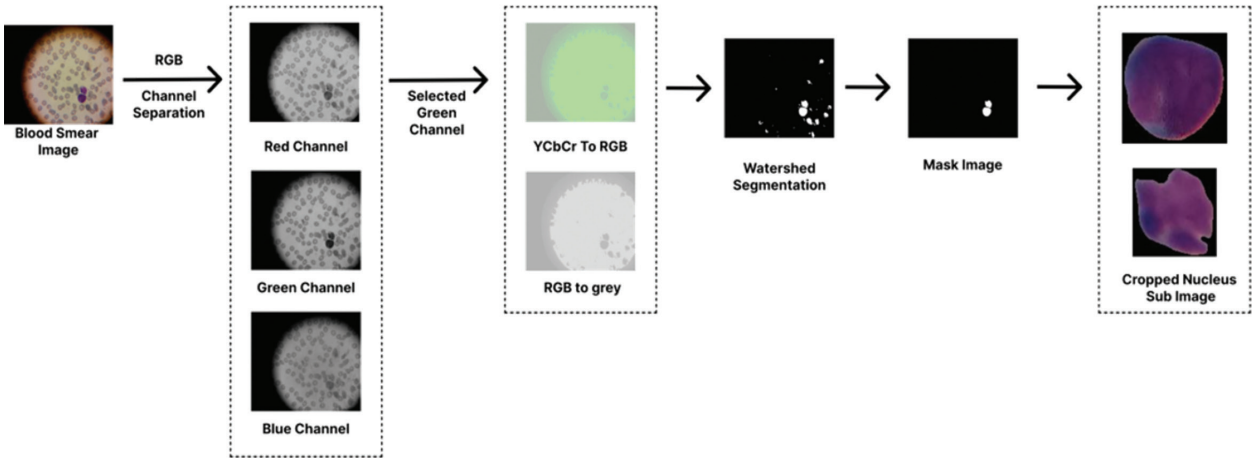



**Fig. 3.** White blood cell (WBC) segmentation and cropping

**Tabel 2.** List of features

| | Feature Method | Extracted Feature |
|---|---|---|
| Texture Features | Gray-Level Co-occurrence Matrix (GLCM) | Contrast, Correlation, Energy, Entropy, Homogeneity, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation1, Informaiton measure of correlation2 |
| | Local Binary Pattern (LBP) | LBP values for the histogram bins |
| Geometrical Features | | Area, Perimeter, EquivDiameter, EulerNumber, MajorAxisLength, MinorAxisLength, Solidity, Eccentricity, Circularity, ConvexArea, Extent |
| Color Features | Color Moments features in LAB color space | mean, Standard Deviation of Color Channels, skewness, kurtosis |
| Statistical Features | Higher-Order Statistical Moments | Skewness, Kurtosis |

### 4.2. FEATURE EXTRACTION

In medical image processing feature extraction is crucial for condensing intricate pixel data into a meaningful representation of patterns, textures, and anatomical structures. This research considers 48 features as listed in Table 2 extracted from nucleus sub-images contributing to leukemia detection from blood smear images. Subtypes of AML and ALL are identified based on morphological features. Shape, size and chromatic

pattern of nucleus vary for different subtypes. Some cells appear primitive while few cells show more mature cell features. Texture features, specifically Local Binary Pattern (LBP) and Gray-Level Co-occurrence Matrix (GLCM), amplify the effectiveness of leukemia detection from blood smear images. LBP excels in capturing intricate local texture patterns by comparing pixel intensity variations, while GLCM comprehensively analyzes spatial relationships between intensity values. Geometrical features provide quantitative insights into cell shape, size, and arrangement, incorporating parameters such as cell area, perimeter, circularity, and eccentricity, thus presenting a comprehensive depiction of cell morphology.

Color Moments features extracted from the LAB color space, significantly contribute to leukemia detection. LAB color space's separation of luminance (L) from chromatic information enhances robustness to lighting variations. These Color Moments, including mean, standard deviation, and skewness, effectively capture color distribution characteristics, thereby highlighting variations in cell staining and aiding the differentiation of normal and leukemia-affected cells. Moreover, higher-order statistical moments encompassing skewness and kurtosis go beyond mean and variance, revealing complex statistical patterns. These moments provide insights into distribution asymmetry and peakedness, uncovering subtle textural irregularities indicative of abnormal cell structures. By quantifying non-uniformities and deviations from a standard distribution, higher-order statistical moments further enhance the ability to differentiate between normal and leukemia-affected cells.

All the extracted features are normalized to ensure they are on similar scale to improve efficiency of classification algorithm and reduce the effect of features with large numerical values dominating the learning process. After features are extracted and normalized they are concatenated into a single feature vector. This feature vector contains unique characteristics of each image in a compact format. This diverse feature integration significantly enhances the precision and sensitivity of automated disease detection systems.

### 4.3. CLASSIFICATION

The utilization of the Random Forest algorithm presents a potent approach for the detection of leukemia from blood smear images. This algorithm, characterized by an ensemble of decision trees, is adept at handling the intricacies of medical image analysis. Random Forest algorithm can be formulated as:

$$Y_{prediction} = \arg\max_{j} \sum_{i=1}^{N_{trees}} I(y_i = j) \quad (1)$$

Where: $Y_{prediction}$ is the predicted class for the sample, $N_{trees}$ is the number of decision trees in the Random Forest. $Y_i, j$ is the predicted class of the i-th decision tree for class j. $I(.)$ is the indicator function that returns 1 if the condition is true and 0 otherwise.

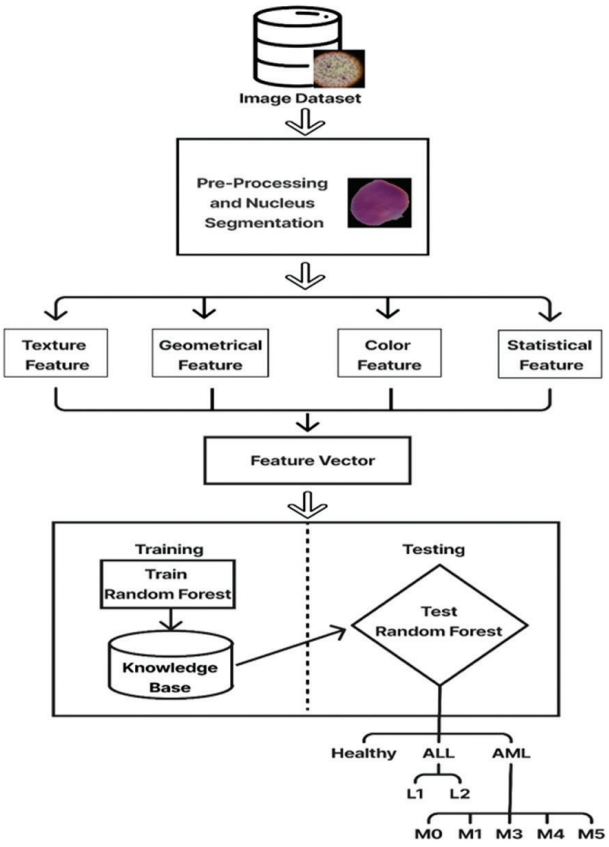



**Fig. 4.** Automated classification system for Acute leukemia

By employing features extracted from the images, such as texture features like GLCM and LBP, geometrical attributes, and Color Moments from the LAB color space, the algorithm can learn intricate patterns associated with normal and leukemia affected cells. The feature vector created by concatenating 48 features is given as an input for the classifier. Each feature is significant in capturing distinct aspects of cell morphology crucial for subtype classification in leukemia. All features are weighted equally in the classification process, reflecting their collective importance in discerning subtle variations indicative of different leukemia types and subtypes. The Random Forest model is trained on these features, and its collective decision-making process, combining outputs from individual trees, enables accurate classification. Through evaluation metrics like recall, precision, and F1-score, the effectiveness of the model is estimated, offering a robust and efficient tool for automated leukemia detection

## 5. EXPERIMENTAL RESULTS AND DISCUSSION

The comprehensive approach outlined for the detection of acute leukemia, including its types and subtypes, is visually represented in Fig 4. The focus of the proposed research is on multi-label classification, specifically involving the categorization of blood smear images into two main classes: healthy and cancerous. Within the cancerous category, a further distinction is made among different acute leukemia types and subtypes.

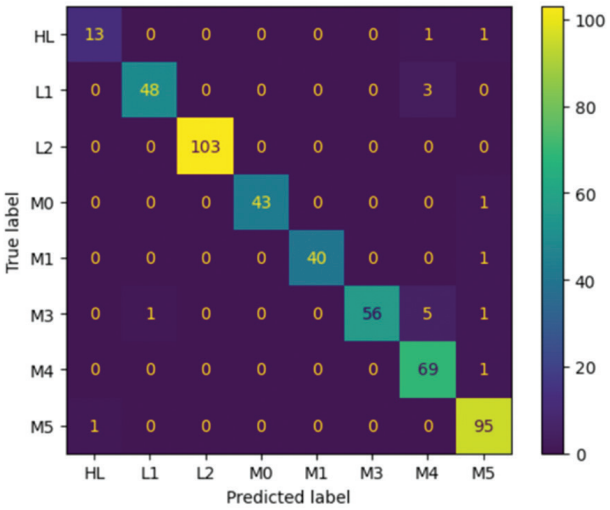
The effectiveness of the model is gauged through different performance parameters such as accuracy, precision, recall, and the F1 score. These results are presented in Table 3. For the experimentation, a dataset comprising 2415 nucleus sub-images is employed. To ensure robust evaluation, 20% of these images (483 in total) are designated for testing, while the remaining 80% constitute the training dataset. Among the testing nucleus images, 451 are successfully assigned to their respective categories. Notably, due to the presence of images with varying sizes for each label within the testing dataset, the consideration of the F1 score is crucial to accurately assess the model's performance.

**Tabel 3.** Performance measure

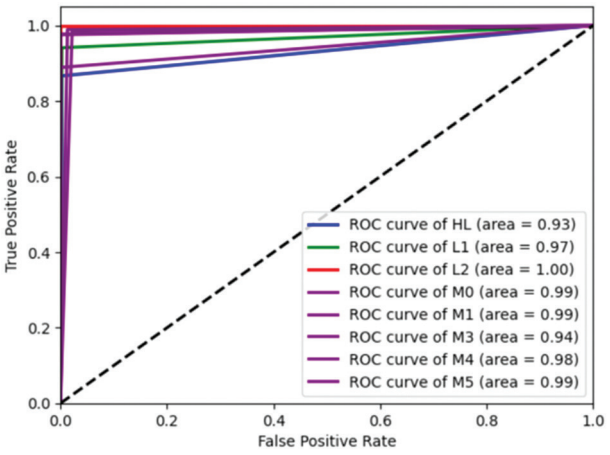
| | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| HL | 0.93 | 0.87 | 0.90 | 15 |
| L1 | 0.98 | 0.94 | 0.96 | 51 |
| L2 | 1.00 | 1.00 | 1.00 | 103 |
| M0 | 1.00 | 0.98 | 0.99 | 44 |
| M1 | 1.00 | 0.98 | 0.99 | 41 |
| M3 | 1.00 | 0.89 | 0.94 | 63 |
| M4 | 0.88 | 0.99 | 0.93 | 70 |
| M5 | 0.95 | 0.99 | 0.97 | 96 |
| Accuracy | | | 0.97 | 483 |
| macro avg | 0.97 | 0.95 | 0.96 | 483 |
| weighted avg | 0.97 | 0.97 | 0.97 | 483 |

The F1 score, a metric that harmonizes precision and recall, becomes particularly valuable when dealing with datasets characterized by imbalanced class distributions. This is a common scenario where one class vastly outweighs the other. Notably, the F1 score is calculated by balancing the trade-off between precision and recall. In this research, the F1 scores exhibit notable variation across the different categories. Specifically, healthy cells achieve an F1 score of 90%, while L1 and L2 types achieve 96% and 100%, respectively. Similarly, M0 and M1 categories have F1 scores of 99%, whereas M3, M4, and M5 exhibit scores of 94%, 93%, and 97%, respectively. In totality, the model achieves an overall accuracy of 97%. A comprehensive understanding of the model's performance, encompassing correctly classified instances, false positives, and false negatives, is provided through the confusion matrix (refer to Fig. 5(a)).

The ROC curve is created by plotting the True Positive Rate (TPR) on the vertical axis and the False Positive Rate (FPR) on the horizontal axis, while varying the threshold values (refer to Fig. 5(b)). Every point on the ROC curve corresponds to a specific threshold setting. The primary purpose of the ROC curve lies in evaluating the balance between sensitivity (true positive rate) and specificity (true negative rate) across different threshold choices. In addition to the curve itself, the Area Under the ROC Curve (AUC-ROC) is a widely utilized summary metric. The AUC-ROC value serves as an extensive assessment of the model's performance in distinguishing between classes, regardless of the specific threshold employed.




(a)




(b)

**Fig. 5.** (**a**) Confusion Matrix; (**b**) ROC Curve of classification output

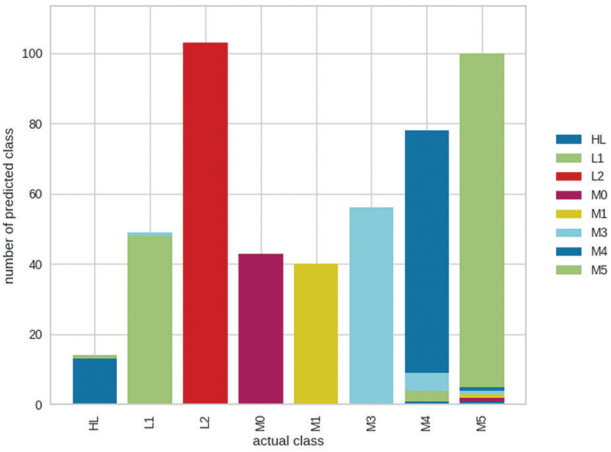



**Fig. 6.** Class Prediction Error Chart

Fig. 6 illustrates the class prediction error for the classifier. Each bar within the chart depicts the proportion of predictions for individual classes, offering a rapid comprehension of the classifier's accuracy in assigning correct classes. This model demonstrates proficient prediction for L2, M0, M1, and M3 subtypes based on

the feature vector. However, it exhibits some inaccuracies in predicting HL, L1, M4, and M5 types.

## 6. CONCLUSION

Acute leukemia is rapid progression disease that leads to generation of white blood cells which are abnormal and can potentially spread to other organs, causing health issues. Timely diagnosis and treatment are essential for effective management. This paper suggests an automated machine learning approach for the early detection of acute leukemia, as well as its classification into the two primary types, ALL and AML. Furthermore, it aims to categorize these types into their respective subtypes, which include L1, L2, M0, M1, M3, M4, and M5. The suggested model comprises three key phases. The initial phase involves Segmentation and Nucleus Cropping, which encompasses extracting the green channel from the original image and converting it to the YCbCr color space. A Median filter is then applied to eliminate any noise, creating a nucleus mask through watershed segmentation, and ultimately cropping the nucleus using a bounding box. The second phase focuses on Feature Extraction, encompassing the extraction of various features, such as texture, geometric, color, and statistical attributes, from the region of interest. The final phase involves Classification, employing a Random Forest classifier to categorize the samples into ALL, AML, ALL Subtypes, AML Subtypes, and healthy cells. With an impressive accuracy rate of 97% and hamming loss of 0.03%, this model has the potential to significantly assist pathologists and medical professionals in swiftly identifying cases of Acute Leukemic cancer, thereby facilitating prompt decision-making for accurate diagnoses. Furthermore, there is room for further enhancements and refinements in the methodology to potentially raise the classification accuracy even higher, ultimately improving the model's utility in clinical settings.

## 7. REFERENCES:


[1] Leukaemia, https://gco.iarc.fr/today/data/factsheets/cancers/36-Leukaemia-fact-sheet.pdf (accessed: 2023)

[2] D. A. Arber, A. Orazi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. M. Le Beau, C. D. Bloomfield, M. Cazzola, J. W. Vardiman, "The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia", Blood, Vol. 127, No. 20, 2016, pp. 2391-405.

[3] S. Shimony, M. Stahl, R. Stone, "Acute myeloid leukemia: 2023 update on diagnosis, risk-stratification, and management", American Journal of Hematology, Vol. 98, No. 3, 2023, pp. 502–526.

[4] K. Gupta, N. Jiwani, P. Whig, "Effectiveness of Machine Learning in Detecting Early-Stage Leukemia", Proceedings of the International Conference on Innovative Computing and Communications, Delhi, India, 19-20 February 2022, pp. 461-472.

[5] L. Putzu, G. Caocci, C. Di Ruberto, "Leucocyte classification for leukaemia detection using image processing techniques", Artificial Intelligence in Medicine, Vol. 62, No. 3, 2014, pp. 179-191.

[6] S. Mohapatra, D. Patra, S. Satpathy, "An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images", Neural Computing and Applications, Vol. 24, 2014, pp. 1887-1904.

[7] J. Rawat, A. Singh, B. Hs, J. Virmani, J. S. Devgun, "Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia", Biocybernetics and Biomedical Engineering, Vol. 37, No. 4, 2017, pp. 637-654.

[8] H. H. Inbarani, A. T. Azar, G. Jothi, "Leukemia image segmentation using a hybrid Histogram-Based soft covering rough K-Means clustering algorithm", Electronics, Vol. 9, No. 1, 2020, p. 188.

[9] T. G. Devi, N. Patil, S. Rai, C. S. Philipose, "Gaussian Blurring Technique for Detecting and Classifying Acute Lymphoblastic Leukemia Cancer Cells from Microscopic Biopsy Images", Life, Vol. 13, 2023, p. 348.

[10] M. Ashok, K. Tharani, S. VenkataSriram, K. Ramasamy, "An Investigational Study of Detecting Acute Lymphoblastic Leukemia using Computer Vision", Proceedings of the 2nd International Conference on Smart Technologies and Systems for Next Generation Computing, Villupuram, India, 21-22 April 2023, pp. 1-6.

[11] P. K. Das, D. V. A, S. Meher, R. Panda, A. Abraham, "A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia", IEEE Access, Vol. 10, 2022, pp. 81741-81763.

[12] R. Raina et al. "A Systematic Review on Acute Leukemia Detection Using Deep Learning Techniques", Archives of Computational Methods in Engineering, Vol. 30, 2023, pp. 251-270.

[13] S. Ansari, A. H. Navin, A. B. Sangar, J. V. Gharamaleki, S. Daneshvar, "A customized efficient deep learning model for the diagnosis of acute leukemia cells based on lymphocyte and monocyte images", Electronics, Vol. 12, No. 2, 2023, p. 322.

[14] L. Boldú, A. Merino, A. Acevedo, A. Molina, J. Rodellar, "A deep learning model (ALNet) for the diagnosis of acute leukaemia lineage using peripheral blood cell images", Computer Methods and Programs in Biomedicine, Vol. 202, 2021.

[15] S. M. Abass, A. M. Abdulazeez, D. Q. Zeebaree, "A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia", Indonesian Journal of Electrical Engineering and Computer Science, Vol. 25, No. 1, 2022, pp. 200-213.

[16] S. Saleem, J. Amin, M. Sharif, G. A. Mallah, S. Kadry, A. H. Gandomi, "Leukemia segmentation and classification: A comprehensive survey", Computers in Biology and Medicine, Vol. 150, 2022, p. 106028.

[17] P. Manescu et al. "Detection of acute promyelocytic leukemia in peripheral blood and bone marrow with annotation-free deep learning", Scientific Reports, Vol. 13, 2023, p. 2562.

[18] D. S. Depto, Md. M. Rizvee, A. Rahman, H. Zunair, M. S. Rahman, M. R. C. Mahdy, "Quantifying imbalanced classification methods for leukemia detection", Computers in Biology and Medicine, Vol. 152, 2023, p. 106372.

[19] W. H. Abir, M. F. Uddin, F. R. Khanam, T. Tazin, M. M. Khan, M. Masud, S. Aljahdali, "Explainable AI in Diagnosing and Anticipating leukemia using transfer Learning Method", Computational Intelligence and Neuroscience, Vol. 2022, 2022, pp. 1-14.

[20] T. A. M. Elhassan, M. S. M. Rahim, T. T. Swee, S. Z. M. Hashim, M. Aljurf, "Feature Extraction of White Blood Cells Using CMYK-Moment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images", IEEE Access, Vol. 10, 2022, pp. 16577-16591.

[21] K. K. Jha, H. S. Dutta, "Mutual Information based hybrid model and deep learning for Acute Lymphocytic Leukemia detection in single cell blood smear images", Computer Methods and Programs in Biomedicine, Vol. 179, 2019, p. 104987.

[22] Z. M. Kouzehkanan et al. "A large dataset of white blood cells containing cell locations and types, along with segmented nuclei and cytoplasm", Scientific Reports, Vol. 12, No. 1, 2022, p. 1123.

# FedExLSA: Design and Development of algorithms on Federated Data Exploration of Topic Prediction Using Latent Semantic Analysis

Original Scientific Paper

**Saranya M***
SRM Institute of Science and Technology, Kattankulathur
Research Scholar, Department of Computing Technologies, School of Computing
Kattankulathur, Chennai, India
sm2317@srmist.edu.in

**Amutha B**
SRM Institute of Science and Technology, Kattankulathur
Professor, Department of Computing Technologies, School of Computing
Kattankulathur, Chennai, India
amuthab@srmist.edu.in

*Corresponding author

***Abstract*** *– Every government in the world has multiple departments that must function and operate to address the various inquiries raised by the population. The government's diverse range of websites offers citizens a platform to submit inquiries, thereby facilitating the fulfilling of their requirements. Comprehending the subjects addressed in People Query is essential for government services. Unstructured query data is analyzed using extracting information from text techniques such as allocation of Latent diffuser (LDA) and analysis of hidden semantics (LSA). LSA outperforms other methods in terms of performance because of its minimal complexity and quick installation process. Research on decentralized learning techniques for natural language processing (NLP) is necessary due to concerns about limited data availability and privacy. Federated learning (FL) employs methods that enable different users to collectively train an integrated broad model while maintaining their information regionally stored and accessible. Nevertheless, the current body of literature lacks a thorough examination and evaluation of FL techniques. Data federation is an approach to data integration that allows the government to access and query data from multiple diverse sources as if they were a single, unified repository. Functioning as a form of data virtualization, it facilitates the creation of a comprehensive representation of data, thereby enhancing operational efficiency and the accuracy of decision-making. FedEx utilizes Federated Learning to apply topic modelling techniques to common NLP tasks. The proposed structure integrates the FL Methodology with Latent Semantic Analysis to deliver outcomes for intelligent data analysis and management.*

***Keywords***: *Natural Language Processing, Topic Modeling, Federated Learning, Latent Semantic analysis, Text Mining*

*Received: March 27, 2024; Received in revised form: June 7, 2024; Accepted: June 7, 2024*

## 1. INTRODUCTION

Open governments and easy-to-contact are usually the most effective. The government is one of several that have recognized the significance of this. This is demonstrated by the creation of the Government Website, which addresses the concerns and desires of the general public. Facebook, email, and a personal visit to the government center's office are some other ways to get in touch with the website. Researchers examine the Internet for possible data sources. Researchers can find data on websites in various ways. A key component is the website's focus on text messaging and its restriction on character counts. One more thing about the website is that it has an API that makes it accessible from any location in the globe. In addition, the Separate Website program has a large user base throughout several states. Public complaints in the government are also received by numerous regional offices through online platforms. By evaluating all the searches, the government might access the most recent data on the website information that the users themselves have contributed. If the government is serious about improving its performance, it should hear the recommendations, comments, and

opinions of its citizens. A great deal of information was retrieved from these individuals during the interrogation. It would take a very long time to read all of this in sequence. However, the government needs answers to these questions to move forward; therefore, it would be a great shame if they were disregarded. The results of this study could help the government do its job better. With public support, new policies can be sustained. Besides the accelerated procedure, the government also can deal with any concerns that might arise. Take the licensing procedure as an example. If someone has a problem with it, the government should try to fix it by making it more clear and transparent.

This investigation utilizes textual data. When text documents are grouped, overlapping data results. A substantial amount of ground could potentially be explored with a single inquiry. Consequently, this research employed topic modelling methodologies. The objective of this study is to investigate the application of latent semantic analysis (LSA) and latent Drichlet allocation (LDA) to topic modelling. A set of methods known as topic modelling is employed to uncover concealed subjects within a query [1]. There are two perspectives from which to examine topics: probabilistic and linear. Linear topic modelling is surpassed by positivity topic modelling. This is illustrated by the "Latent Semantic Indexing" linear topic model. The technique in question is referred to as "Latent Semantic Analysis" (LSA). Probabilistic topic modelling is illustrated by the works of Latent Drichlet Allocation (LDA) and Probabilistic Latent Semantic Analysis (PLSA) [2]. The LSA model yields results that do not account for the correlation between the query and the corpus. Using LDA, one could investigate the interrelationships between the documents in the corpus. Numerous scholars have implemented LDA on datasets other than Query data. By applying the LDA method for topic modelling, [2] was able to ascertain the content that was discussed in People Queries about two distinct industries.

One way to train ML models [3] collaboratively without revealing any local data is by using Federated Learning (FL). It often requires several clients to collaborate with one or more servers that mediate the setting of agreements, privacy assurances, and the aggregation of node updates. Many researchers are looking into FL's possible application in topic modelling because of its privacy-preserving and decentralized data-leveraging features. Many scholars have focused on developing LDA-like or federated frameworks [4], whereas others have proposed federated general-purpose topic models [5]. FedLSA, also known as Federated Latent Semantic Analysis, contributes significantly to the large field of data science and natural language processing (NLP) research by delivering innovative solutions to pressing challenges. FedLSA is a system that allows for the analysis of distributed text data while adhering to tight privacy constraints, which is especially significant in an era where privacy is becoming increasingly crucial. Furthermore, as the number of edge devices and Internet of Things (IoT) technologies grows, its decentralized learning method enables collaborative analysis across several data sources without centralizing the data. This not only ensures the ability to handle large volumes of text data but also encourages the creation of applications that can be utilized in various fields, including healthcare, finance, and social media analysis. FedLSA marks a fundamental shift in data science and NLP research. It encourages collaboration and information sharing through federated analysis while maintaining data privacy. This technique user in a new era of collaborative, privacy-preserving analysis. However, no research has been conducted on federated latent semantic analysis implementations. Finding the best topic modelling technique for data derived from citizen queries in government is the goal of this research. Topic prediction often requires a large training dataset, which is not always readily available.

Below is a summary of the work's contributions:

- Our proposed framework is Fed LSA, which stands for federation. With its help, a large number of users could train a topic model with LSA and SVD.
- Fed LSA to perform even better, recommend combining it with machine learning. The goal of this method is to increase the degree to which text features resemble their abstract counterparts.
- Using the DigiLocker NAD, IHMCL, Kerala Startup Mission, and KILA datasets all of which are publicly available online, test our methods to make sure it works and also investigate its performance in various federated settings.

In this, Federated Latent Semantic Analysis (FedLSA) provides a strong framework for studying latent themes in text data from many government departments while protecting data privacy and confidentiality. FedLSA uses secure communication and aggregation techniques to allow for collaborative data analysis without centralizing sensitive information. The benefits of greater privacy, cooperative insights, and scalability make FedLSA an interesting technique for governmental data analysis, despite potential downsides like as communication costs and aggregation problems.

## 2. LITERATURE SURVEY

Text mining is the practice of using appropriate analysis to extract valuable information from a database of documents. Text mining uses query data, which is essentially unstructured text data. Data sources can be mined for important information using extraction, which involves discovering and analyzing interesting patterns. Using preprocessing procedures, text miners can convert the query's unstructured data into an intermediate form that is more specifically arranged [6]. The primary objective of preprocessing is to enhance accuracy. There are several People Queries collections on sites like the web. In particular, there is a growing need for automated

methods that can read, evaluate, and summarize large document sets. Various topic modelling techniques are used in numerous applications. Each topic modelling technique differs from both centralized and federated learning. Table 1 displays the comparison.

**Table 1.** Comparison of the various topic models

| Technique | Context | Advantages | Disadvantages |
|---|---|---|---|
| LDA | Centralized | Topics that could be easily understood and are commonly utilized | Requires significant computational resources and highly influenced by hyper parameters. |
| GD | Centralized | Efficiently computes straightforward and positive outcomes, capable of processing matrices with many zero elements | Requires the specification of the number of Topics and the convergence of local minima. |
| LSA | Centralized | The concept is straightforward and does not require a specific Topic number. | Assumes linear correlations, more difficult to analyze |
| RmsProp | Centralized | Documents the progression of a Logic | Requires data with time stamps due to its high level of intricacy. |
| Federated LDA | Federated | Ensures confidentiality, decentralized analysis | Overhead in communication and intricate aggregation |
| Federated GD | Federated | Ensures confidentiality, manages decentralized information | Secure aggregation is necessary to address convergence difficulties. |
| Federated LSA | Federated | Efficient and expeditious data privacy | There are difficulties in communication and aggregation. |
| Federated RmsProp | Federated | Effectively manages differences in local datasets between clients. | Requires a high level of intricacy and necessitates synchronization. |

Each topic modelling approach has a unique set of tradeoffs. Centralized approaches such as LDA, GD, LSA, and RmsProp are effective, but they might be constrained by computing resources and data privacy issues. Federated techniques alleviate privacy concerns and enable collaborative study of distant datasets; however, they add complexity to communication and aggregation.

As proposed by [7], the objective of Federated Topic Modelling (FL) is to train a model effectively using decentralized data from multiple clients. To reduce communication costs without sacrificing performance, numerous algorithms have been developed, including Fed Avg [8] and FedRmsProp [9]. In general, these algorithms operate in a two-step process: initially, they train the model using local data while simultaneously synchronizing the server with the latest model weights. Each client transmits the modified weights to the server for aggregation once training is complete. These potent FL algorithms are compatible with our Federal LSA architecture. Federated topic models have received limited attention from researchers. The federated topic modelling approach, as illustrated in reference [10], showcases the implementation of novel methodologies including heterogeneous model integration and topic-wise normalization. They implemented an innovative local differential privacy (LDP) technique to federate the LDA. The focus of this study is topic modelling via LDA. The focus of this study, on the other hand, is topic modelling in federated environments via LSA. Recently, there has been much buzz around latent topic modelling, which is an unsupervised approach to topic discovery in large document collections. An example of such a model is the LDA [11]. Using statistical (Bayesian) topic models, Latent Drichlet Allocation (LDA) is a well-liked approach for text mining. A generative model of writing is what LDA achieves. Consequently, it strives to generate a document that is relevant to the given subject. This approach can also process other types of data. Various methods, such as latent Drichlet co-clustering, topic modelling, author-topic analysis, and temporal text mining, use topics to express queries; each topic is a discrete probability distribution that specifies the likelihood of each word appearing in that topic. These subject probabilities can be used to describe a document. In this sense, a "Query" is just a "bag of words" sorted just by Topic and word count. As shown graphically in Fig. 1 and the LDA Mathematical Notation shown in Table 2. The LDA is an example of a generative probabilistic model, the next step for LDA to produce a specific corpus is to follow the following procedure:

1. Choose a distribution, $\theta t \sim D(y)$, over words for any subject t, where t is in the interval {1… t}.
2. For every query $Ds$, where q is an integer from 1 to q, Select a distribution where topics are $\delta q \sim D(\alpha)$. As if they were random variables obeying Drichlet distributions with parameters $x$ and $y$, respectively, the word distributions for themes and the document topic distributions were considered. The likelihood of the corpus in Equation 1 is given by a set of $Q$ Queries denoted as $D = \{D1, D2, \ldots, DQ\}$.

$$p(D|x,y) = \prod_{q=1}^{q} \int p(\alpha|\beta) \prod_{w=1}^{W} \sum_{a=1}^{t} (p(aqw|\delta q)p(\theta a|y))d\alpha q, d\alpha a \quad (1)$$

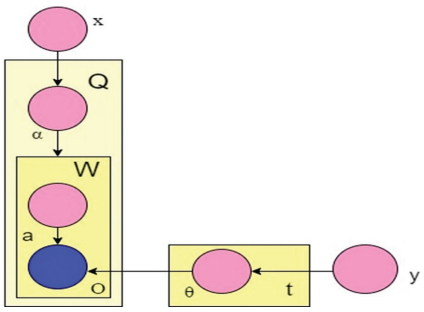



**Fig.1.** Graphical Representation of the LDA Model. The blue shade represents the Observed Model. Pink Shaded represent the Latent Variable.

The corpus was subjected to multiple runs of the LDA algorithm with varying numbers of topics. They started each experiment with the two hyper parameters set to $x = 0.1$ and $y = 0.01$

Singular Value Decomposition (SVD) is also utilized by Latent Semantic Analysis (LSA) to reorganize information. Using a matrix-based technique, SVD reorganizes and calculates all contractions of vector space. In addition, compute the reductions in vector space and arrange them in descending order of importance. The meaning of the text can be inferred using the most significant assumption if the LSA [12] assumption phase does not use the least important assumption. Finding words with a similar vector is one approach to finding words with many similarities. An important initial stage in LSA is to gather a large amount of relevant content and arrange it according to topics. In the second step, create a matrix that shows how often each word and document appears. Kindly provide the cell names (e.g., "document a," "terms b") and dimensional values (m for terms and n for documents) for each entry so that there is no room for misunderstanding.

**Table 2.** Table of Mathematical Notation in Latent Drichlet Allocation

| Notation | Description |
|---|---|
| Q | Number of Queries in Corpus |
| N | Number of topics |
| W | Number of words in one Query |
| X | hyper Topic-specific parameter (if symmetric scalar vector t) |
| Y | pre-word distribution hyper parameter (if symmetric scalar vector) |
| α | Topic Mixture ratio QXT Matrix (one row per Q Query) |
| θ | Word distribution TXS Matrix ( S is the size of Vocabulary) topic t with θt |
| a | topics generating word Q vectors(one per w words)value of at is in (1…t) |

Running the calculations and making adjustments to each cell is the third stage. To conclude, SVD is going to be an enormous assistance in calculating all the diminutions and creating the three matrices. The SVD operational principle was discussed in Section III.

This research has connections to three different fields: federated learning, similarity information, and LSA-based [13] topic modelling. a) One method for training models proposed by [14] is FL, which stands for Federated Topic Modeling. Its objective is to facilitate the effective utilization of distributed data among several People. Since its inception, numerous algorithms have been developed to reduce communication costs without compromising performance. These include FedAvg's and FedRmsProp. There are usually two steps to these algorithms in a typical implementation: first, the clients use local data to train the model while synchronizing the most recent model weights with the server. After training is completed, the clients send the updated weights back to the server for aggregated data. These cutting-edge FL algorithms are compatible with the proposed Fed LSA design. Few academics have focused on federated topic modelling. As an illustration, federated topic modelling has been described in [15]. It integrates innovative approaches such as topic-wise Semantic analysis, private Metropolis-Hastings, and heterogeneous model integration. Constructed federated LDA using a novel local differential privacy method. The major focus of this study is LDA-based topic modelling. In contrast, this study delves into LSA-based topic modelling in federated settings.

## 3. METHODOLOGY

Federated Latent Semantic Analysis (FLSA) is an approach that ensures the preservation of privacy while analyzing vast quantities of text data distributed across multiple clients, including institutions and smartphones. The methodology commences by assigning unique local datasets to each client and initializing local models with shared initial parameters. At the local level, individual clients perform data preprocessing, compute Term Frequency-Inverse Document Frequency (TF-IDF) vectors, and employ Singular Value Decomposition (SVD) to extract latent semantic structures. Based on their LSA results, clients subsequently generate local model updates and implement privacy-preserving strategies. Local updates are transmitted to a centralized server, which securely aggregates them while maintaining the confidentiality of individual data. By merging local modifications, the server modifies the global model and returns the updated model to the clients. The updated global parameters are subsequently incorporated into the local models by the clients, thereby enhancing the local LSA tasks. The process is iterative, consisting of multiple iterations of local processing and global aggregation. This iterative approach progressively enhances the global model, thereby improving the accuracy and generalizability of the latent semantic structures, all the while safeguarding data privacy. Several FedLSA stages are illustrated in Fig. 2.

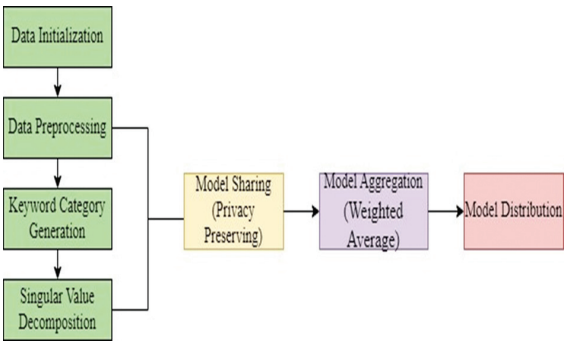



**Fig 2.** Stages of Federated Latent Semantic Analysis

By utilizing Latent Semantic Analysis (LSA), the dimensionality of a document representation is diminished. In a word vector, LSA employs a vector comprising latent semantic concepts. A large word-document matrix is subjected to singular value decomposition (SVD) by LSA [14] to reduce the dimensionality of the data.

Three matrices comprise a massive term-document matrix: one for documents, one for singular values, and one for concepts and terms. To reduce the dimension of the word document matrix, singular value decomposition (SVD) is abstained from in this instance. This methodology is founded upon two fundamental assumptions: (1) the number of subjects addressed in each document and (2) the vocabulary size corresponding to each subject as determined by Equations (2) and (3), respectively. Let us consider two variables: the number of subjects (T) and the size of the vocabulary (V). Within the given context, the notation [3] (t, d) signifies the occurrence of topic t in document d, whereas γ (w, t) denotes the creation of term w by topic t. The following is one possible configuration for the two assumptions:

$$r\,(t|dr) = \mu_{(t,d} \ \sum_{w\in T}^{t} \ \mu_{(t,dr)} \ = 1 \tag{2}$$

$$r\,(w|t) = \beta_{(w,t)} \ \sum_{w\in V}^{v} \ \beta_{(w,t)} = 1 \tag{3}$$

Additionally, the topic of the paper and each word are generated using Singular Value Decomposition. Finding two semantic vector matrices $A$ and $B$ for a matrix $M$ such that $M{\sim}AB$ and two matrices $W$ and $H$ such that $M \approx WX$ are the three main goals of LSA. Cutting down on the following $L\,(\varphi)$ loss concerning $W$ and $X$ is an easy way to do it.

$$L\,(M, W, X) = (M_{i,j} - M\hat{}_{i,j})^{2} \tag{4}$$

It is possible to express each document using a count (column) vector overlaid on top of the bag-of-word representation. Concerning the i-th client, $M_i$ represents the count feature matrix for documents. The union of all matrices with $i = 0,..., X$ allows for the decomposition of this matrix.

### 3.1. FEDERATED LSA

The FedLSA factorization procedure for client-distributed matrices is illustrated in Fig. 3. Use the GD algorithm to minimize loss within the federated learning architecture. However, as shown in Section IV studies, using FedAvg's approach to optimize the loss on each client alone results in poor topic models. Below are the factors that FedLSA follows. Imagine a network of X client devices, where the i-th device's data distribution function (ddi) might vary for different values of i. In distributed learning environments, X clients are usually trained using a single global model. Finally, under the assumption of a Federated Semantic Analysis (FedLSA) architecture with layers $= 0,..., L$, all clients share the set of weights $\varphi = \{W\}\ L{=}0$. Mastering the art of limiting the average loss for every client could help achieve the global goal. This is the general principle behind many federated learning approaches.

For instance, fedLSA aims to minimize the following goals in Equation (5):

$$Min\ L\,(\varphi) = \sum_{i=0}^{X} Ni\ Li(\varphi) \tag{5}$$

The Weight of each device $i$, represented as $Ni > 0$, and the number of Peoples, $X$, are input into the local objective function, $Li\,(\varphi){:=}\ Pxi{\sim}ddi\ [li\,(xi; \varphi)]$.

Unfortunately, statistical heterogeneity means that there is no silver bullet when it comes to fitting the global model to individual clients. This, in turn, impacts the degree to which a client's local distribution resembles the population distribution. People who share fewer attributes might view this strategy as unfair. In comparison, $X$ local models $= \{W'\ i\}\ L = 0$ are learned, where each model is trained using only ddi.

The data distribution of each client i determines the set of weights $\chi i$ to the maximum extent possible. Considering that each client typically has limited data that may not be sufficient to train a comprehensive model without over fitting, the total number of parameters that must be learned across all clients increases as $X$ decreases. Using shared learning problems or comparable client data distributions, simultaneously learn $X$ distinct models.

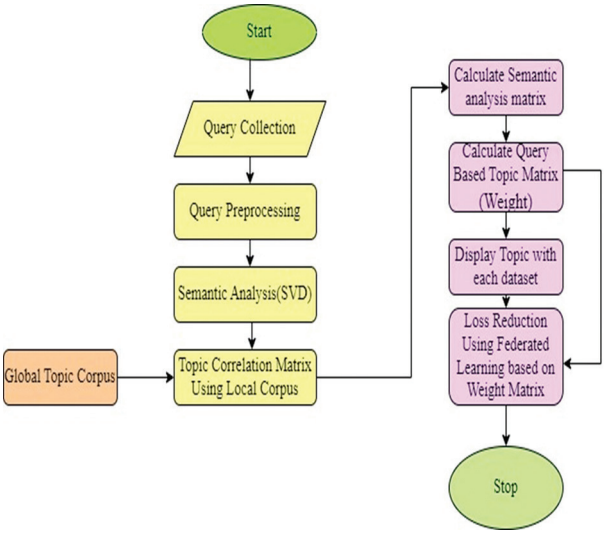



**Fig. 3.** Architecture of Client distributed Matrices using Federated LSA Model

Aiming for a compromise between $X$ distinct local models and one global model is the optimal approach to data utilization. Ensure that all models utilize the same vocabulary of combined components, but have each client modify their model by their specific distribution in their local area. By factorizing the subsequent equations (1) and (2) using layer-wise decomposition, they construct each weight matrix.

**Algorithm1.** Federated Latent Semantic Analysis for $X$ Number of Clients Communication

**Server Side Execution**

1 Server Executes:
2 Initialize $W\,(0)$ and $\varphi$
3 for each round $r{=}$ 0, 1,2 . . . do
4 $Qr \leftarrow$( group of $X$ clients)
5 for each client $i \in Qr$ in parallel do
6 ($W(r{+}1)$
$i, \varphi\ (r{+}1)$
) ← Update ($i$, $W(r)$, $\varphi\ (r)$)

7 end for

**Client Side Execution**

1 Client Update ($i$, $W$, $\varphi$): //Run on client $i$
2 for each local updation from 0 to $U$ do
3 For each $Wx \in \varphi$
4 $\varphi M$ Compute the Similarity Measure
5 loss $L$ is defined in Eq. (3)
6 return ($W$, $\varphi$) to serve
7 end for each
8 end

Algorithm 1 displays the entire pseudo code of the proposed FedLSA framework. This method describes execution on the server and client sides. Existing Federated learning algorithms like FedAvg's and FedRmsProp are compatible with this architecture, which is dubbed FedLSA.

## 4. EXPERIMENTAL RESULTS

Demonstrate the efficacy of our algorithms by running them on many publicly available datasets and comparing their results with those of the state-of-the-art FedLSA. Python Tensor Flow is used to implement all the models.

### 4.1. DATASETS

In the tests, the four real-world text datasets represent People's Queries related to different topics. Tesz, which stands for "Questions and Answers in Various People Queries," is associated with the following four datasets.

a. Digi locker NAD: This dataset is a subset of Dig locker, which contains user queries about many domains, including education and various schemes. Around 1,250 authentic and encrypted queries written in English make it up.

b. IHMCL: Transportation-related queries, such as Fas tags, are a component of the Government Highway Management, which includes this data set. It has 800 queries written in English and authorized by users.

c. Kerala Startup Mission: One thousand queries for the Kerala State's Medical and Educational Schemes are contained in this collection. The Government of Kerala was responsible for its upkeep.

d. KILA: This is a database of 2000 questions about various forms of education (seminars, conferences, workshops, etc.) that have been posted on People. The Kerala government made this dataset available.

Table 3 shows the fundamental statistics of the datasets. "Files" indicates the total number of queries in the Records dataset, "terminology" indicates the total number of terms in the dataset, and "types" indicates the total number of categories in the dataset. A dataset's "Record length" is its mean Record length. After removing stop words and tokens using text preprocessing techniques, the figures were calculated. For our experiments, they used 70% of the datasets for training and 30% for testing.

**Table 3.** Data Preprocessing

| Dataset | Files(Queries) | Terminology | Types(Fields) |
|---|---|---|---|
| DigiLockerNAD | 1200 | 2300 | 3 |
| IHMCL | 800 | 1280 | 4 |
| Kerala Startup Mission | 1000 | 1600 | 2 |
| KILA | 2000 | 3460 | 5 |

### 4.2. COHERENCE METRICS

For query-based data, three separate types of coherence metrics should be used: PMI, LSA, and Word Embedding (WE) [11]. Here, we will go over the 9 metrics that come out of these measurements. It begins by outlining the existing PMI- and LSA-based metrics for theme consistency assessment and then provides a novel Word Embedding-based statistic. The top $n = 20$ words ({$w1$, $w2$,..., $w10$}) chosen based on their probabilities ($p$ ($w|z$)) in the collection $\mu$ could represent a topic t inside this subject. One could determine the coherence of a topic by averaging the semantic similarity of the word pairs related to it (Equation (6).

$$Topic\ Coherence(x) = 1/\sum_{x=1}^{y-1} a = x + 1CS\ (w_a, w_b) \quad (6)$$

Demonstrated that the pair-word PMI might represent the coherence [16] of topics identified in both the standard and Query corpora. For additional accuracy, they could use Equation (6) to determine how similar wa and wb are. This is accomplished by pulling co-occurrence statistics from a backdrop corpus that contains DigiLocker NAD, IHMCL, KILA, and the Kerala Startup Mission. Equation (1) is also used for this objective.

Note that to precalculate the PMIs of word pairs, certain extra datasets are required. Finding out how similar two-word pairings are in meaning is another usage of LSA [17]. Words are represented by dense vectors in the reduced LSA [18] space ($Vxi$) to apply LSA. To obtain this vector from a background corpus, Singular Value Decomposition is employed. The degree of similarity between two words can be assessed using the LSA metric [19], which uses a cosine function to calculate the distance between the word vectors. The following is substituted into Equation (6) within Equation (7) to accomplish this:

$$C(w_a, wb) = PMI(w_a, wb) = logp(w_a, wb)/p(w_a)Xp(w_b) \quad (7)$$

As indicated before, word embedding is more accurate than LSA. Table 5 displays the results of the coherence score for several topic modelling techniques. The topics generated by federated learning approaches, such as FedLDA, FedAvg, FedGD, and FedRmsProp algorithms, are of high quality. These algorithms provide instances of topic terms on the dataset. Table 5 demonstrates that FedLDA tends to produce repetitive subjects.

While the topics of FedAvg and FedRmsProp appear diversified, they are less informative and lack coherence. The Topic Quality of FedLSA is superior to those of other methods. Fig. 4 displays the coherence score graphically. Not long ago, this investigation was conducted. Application of Word Embedding vectors Vwa, which are obtained using a Word Embedding model that has been pre-trained on a large text dataset. The Topic Coherence Notation is displayed in Table 4.

**Table 4.** Table of Notations for Topic Coherence

| Notation | Description |
|---|---|
| $CS$ | Coherence Score |
| $PMI$ | Point Wise Mutual Information |
| $W_a$, $w_b$ | Two Different Kind of Words |
| $Vxi$ | Vector Space |
| $p(w_a)$, $p(w_b)$ | Probability of a Particular word |

According to Equation (8), the cosine similarity of two words' word vectors is larger when the words have a comparable semantic meaning.

$$CS\ (wa, wb) = cosine\ (Va, Vb) \tag{8}$$

**Table 5.** Coherence Score of Topic Modeling Methods

| Technique | DigiLockerNAD | IHMCL | Kerala Startup Mission | KILA |
|---|---|---|---|---|
| LDA | 0.524 | 0.410 | 0.523 | 0.467 |
| LSA | 0.536 | 0.456 | 0.510 | 0.426 |
| LDAGD | 0.517 | 0.345 | 0.444 | 0.356 |
| LSAGD | 0.578 | 0.444 | 0.543 | 0.432 |
| FedLDA | 0.657 | 0.334 | 0.437 | 0.543 |
| FedLSA | 0.432 | 0.523 | 0.324 | 0.326 |
| FedGD | 0.523 | 0.432 | 0.324 | 0.467 |
| FedAvg | 0.434 | 0.433 | 0.456 | 0.345 |
| FedRMSprop | 0.343 | 0.435 | 0.439 | 0.346 |

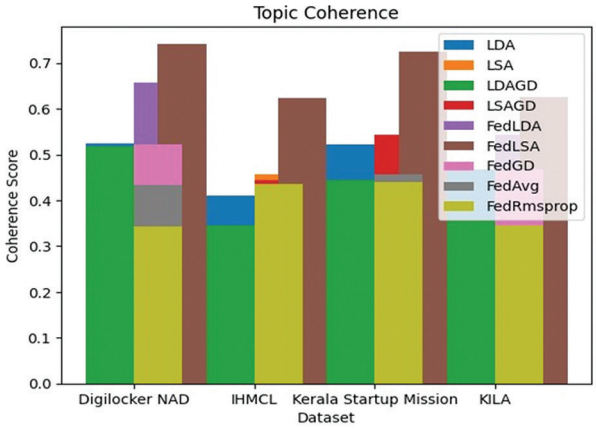



**Fig. 4.** Topic Coherence of Various Topic Modeling Techniques

They examined the similarity between WE, PMI, and LSA measures and human evaluations using the methodologies described in Section III. They also examined whether the WE-based metric could capture the coherence of People Queries topics.

### 4.3. COMPARISON METHODS

From centralized to federated, these are the topic modelling strategies they used in our studies.LDA and LSA are examples of traditional topic models used with centralized text data.

- Centralized GD-based Methods: To confirm that Semantic analysis is effective for the centralized LSA topic modelling, they incorporate GD-based LSA [9] methods (LSA+GD) into our trials. The main idea of LSA+GD is to maximize LSA's least square loss using mini-batch GD.
- Three federated topic modelling approaches are put into practice by us: FedAvg, FedGD, and FedRmsProp; FedLSA, which is based on variational inference; and FedRmsProp.

### 4.4. EXPERIMENTAL SETUP

In this experiment, they constructed datasets for several clients according to the methods described in [20]. For the sake of precision, it is assumed that the training samples for each client are chosen at random with class labels based on a categorical distribution over l classes, where v is a vector with elements ($vi > 0$, $i \in [1, l]$ and $Pvi = 1$). They pull $v \sim Dir$ (μq) from a Drichlet distribution, where q is the label distribution of a specific dataset and $\beta$ controls the degree of client identity, to generate a set of clients that are not identical. Every client has the same distribution relative to q as $\beta$ gets closer to infinity. On the other hand, as $\beta$ gets closer to zero, each client only saves instances from one label. To conduct our experiments, they manipulated the heterogeneity of the client data using $\beta$ and generated different FL settings by changing the client number $N$. Throughout the experiment, they allocated $N$ to the set {10, 20, 30, 40}. They divided the overall sample size by the People number $N$ to obtain the number of documents (Queries) given to each client. This ensured that our results would be similar. They then create a test and training set using the aggregated topic weight vectors from all documents. They then used the findings to calculate the accuracy and macro F1 score using a Logistic Regression (LR) classifier. Federated topic modelling is the next step in this procedure. By adjusting the value of n to 10, 20, 30, and 40 for all datasets, thorough results were obtained. This federated topic modelling method uses a constant participant fraction of $P = 1$ throughout all iterations. With each cycle, they tweaked the local batch size from the set {20, 40, 60, and 80} and the local GD training epoch count from the set {10, 20, 30, and 40}. When FedAvg and FedRmsProp are run by default, the hyper parameters are set to [21].

## 5. RESULT AND DISCUSSION

Following this, discuss more about the topics produced by the top model, which is the optimized LSA model that performed best across all of these criteria in terms of the topic Coherence Score. Not only are the five resulting subjects easily distinct but they are also

thoroughly relevant and cohesive. Look at Table 6 for the subject keywords for every single idea. Additionally, it delves further into each subject: • DigiLocker Non-Disclosure Agreement: This covers matters about the realm of education and student conduct. The fact that users provide comments on elements about education suggests a strong connection with the Dig Locker app, and this app in particular. User comments regarding app performance and troubleshooting in the education domain are the focus of this section.

- IHMCL People Support: This section addresses Transportation-Related Questions, Concerns, and management, service-related problems, fast tags, and Toll Information. The significance of dependable and trouble-free transportation services for individuals is showcased. Regarding the Kerala Startup Mission, this section addresses public comments on health and education initiatives. Surprisingly, it is highly related to the knowledge and expertise acquired by the Kerala Government's various schemes. Disagreements over the app's usability and user interface are widely discussed.
- KILA: This category houses all of the questions that attendees of educational events like seminars, conferences, and workshops may have.

**Table 6.** Keywords related to each Topic based on the Dataset

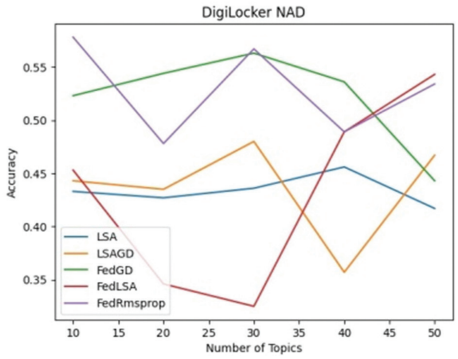
| Topic | Topic Keyword |
|---|---|
| 1 | ['Student'' Certificate', 'Digital''new', 'download'] |
| 2 | ['People', 'Transport', 'Renew', 'FasTag', 'car'] |
| 3 | ['Citizen', 'Medisep', 'Scheme'' Scholarship'] |
| 4 | ['Education', 'Conference', 'Seminar'' workshop'] |

Tables 5 and 7 show the results of text classification metrics and coherence scores for different topic modelling methods across the four datasets and the graphic representation shown in Fig. 5. On four separate topics ($X$=10, 20, 30, and 40), they averaged the provided coherence scores, F1 scores, and accuracy values. The setting for the FL environment is $\varphi = 1.5$ and $X = 10$.

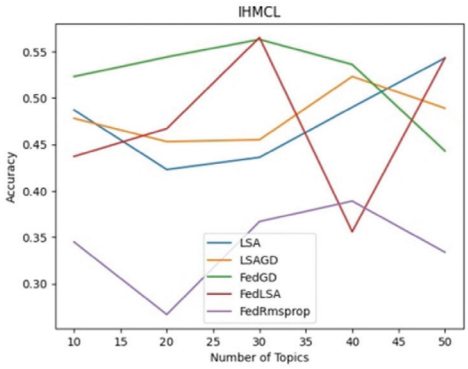
These are the key points to remember. First, these papers are typically shorter than 20 words in length, and when they compare classical LDA with LSA, They see that LSA performs better on all datasets. Therefore, when it comes to People Query data, LSA typically performs better than LDA. 2) Modeling LSA topics in federated and centralized environments. The coherence score and classification both reveal this. As an example, Table 5 shows that across all four datasets, LSA+GD produces consistently higher F1 scores than LSA+GD in centralized learning, with the 15% gap being most pronounced on KILA. The efficacy of LSA-based topic modelling is demonstrated. 3) On all four datasets, FedLSA techniques (including FedAvg, FedGD, and FedRMSProp) outperform FedLDA among the federated topic models. Both the F1 Score and the Accuracy display this. Table 7 shows that across all four datasets used for centralized learning, LSA+GD consistently outperform LSA in terms of F1 scores.

**Table 7.** Evaluation Metrics for Four Different Datasets with Four Topic

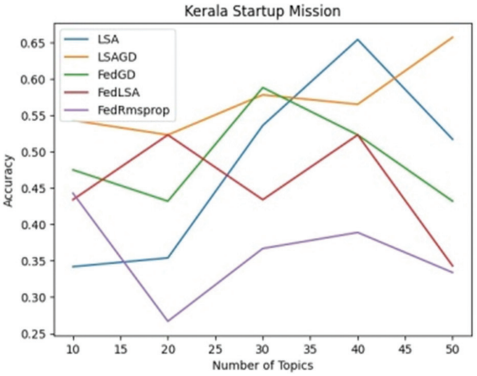
| Dataset Metrics | DigiLockerNAD F1score | DigiLockerNAD Acc | IHMCL F1score | IHMCL Acc | Kerala StartupMission F1score | Kerala StartupMission Acc | KILA F1score | KILA Acc |
|---|---|---|---|---|---|---|---|---|
| LDA | 0.424 | 0.433 | 0.487 | 0.423 | 0.342 | 0.354 | 0.468 | 0.456 |
| LSA | 0.436 | 0.456 | 0.490 | 0.543 | 0.536 | 0.654 | 0.593 | 0.482 |
| LDAGD | 0.417 | 0.443 | 0.478 | 0.453 | 0.517 | 0.543 | 0.467 | 0.432 |
| LSAGD | 0.435 | 0.480 | 0.455 | 0.523 | 0.578 | 0.565 | 0.478 | 0.453 |
| FedLDA | 0.357 | 0.467 | 0.489 | 0.437 | 0.657 | 0.475 | 0.512 | 0.342 |
| FedLSA | 0.523 | 0.554 | 0.467 | 0.565 | 0.432 | 0.588 | 0.489 | 0.553 |
| FedGD | 0.563 | 0.356 | 0.356 | 0.543 | 0.523 | 0.432 | 0.543 | 0.454 |
| FedAvg | 0.443 | 0.453 | 0.498 | 0.453 | 0.434 | 0.523 | 0.465 | 0.431 |
| FedRMSprop | 0.346 | 0.325 | 0.489 | 0.465 | 0.343 | 0.443 | 0.343 | 0.345 |



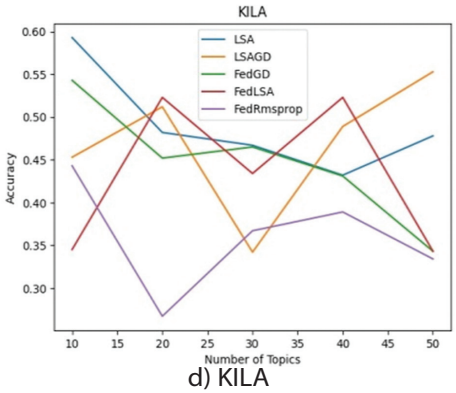

a) Digi Locker NAD




b) IHMCL




c) Kerala Startup Mission



d) KILA

**Fig. 5.**The Performance of Various Federated Latent Semantic Analysis Techniques

Comparison of Overall Performance to that of Related Articles Using locally stored documents, this research presents FedLSA, a framework for federated topic modelling algorithms based on LSA that generate high-quality topics. To mitigate the impact of client-side data heterogeneity on performance. Decentralized short text analysis and short document content mining are just two of the many potential uses for our FedLSA algorithms in light of the rising tide of privacy concerns.

## 6. CONCLUSION

Fed LSA is a framework that is introduced in this article to support federated topic modelling approaches that are based on LSA. Whether the documents are stored locally or not, these approaches could still produce high-quality topics. They provide the FedLSA design to fix performance problems brought on by data heterogeneity on the client side. Semantic analysis further optimizes the relationship between topic weights and the amount of input text. This elucidates the possible benefits of LSA for subject modelling. In light of the growing number of privacy concerns, our FedLSA algorithms have numerous potential uses, one of which is the distributed analysis of People Query documents.

## 7. REFERENCES

[1] G. Masson, N. Sneddon, R. Alghamdi, K. Alfalqi, "A survey of topic modelling in text mining", International Journal of Advanced Computer Science and Applications, Vol. 6, No. 1, 2015.

[2] D. M. Blei, A. Y. Ng, M. I. Jordan, "Latent Drichlet allocation", Journal of Machine Learning Research, Vol. 3, 2003, pp. 993-1022.

[3] A. Fallah, A. Mokhtari, A. Ozdaglar, "Personalized federated learning: A meta-learning approach", arXiv:2002.07948, 2020.

[4] J. Xu, B. S. Glicksberg, C. Su, P. Walker, J. Bian, F. Wang, "Federated learning for healthcare informatics", Journal of Healthcare Informatics Research, Vol. 5, No. 1, 2021, pp. 1-19.

[5] D. Newman, J. H. Lau, K. Grieser, T. Baldwin, "Automatic evaluation of topic coherence", Proceedings of Human language technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics, Los Angeles, CA, USA, 2-4 June 2010, pp. 100-108

[6] C. C. Aggarwal, C. Zhai, "A survey of text classification algorithms", Mining text data, Springer 2012, pp. 163-222.

[7] B. McMahan, E. Moore, D. Ramage, S. Hampson, B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data", Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, Fort Lauderdale, Fl, USA, 2017, pp. 1273- 1282.

[8] D. Jiang, Y. Song, Y. Tong, X. Wu, W. Zhao, Q. Xu, Q. Yang, "Federated topic modeling", Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3-7 November 2019, pp. 1071-1080.

[9] Y. Wang, Y. Tong, D. Shi, "Federated latent drichlet allocation: A local differential privacy based framework", Proceedings of the AAAI Conference on Artificial Intelligence, Vol. 34, No. 4, 2020, pp. 6283-6290.

[10] S. Si, J. Wang, R. Zhang, Q. Su, J. Xiao, "Federated Non-negative Matrix Factorization for Short Texts Topic Modeling with Mutual Information", Proceedings of the International Joint Conference on Neural Networks, Padua, Italy, 18-23 July 2002.

[11] Y. Li, T. Yang, "Word embedding for understanding natural language: A survey," Guide to Big Data Applications, Springer, 2018, pp. 83-104.

[12] S. Deerwester, S. T. Dumais, G. W. Furnas, T. K. Landauer, R. Harshman, "Indexing by latent semantic analysis," Journal of the American Society for Information Science, Vol. 41, No. 6, 1990, pp. 391-407.

[13] T. Williams, J. Betak, "A Comparison of LSA and LDA for the Analysis of Railroad Accident Text", Procedia Computer Science, Vol. 130, 2018, pp. 98-102.

[14] J. Stremmel, A. Singh, "Pretraining federated text models for next word prediction", arXiv:2005.04828, 2020.

[15] K. Cheng, T. Fan, Y. Jin, Y. Liu, T. Chen, Q. Yang, "Secure boost: A lossless federated learning framework", arXiv:1901.08755, 2019.

[16] A. Fang, C. Macdonald, I. Ounis, P. Habel, "Using word embedding to evaluate the coherence of topics from twitter data", Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval, Pisa, Italy, 17-21 July 2016, pp. 1057-1060.

[17] R. Barzilay, M. Lapata, "Modeling local coherence: An entitybased approach", Computational Linguistics, Vol. 34, No. 1, 2008, pp. 1-34.

[18] S. Zhou, X. Xu, Y. Liu, R. Chang, Y. Xiao, "Text Similarity Measurement of Semantic Cognition Based on Word Vector Distance Decentralization With Clustering Analysis", IEEE Access, Vol. 7, 2019, pp. 107247-107258.

[19] J. Hoblos, "Experimenting with Latent Semantic Analysis and Latent Drichlet Allocation on Automated Essay Grading", Proceedings of the Seventh International Conference on Social Networks Analysis, Management and Security, Paris, France, 14-16 December 2020.

[20] T.-M. H. Hsu, H. Qi, M. Brown, "Measuring the effects of nonidentical data distribution for federated visual classification", arXiv:1909.06335, 2019.

[21] R. Alghamdi, K. Alfalqi "A Survey of Topic Modeling in Text Mining", International Journal of Advanced Computer Science and Applications, Vol. 6, No. 1, 2015.

# Real-Time Solid Waste Sorting Machine Based on Deep Learning

Original Scientific Paper

**Imane Nedjar***
University of Tlemcen,
Biomedical Engineering Laboratory
Ecole Supérieure en Sciences Appliquées de Tlemcen, ESSA-Tlemcen,
BP 165 RP Bel Horizon, Tlemcen 13000, Algeria
imane.nedjar@univ-tlemcen.dz

**Mohammed M'hamedi**
University of Tlemcen,
Faculty of Sciences, Department of Computer Science
Ecole Supérieure en Sciences Appliquées de Tlemcen, ESSA-Tlemcen,
BP 165 RP Bel Horizon, Tlemcen 13000, Algeria
mohamed.mhamedi@univ-tlemcen.dz

**Mokhtaria Bekkaoui**
University of Tlemcen,
Manufacturing Engineering Laboratory of Tlemcen
Ecole Supérieure en Sciences Appliquées de Tlemcen, ESSA-Tlemcen,
BP 165 RP Bel Horizon, Tlemcen 13000, Algeria
mokhtaria.bekkaoui@univ-tlemcen.dz

*Corresponding author

***Abstract*** *– The collection and separation of solid waste represent crucial stages in recycling. However, waste collection currently relies on static trash bins that lack customization to suit specific locations. By integrating artificial intelligence into trash bins, we can enhance their functionality. This study proposes the implementation of a sorting machine as an intelligent alternative to traditional trash bins. This machine autonomously segregates waste without human intervention, utilizing deep learning techniques and an embedded edge device for real-time sorting. Deploying a convolutional neural network model on a Raspberry Pi, the machine achieves solid waste identification and segregation via image recognition. Performance evaluation conducted on both the Stanford dataset and a dataset we created showcases the machine's high accuracy in detection and classification. Moreover, the proposed machine stands out for its simplicity and cost-effectiveness in implementation.*

***Keywords***: *Waste, deep leaning, raspberry pi, artificial intelligence, sorting machine*

*Received: March 15, 2024; Received in revised form: June 12, 2024; Accepted: June 12, 2024*

## 1. INTRODUCTION

The surge in population and industrialization has led to a significant increase in daily waste production. According to statistics from the World Bank, global municipal solid waste generation exceeds 2 billion tons annually, a figure expected to soar to 3.4 billion tons by 2050 [1].

Waste in gaseous, liquid, or solid form poses a significant environmental threat if not properly managed and segregated. The Senior Director of the Social, Urban, Rural, and Resilience Global Practice at the World Bank underscores the urgency of implementing effective solid waste management practices to achieve sustainable development goals, which prioritize waste reduction and recycling [2].

Recycling is crucial for ecological preservation and promoting circular economies [3]. However, its benefits are limited because only 13.5% of global waste is recycled, primarily due to insufficient collection and sorting infrastructure. Furthermore, 33% of waste is openly discarded without preliminary sorting, while mixed waste disposal remains prevalent [4].

While manual sorting persists in some waste management systems, it introduces various challenges. These include the risk of contamination from bacteria and viruses, the requirement for a large workforce, and the associated expenses of training and oversight [5].

Developed nations are steadily adopting automated waste management systems to facilitate the advancement of intelligent and sustainable cities. These systems harness cutting-edge technologies like robotics [6], artificial intelligence [7], and the Internet of Things [8], sparking considerable research interest in this field.

In robotics, Aitken et al. [9] have devised an automated system for nuclear waste treatment utilizing a robotic arm. This system aims to execute tasks that are repetitive and hazardous for humans. On the other hand, Gupta et al. [10] have presented a cost-efficient solution for defining routes for the mobile autonomous robot assigned to litter emptying. This proposal aims to alleviate the impact on workers' health, reduce greenhouse gas emissions, and minimize operational expenses.

Given the success of artificial intelligence in various fields like medicine [11], biology [12], and environment [13], recently several researchers have conducted studies on using AI for automatic waste management. For example, Majchrowska et al. [4] used deep learning to detect waste in natural and urban environments. Furthermore, the studies presented by [14-18] have proposed a deep-learning model for waste classification. Mittal et al. [19] created a smartphone application called Spotgarbage, which utilizes a Convolutional Neural Network (CNN) and enables citizens to monitor the cleanliness of their neighborhood by tracking garbage. In contrast, The dangers posed by medical waste, such as viruses and bacteria, motivated Zhou et al. [5] to propose a classification method based on deep learning for medical waste classification. This method identifies eight types of medical waste: gauze, gloves, infusion bags, bottles, infusion apparatus, syringes, needles, and tweezers. Machine learning methods are employed in municipal solid waste classification [20] and are also utilized in container management through sensor measurements [21, 22]. Yang and Thung [23] combined machine learning and deep learning techniques by using the support vector machine and convolutional neural network to classify waste into six classes: glass, paper, metal, plastic, cardboard, and trash. On the other hand, the Internet of Things (IoT) technology has been employed with both machine learning in [24] and deep learning for waste management systems in [25].

A growing trend in automatic waste management involves the implementation of Smart Waste Bins [26]. Initially, researchers suggested improving waste bin control through level detection sensors [27, 28] and utilizing remote control via mobile applications or Global System for Mobile Communications (GSM) technology. For example, Monika et al. [29] employed an intelligent bin equipped with an ultrasonic sensor to monitor the saturation of the dustbin. The authors employed the GSM technology to alert the authorities to manage the dustbin. Additionally, convolutional neural networks have been integrated into waste bins for efficient trash segregation [30, 31].

Despite ongoing research, the current waste collection and sorting involves using trash bins equipped with labels or colors to help individuals correctly dispose of waste into designated containers. However, variations in these labels across countries can confuse some individuals.

Moreover, citizens frequently make errors in their waste disposal practices. Another limitation of existing trash bins is their lack of customization based on specific deployment locations. For instance, waste generated in hospitals differs significantly from that produced in stadiums, public gardens, educational institutions, and other settings.

In this research, our focus is on educational institutions, where we propose the implementation of a sorting machine as an alternative to conventional trash bins. We designed this machine to autonomously segregate waste without requiring human intervention. Since plastic bottles and paper are the primary waste generated in educational institutions and universities, efficiently collecting and sorting these items has become crucial for streamlining the recycling process.

Plastic waste constitutes a significant threat to the environment. The United Nations Environment Programme (UNEP) estimated that there could be more plastic than fish in the ocean by 2050 [32]. Plastic waste not only affects our oceans but also infiltrates our food supply as microplastics and nanoplastics, posing a significant threat to human health. Paper is also a part of our daily waste and impacts the environment. Discarding paper due to printing errors, packaging, and advertising posters is a daily practice that we engage in without fully recognizing its ecological and economic impact.

The primary contributions of this article include:

- Introducing the structure of the sorting machine along with its operational diagram.
- Evaluating two mobile architectures — MobileNet and NASNet-Mobile — as the backbone for our final model.
- Sharing the paper and plastic waste images dataset, and the Python source code of the machine.

The manuscript is structured as follows: The Materials section presents an overview of the machine's components and functionality. In the Method section, we will elaborate on the convolutional neural network model utilized in our proposal. We will present the evaluation results in the Results and Discussion section. Section 5 presents the conclusion of the article.

## 2. MATERIALS

### 2.1. GENERAL OVERVIEW OF THE PROTOTYPE MACHINE

The machine aims to identify and categorize paper and plastic bottles waste. By doing so, the machine can aid in simplifying the waste segregation process for municipal corporations. We chose these two types of waste due to their high frequency of being discarded in educational institutions and their recyclable nature.

The prototype machine works as follows:

As a student approaches the waste disposal machine, a motion sensor detects their presence, and the machine's window opens to allow them to dispose of their waste. After the disposition of the waste, a Raspberry Pi camera captures an image, and a convolutional neural network (CNN) model identifies and classifies the waste type. The servomotor then swings the support that holds the waste towards the appropriate container. The machine contains two compartments: one for paper waste and the other for plastic bottles.

To further encourage students to recycle waste, we designed the prototype machine to provide a reward after throwing a fixed number of plastic bottles, for example, nb=4. A box containing pens opens, allowing them to claim a pen as a reward.

This approach creates an educational and interactive way to promote waste segregation and recycling, encouraging students to engage in sustainable practices. Fig. 1 presents the UML state machine diagram that illustrates the various states and the responses to different events. We create this diagram using Astah software.

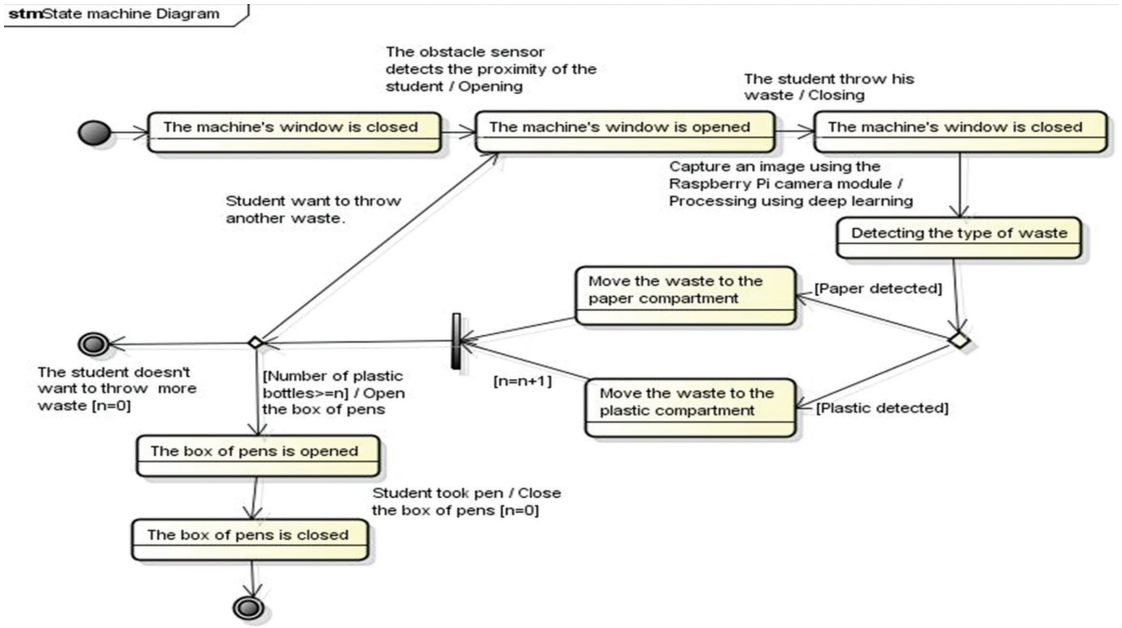



**Fig. 1.** The state machine diagram

### 2.2. MACHINE COMPOSITION

The machine has two main components: the electronic component box and two physical parts.

The top part of the machine handles the processing tasks, including the detection, recognition, and classification of waste.

The bottom part houses the containers designated for paper and plastic waste. The design of the prototype machine shown in Fig. 2 was modeled using SolidWorks software.

- Electronic component box

The electronic components used in the prototype machine, including the Raspberry Pi, Pi Camera, and servomotors, are housed in the electronic component box for protection. Fig. 3 shows the electronic circuit of the prototype machine.

The source code for the prototype machine is accessible at the following link: https://github.com/Nedjar-Imane/Sorting-Machine/tree/main

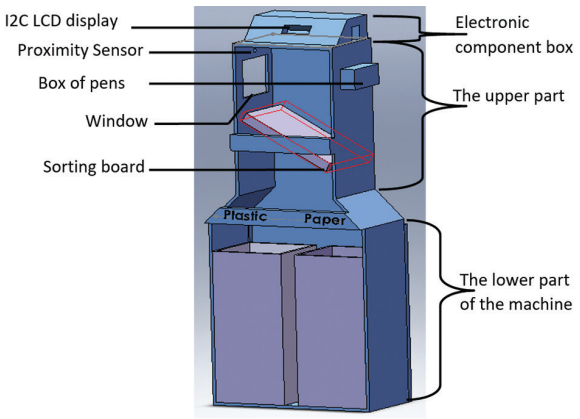



**Fig. 2.** Machine's design

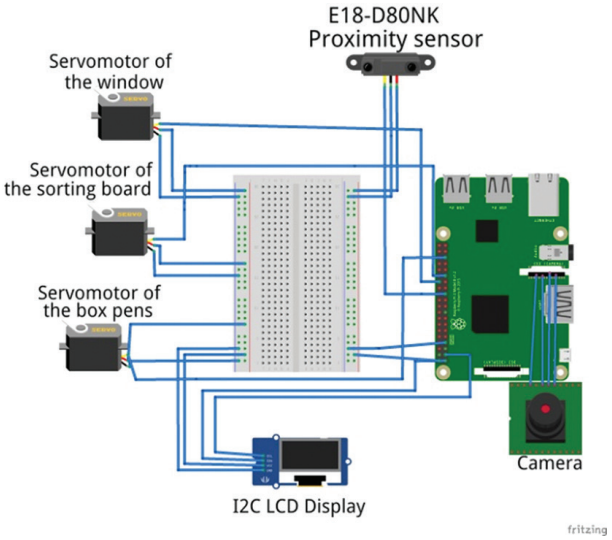



**Fig. 3.** Electronic circuit of the prototype machine created using FreeCAD software

- The upper part of the prototype machine

Window: the upper part of the prototype machine features an opening where students can deposit the waste. An obstacle avoidance sensor (E18-D80NK) is placed at the top of the window to detect incoming objects. The servomotor (Metal gears RG996R) operates the window to ensure smooth and efficient operation.

Sorting board: once the sensor detects an object, the window opens automatically, and the user can throw their waste into the machine. There is a sorting board inside the prototype machine. It turns to the right for plastic waste and the left for paper waste.

Box of pens: we developed this prototype machine for educational institutions to encourage students to recycle waste. When a student throws a specified number of plastic bottles, the box of pens opens, allowing the student to take a pen. The servomotor operates the opening mechanism of the box.

- The lower part of the prototype machine

This part is designed to sort the waste into two containers.

## 3. PROPOSED METHOD

### 3.1. DATASET

We have collected and organized our dataset titled 'Plastic and Paper Waste' by taking photos with mobile phones in our homes and at the university (see Fig. 4).

The 'Plastic and Paper Waste' dataset contains 400 images for each class, encompassing diverse papers and plastic bottles captured in various positions, states, lighting conditions, and backgrounds.

In our experiment, we also used the Stanford dataset [23], which includes images of trash against a white background organized into six classes.

The Stanford dataset contains 594 images of paper and 319 images of plastic bottles. To balance the dataset, we augmented the number of plastic bottle images to 594 using techniques such as rotation and zoom.

(a) (b)

**Fig. 4.** Images from our dataset with different backgrounds (**a**) Paper, (**b**) Plastic bottle

### 3.2. CONVOLUTIONAL NEURAL NETWORK MODEL

The remarkable achievements of CNN-based architectures are notably outstanding, particularly in computer vision, where accuracy levels often approach perfection.

Our system is specifically tailored for real-time detection and classification of paper and plastic, utilizing CNN for these tasks. In this study, we opted for the MobileNet and NASNet-Mobile neural network architectures for classification, chosen for their suitability in the context of mobile and embedded devices. Several recent studies have used these neural network architectures, such as [33, 34], in addition to real-time video applications where processing speed is crucial [35, 36].

- MobileNet

Google's MobileNets architecture [37] is tailored for mobile and embedded vision applications due to its lightweight design. Its efficiency stems from the depthwise separable convolutions instead of full convolutions. MobileNet introduces two parameters: Width Multiplier (α) and Resolution Multiplier (ρ), which enhance the architecture's flexibility.

- NASNet-Mobile

Neural Architecture Search Network aims to discover an optimal CNN architecture using reinforcement learning. NASNet is a technique developed at Google Brain for searching through a space of neural network configurations [38]. The optimized version, based on Normal and Reduction-Cells, is known as NASNet-Mobile. Normal Cells are convolutional cells that return a feature map of the same dimension as the input, while Reduction Cells are convolutional cells that reduce the feature map's height and width by a factor of two. These cells are combined to create the complete neural network optimized for a specific task while minimizing the computational resources needed for training and inference.

### 3.3. RASPBERRY PI AND TENSORFLOW LITE

- Raspberry Pi

Is is a single-board computer developed by the Raspberry Pi Foundation. The Raspberry Pi boards are about the size of a credit card and feature a range of input/output pins that connect to sensors, motors, and other electronic components.

- TensorFlow Lite

It is a lightweight framework for deploying deep learning models on mobile and embedded devices. It is an optimized version of the popular TensorFlow library. With TensorFlow Lite, models can run locally on the device without relying on cloud-based services, allowing for real-time processing and lower latency.

In this work, we installed the proposed system on a Raspberry Pi to ensure real-time detection and classification. Additionally, we have used the library TensorFlow Lite as background for the CNN model.

### 3.4. EVALUATION METRICS

The measures considered for evaluating the CNN models rely on various metrics, including accuracy, precision, recall, F1 score, and kappa statistic.

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

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

$$\mathrm{Re}\ call = \frac{TP}{TP+FN} \quad (3)$$

$$F1\,score = 2\times\frac{\Pr ecision \times \mathrm{Re}\ call}{\Pr ecision + \mathrm{Re}\ call} \quad (4)$$

where: *TP* refers to True Positive, and *TN* refers to True Negative, which indicates the correct classification of plastic bottles and paper images.

On the other hand, *FP* refers to False Positive, and *FN* refers to False Negative, which indicates the misclassification of plastic bottles and paper images.

$$\mathrm{Kappa} = \frac{(P_0 - P_e)}{(1-P_e)} \quad (5)$$

The Kappa Statistic [39] is calculated from the variance between the observed agreement (*Po*) and the expected agreement (*Pe*), highlighting the difference between the actual agreement and what would be expected by chance alone.

## 4. RESULTS AND DISCUSSION

We conducted the evaluation using both our dataset and the Stanford dataset. During the training process, we applied transfer learning, whereby we initialized the CNNs with pre-trained ImageNet weights.

We also employed a fine-tuning strategy to improve the prediction by adding extension layers to the CNNs. These extensions consist of a Global Average Pooling layer, a Dense layer, and a Dropout layer.

The stochastic gradient descent optimizer [40] was used, with a momentum equal to 0.9, a learning rate of 1e-4, and 20 epochs for training. We used the cross-entropy loss function $L$, as shown in equation (6), which increases when the predicted probability diverges from the correct label.

$$L = -(y\log(p) + (1-y)\log(1-p)) \quad (6)$$

where $y$ is a binary indicator, with values of 0 or 1, denoting whether the class label '$c$' accurately identifies observation '$o$'. Similarly, $p$ represents the predicted probability of observation '$o$' belonging to the '$c$' class.

The experiment consisted of testing each dataset individually, followed by combining them.We split the datasets into a training set (80%) and a validation set (20%). Fig. 5 shows that both models achieved accuracy levels above 96% and 98%.

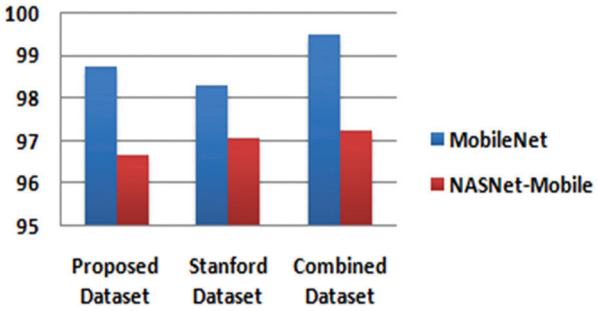



**Fig. 5.** The accuracy obtained for each dataset

MobileNet outperformed NASNet-Mobile in all datasets. We combined the Stanford and the proposed dataset to address the overfitting issue in classification.

MobileNet achieved the highest score of 99.50% with an error rate of 0.0136 on the combined dataset (see Fig. 6 and Fig. 7 ). In Table 1, we have compared the metric values of kappa, precision, recall, and F1 score for MobileNet and NASNet-Mobile. The results obtained showed that MobileNet outperforms NASNet-Mobile for all the metrics used. Based on the results of the experiments, MobileNet has been chosen as the base model for our machine (see Fig. 8 ).

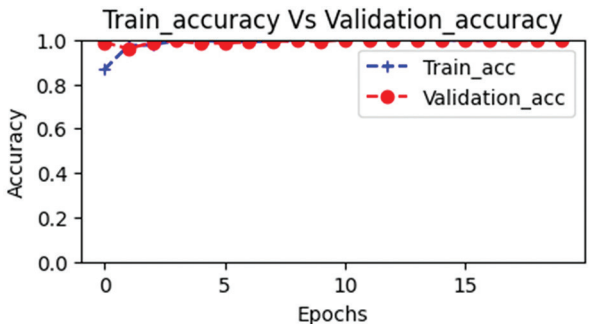



**Fig. 6.** The accuracy of MobileNet on the combined dataset

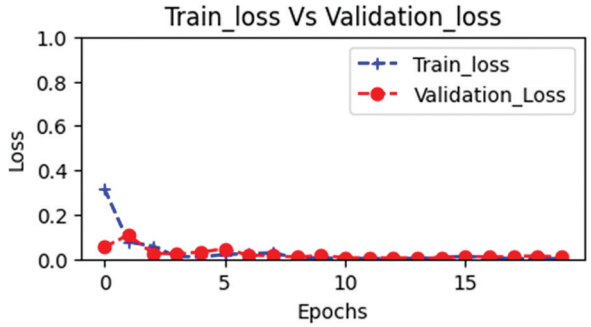



**Fig. 7.** The loss of MobileNet on the combined dataset

**Table 1.** Kappa, Precision, Recall and F1 score obtained for combined dataset

| | Kappa | Precision | | Recall | | F1 score | |
|---|---|---|---|---|---|---|---|
| | | Paper | Plastic | Paper | Plastic | Paper | Plastic |
| MobileNet | 0.98 | 0.99 | 1.00 | 1.00 | 0.99 | 0.99 | 0.99 |
| NASNet-Mobile | 0.94 | 0.96 | 0.98 | 0.98 | 0.96 | 0.97 | 0.97 |

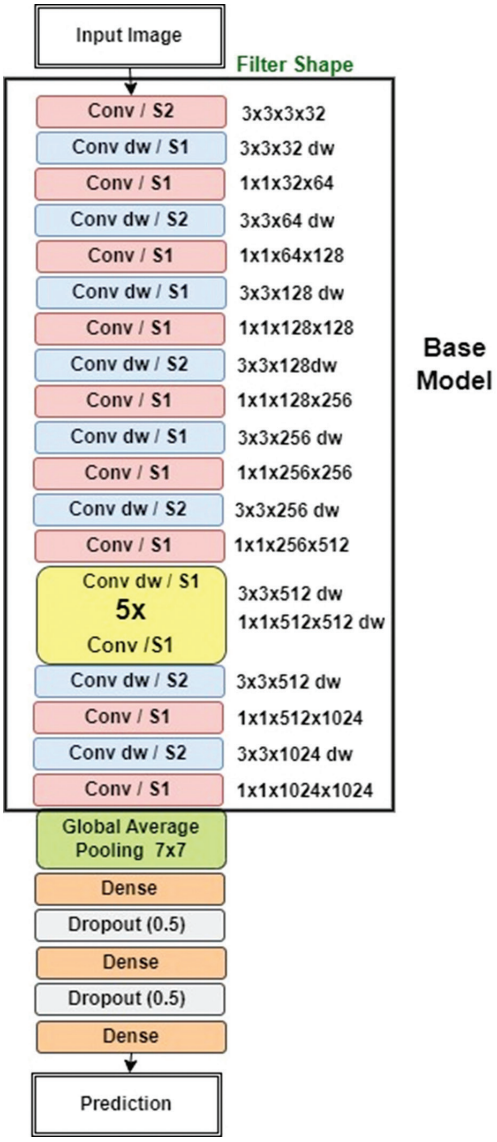



**Fig. 8.** The architecture of our model based on MobileNet

To enhance the convergence and generalization performance of the model, we applied Cyclical Learning Rates (CLR) to our model.

The concept behind CLR is to discover an optimal learning rate schedule by systematically varying the learning rate throughout the training process. Instead of using a fixed learning rate in training neural networks, CLR dynamically adjusts the learning rate cyclically, oscillating between a minimum and maximum value over a predefined number of iterations [41].

We have chosen a minimum learning rate of 1e-4 and a maximum learning rate of 1e-1.

We present the CLR schedule obtained in Fig. 9.

Applying CLR to our model has improved the accuracy obtained while accelerating the training process. Fig. 10 shows that all the metrics have improved, reaching 100%.

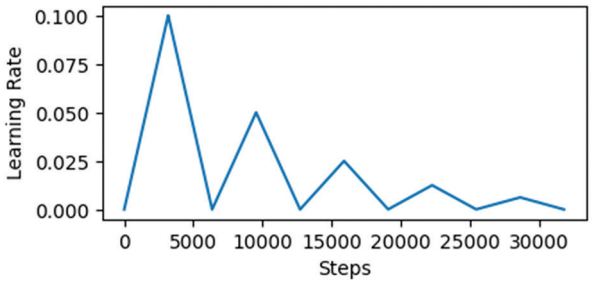



**Fig. 9.** Cyclical Learning Rates schedule

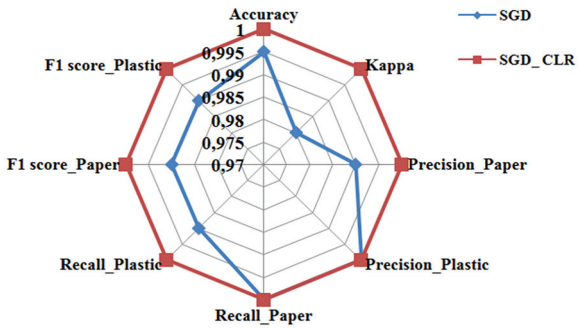



**Fig. 10.** The metrics obtained using the SGD optimizer with and without CLR on our model for the combined dataset

From 398 test images, which included 199 images of plastic bottles and 199 paper images, there were only two misclassified images by our model without CLR.

We obtained a paper precision of 99%, plastic bottle precision of 100%, paper recall of 100%, plastic bottle recall of 99%, Kappa of 98%, and F1 score of 99%. In addition, When we used the proposed model with CLR, the two misclassified images were correctly classified, resulting in an improvement in the values of all metrics, reaching 100%.

Since our model runs on a Raspberry Pi, balancing performance and computation time was crucial. To achieve this, we opted for TensorFlow Lite instead of TensorFlow.

In real-time, the Raspberry Pi camera captures an image of the waste when the student deposits it into the prototype machine. The proposed model then classifies the waste based on the image, and finally, the prototype machine directs the waste to the appropriate container.

Several researchers have focused on developing models for waste recognition, including medical waste [5], construction and demolition waste [42], and municipal solid waste [14-18].

In this work, we propose both the model for recognition and classification and a prototype machine to make the idea more practical and feasible.

### 4.1. FUTURE PERSPECTIVE

Our idea is to enhance the existing trash bins with intelligent machines capable of detecting, recognizing, and sorting waste.

The proposed machine has been designed initially for educational institutions, but its application is not limited to them; it can also adapted for use in public spaces and even in houses. For this purpose, certain modifications need to be made to the machine's system and mechanism.

The machine system must identify and classify other types of waste, such as glass, metal, organic waste, food scraps, and non-recyclable waste. The machine can be customized for the place where it will be used. For example, if we want to use the machine in healthcare facilities, biomedical waste generated, such as needles, syringes, and other medical equipment, must be included.

In our proposed prototype machine the sorting board moves the waste to the appropriate container. In cases where there are more than two types of waste, the machine needs to have a rotation mechanism that allows the containers to rotate, allowing the sorting board to direct the waste into the correct container.

## 5. CONCLUSION

Intelligent waste management is considered a viable solution for achieving sustainable development goals. In this study, we introduce the design of a real-time sorting prototype machine that leverages artificial intelligence for effective solid waste collection and separation.

The machine comprises physical components for waste sorting and a software component for identification and classification. The physical components incorporate an object detection sensor (E18-D80NK), servomotors (Metal Gears RG996R) for movement, and a Raspberry Pi for real-time detection and classification. To identify and classify waste, we tested two baseline models, namely MobileNet and NASNet-Mobile, on both the Stanford dataset and our proposed dataset. The final model chosen was based on MobileNet, achieving an accuracy of 99.50% without employing a cyclical learning rate and 100% when we used it.

To pave the way for potential improvements and industrial realization, we have made the machine's source code readily accessible.

Data Availability Statement

The dataset titled 'Plastic and Paper Waste' is available on GitHub at: https://github.com/Nedjar-Imane/Sorting-Machine/tree/main/Datasets

## 6. REFERENCES


[1] S. Kaza, L. Yao, P. Bhada-Tata, F. Van Woerden, "What a waste 2.0: a global snapshot of solid waste management to 2050", https://openknowledge.worldbank.org/handle/10986/2174 (accessed: 2024)

[2] WorldBank.org, "What a Waste: An Updated Look into the Future of Solid Waste Management", https://www.worldbank.org/en/news/immersive-story/2018/09/20/what-a-waste-an-updated-look-into-the-future-of-solid-waste-management (accessed: 2024)

[3] G. Rozing, D. Jukić, H. Glavaš, M. Žnidarec, "Recyclability and ecological-economic analysis of a simple photovoltaic panel", International Journal of Electrical and Computer Engineering Systems, Vol. 15, No. 1, 2022, pp. 99-104.

[4] S. Majchrowska, A. Mikołajczyk, M. Ferlin, Z. Klawikowska, M. A. Plantykow, A. Kwasigroch, K. Majek, "Deep learning-based waste detection in natural and urban environments", Waste Management, Vol. 138, 2022, pp. 274-284.

[5] H. Zhou et al. "A deep learning approach for medical waste classification," Scientific Reports, Vol. 12, No. 1, 2022, p. 2159.

[6] A. G. Satav, S. Kubade, C. Amrutkar, G. Arya, A. Pawar, "A state-of-the-art review on robotics in waste sorting: scope and challenges", International Journal on Interactive Design and Manufacturing, Vol. 17, No. 6, 2023, pp. 2789-2806.

[7] B. Fang et al. "Artificial intelligence for waste management in smart cities: a review", Environmental Chemistry Letters, Vol. 21, No. 4, 2023, pp. 1959-1989.

[8] T. Anagnostopoulos, A. Zaslavsky, K. Kolomvatsos, A. Medvedev, P. Amirian, J. Morley, S. Hadjieftymiades, "Challenges and opportunities of waste management in IoT-enabled smart cities: a sur-

vey", IEEE Transactions on Sustainable Computing, Vol. 2, No. 3, 2017, pp. 275-289.

[9] J. M. Aitken et al. "Autonomous nuclear waste management", IEEE Intelligent Systems, Vol. 33, No. 6, 2018, pp. 47-55.

[10] A. Gupta, M. J. Van der Schoor, J. Bräutigam, V. B. Justo, T. F. Umland, D. Goehlich, "Autonomous service robots for urban waste management-multiagent route planning and cooperative operation", IEEE Robotics and Automation Letters, Vol. 7, No 4, 2022, pp. 8972-8979.

[11] T. Schaffter et al. "Evaluation of combined artificial intelligence and radiologist assessment to interpret screening mammograms", JAMA Network Open, Vol. 3, No. 3, 2020, pp. e200265-e200265.

[12] B. Richards, D. Tsao, A. Zador, "The application of artificial intelligence to biology and neuroscience", Cell, Vol. 185, No. 15, 2022, pp. 2640-2643.

[13] V. Hocenski, A. Lončarić Božić, N. Perić, D. Klapan, Ž. Hocenski, "Environmental impact estimation of ceramic tile industry using modeling with neural networks", International Journal of Electrical and Computer Engineering Systems, Vol. 13, No. 1, 2022, pp. 29-39.

[14] C. Shi, C. Tan, T. Wang, L. Wang, "A waste classification method based on a multilayer hybrid convolution neural network", Applied Sciences, Vol. 11, No. 18, 2021, p. 8572.

[15] M. Al Duhayyim, T. A. Elfadil Eisa, F. N. Al-Wesabi, A. Abdelmaboud, M. A. Hamza, A. S. Zamani, M. Rizwanullah, R. Marzouk, "Deep reinforcement learning enabled smart city recycling waste object classification", Computers, Materials & Continua, Vol. 71, No. 3, 2022, pp. 5699-5715.

[16] A. Mitra, "Detection of waste materials using deep learning and image processing", University San Marcos, California State, USA, Master Thesis, 2020.

[17] M. Malik, S. Sharma, M. Uddin, C. L. Chen, C. M. Wu, P. Soni, S. Chaudhary, "Waste classification for sustainable development using image recognition with deep learning neural network models", Sustainability, Vol. 14, No. 12, 2022, p. 7222.

[18] A. Altikat, A. Gulbe, S. Altikat, "Intelligent solid waste classification using deep convolutional neural networks", International Journal of Environmental Science and Technology, Vol. 19, 2022, pp. 1285-1292.

[19] G. Mittal, K. B. Yagnik, M. Garg, N. C. Krishnan, "Spotgarbage: smartphone app to detect garbage using deep learning", Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, Association for Computing Machinery, Heidelberg, Germany, 12-16 September 2016, pp 940-945.

[20] S. Chaturvedi, B. P. Yadav, N. A. Siddiqui, "An assessment of machine learning integrated autonomous waste detection and sorting of municipal solid waste", Nature Environment & Pollution Technology, Vol. 20, No. 4, 2021, pp. 1515-1525.

[21] D. Rutqvist, D. Kleyko, F. Blomstedt, "An automated machine learning approach for smart waste management systems", IEEE Transactions on Industrial Informatics, Vol. 16, No. 1, 2020, pp. 384-392.

[22] S. Dubey, P. Singh, P. Yadav, K. K. Singh, "Household waste management system using IoT and machine learning", Procedia Computer Science, Vol. 167, 2020, pp. 1950-1959.

[23] M. Yang, G. Thung, "Classification of trash for recyclability status", Stanford University, Stanford, CA, USA, CS229 project report, 2016.

[24] R. Khan, S. Kumar, A. K. Srivastava, N. Dhingra, M. Gupta, N. Bhati, P. Kumari, "Machine learning and IoT-based waste management model", Computational Intelligence and Neuroscience, Vol. 2021, No. 1, 2021, p. 5942574.

[25] M. W. Rahman, R. Islam, A. Hasan, N. I. Bithi, M. M. Hasan, M. M. Rahman, "Intelligent waste management system using deep learning with IoT", Journal of King Saud University-Computer and Information Sciences, Vol. 34, No.5, 2022, pp. 2072-2087.

[26] A. Noiki, S. A. Afolalu, A. A. Abioye, C. A. Bolu, M. E. Emetere, "Smart waste bin system: a review", Proceedings of the 4th International Conference on Science and Sustainable Development, Ota, Nigeria, 3-5 August 2020, p. 012036.

[27] Y. Zhao, S. Yao, S. Li, S. Hu, H. Shao, T. F. Abdelzaher, "VibeBin: A vibration-based waste bin level detec-

tion system", Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 1, No 3, 2017, pp. 1-22.

[28] S. J. Ramson, D. J. Moni, S. Vishnu, T. Anagnostopoulos, A. A. Kirubaraj, X. Fan, "An IoT-based bin level monitoring system for solid waste management", Journal of Material Cycles and Waste Management, Vol. 23, 2021, pp. 516-525.

[29] K. Monika et al. "Smart dustbin-an efficient garbage monitoring system", International Journal of Engineering Science and Computing, Vol. 6, No. 6, 2016, pp. 7113-7116.

[30] S. Hulyalkar, R. Deshpande, K. Makode, S. Kajale, "Implementation of smartbin using convolutional neural networks", International Research Journal of Engineering and Technology, Vol. 5, No. 4, 2018, pp. 1-7.

[31] I. E. Agbehadji, A. Abayomi, K. H. N. Bui, R. C. Millham, E. Freeman, "Nature-Inspired Search Method and Custom Waste Object Detection and Classification Model for Smart Waste Bin", Sensors, Vol. 22, No. 16, 2022, p. 6176.

[32] UNEP.org, "UN Declares War on Ocean Plastic", https://www.unep.org/news-and-stories/press-release/un-declares-war-ocean-plastic-0 (accessed: 2024)

[33] X. Li, J. Du, J. Yang, S. Li, "When mobilenetv2 meets transformer: A balanced sheep face recognition model", Agriculture, Vol. 12, No. 8, 2022, p. 1126.

[34] D. Fangchun, L. Jingbing, A. B. Uzair, L. Jing, C. Yen-Wei, L. Dekai, "Robust Zero Watermarking Algorithm for Medical Images Based on Improved NasNet-Mobile and DCT", Electronics, Vol. 12, No. 16, 2023, p. 3444.

[35] I. Nedjar, H. M. Sekkil, M. Mebrouki, M. Bekkaoui, "A comparison of convolutional neural network models for driver fatigue detection", Proceedings of the 7th International Conference on Image and Signal Processing and their Applications, Mostaganem, Algeria, 8-9 May 2022, pp. 1-6.

[36] S. Manzoor, E. J. Kim, S. H. Joo, S. H. Bae, G. G. In, K. J. Joo, J. H. Choi, T. Y. Kuc, "Edge deployment framework of guardbot for optimized face mask recognition with real-time inference using deep learning", IEEE Access, Vol. 10, 2022, pp. 77898-77921.

[37] A. G. Howard et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications", arXiv:1704.04861, 2017.

[38] B. Zoph, V. Vasudevan, J. Shlens, Q. V. Quoc V. Le, "Learning transferable architectures for scalable image recognition", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18-23 June 2018, pp. 8697-8710.

[39] A. J. Viera, J. M. Garrett, "Understanding interobserver agreement: the kappa statistic", Family Medicine, Vol. 37, No. 5, 2005, pp. 360-363.

[40] M. Zinkevich, M. Weimer, L. Li, A. Smola, "Parallelized stochastic gradient descent", Advances in Neural Information Processing Systems, Vol. 23, 2010.

[41] L. N. Smith, "Cyclical learning rates for training neural networks", Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Santa Rosa, CA, USA, 24-31 March 2017, pp. 464-472.

[42] K. Lin, T. Zhou, X. Gao, Z. Li, H. Duan, H. Wu, G. Lu, Y. Zhao, "Deep convolutional neural networks for construction and demolition waste classification: Vggnet structures, cyclical learning rate, and knowledge transfer", Journal of Environmental Management, Vol. 318, 2022, p. 115501.

# Improving Spatio-Temporal Topic Modeling with Swarm Intelligence: A Study on TripAdvisor Forum of Morocco

Original Scientific Paper

**Ibrahim Bouabdallaoui***
LASTIMI Laboratory – High School of Technology Salé, Mohammed V University in Rabat
Avenue Le Prince Héritier, Salé, Morocco
ibrahim_bouabdallaoui@um5.ac.ma

**Fatima Guerouate**
LASTIMI Laboratory – High School of Technology Salé, Mohammed V University in Rabat
Avenue Le Prince Héritier, Salé, Morocco
fatima.guerouate@est.um5.ac.ma

**Mohammed Sbihi**
LASTIMI Laboratory – High School of Technology Salé, Mohammed V University in Rabat
Avenue Le Prince Héritier, Salé, Morocco
mohammed.sbihi@est.um5.ac.ma

*Corresponding author

***Abstract*** *– This study introduces innovative methodologies for spatiotemporal topic modeling applied to the TripAdvisor forum of Morocco, leveraging the diverse and geographically tagged user-generated content. We develop and evaluate two schemas integrating Latent Dirichlet Allocation (LDA) with advanced clustering techniques, including a hybrid K-Means algorithm that incorporates Genetic Algorithms and the Artificial Bee Colony method. The first schema independently processes user threads, publication times, and locations using LDA, followed by clustering, while the second schema combines these dimensions into a unified vector for holistic LDA application, facilitating direct comparisons of clustering efficacy. Our findings demonstrate that swarm intelligence significantly boosts clustering performance, especially for larger clusters, and enhances the visualization of complex data relationships. These insights offer actionable intelligence for tourism stakeholders and underscore the practical benefits of advanced computational techniques in harnessing user-generated content for strategic decision-making.*

***Keywords***: *topic modeling, latent Dirichlet allocation, artificial bee colony, genetic algorithms, k-means*

*Received: March 24, 2024; Received in revised form: June 14, 2024; Accepted: June 17, 2024*

## 1. INTRODUCTION

The digital era has catalyzed an unprecedented expansion of online content, transforming forums into invaluable repositories of user-generated data. Among these, TripAdvisor stands out as a premier global travel platform, amassing a vast array of reviews, discussions, and user interactions that are rich in spatial and temporal diversity. This platform provides a unique window into user experiences, offering insights across a spectrum from travel advice to detailed service reviews [1]. The Moroccan TripAdvisor forum, in particular, encapsulates a vivid tableau of the region's cultural, economic, and touristic pulse. However, the complexity and volume of this data pose significant analytical challenges, underscoring the necessity for sophisticated analytical methodologies. In the realm of text analysis, Topic Modeling, particularly through the use of Latent Dirichlet Allocation (LDA), has proven to be a powerful tool for uncovering latent thematic patterns within large text corpora. The integration of spatial and temporal data into topic modeling further enhances our ability to perform dynamic, context-aware analyses—aptly termed Spatio-Temporal Topic Modeling. This approach finds applicability in a myriad of fields such as urban planning, epidemiology, and notably, tourism, where understanding spatial and temporal variations is crucial. Recent studies underscore the growing sophis-

tication in spatio-temporal topic modeling across various domains, affirming the relevance of our proposed methodologies for the TripAdvisor forum of Morocco. Liu et al. (2015) demonstrated the potential of spatio-temporal topic models to analyze social media check-in data, revealing user movements and interests that parallel the tourist behaviors observable in TripAdvisor reviews [2]. Similarly, Min et al. (2014) explored multimodal spatiotemporal themes in landmark studies, a concept that can be adapted to identify and analyze thematic patterns in reviews related to specific tourist landmarks [3]. Luna and Genton (2005) offered a predictive model approach for handling spatially sparse but temporally rich data, an approach that could enhance the understanding of spatial and temporal dynamics in TripAdvisor forum data [4]. Additionally, the work by Chen et al. (2019) on local topic detection using spatio-temporal social media data provides a valuable framework for extracting localized insights from geographically tagged discussions on TripAdvisor [5]. Finally, the methodology proposed by Zhao et al. (2016) for efficiently mining topics from spatio-temporal documents could directly inform our approaches to managing the complex dataset derived from the TripAdvisor forum [6]. Collectively, these studies not only validate the necessity of advanced modeling techniques but also enhance the robustness of our research design, aiming to uncover rich, actionable insights into the tourism dynamics depicted in user-generated content. This study introduces two innovative schemas for applying Spatio-Temporal Topic Modeling to the TripAdvisor forum of Morocco. These schemas aim to synergistically combine text, temporal, and spatial data using advanced methodologies including LDA, vectorization, autoencoding, and a novel hybrid K-Means clustering approach that integrates the capabilities of Genetic Algorithms and the Artificial Bee Colony method. The objective is to evaluate these schemas' effectiveness in generating discernible, meaningful topics and improving clustering performance for large, complex datasets.

## 2. LITERATURE REVIEW

This paper builds upon a rich body of work in the field of spatio-temporal Topic Modeling. It is, therefore, imperative to discuss and understand the relevant research landscape. This section presents a review of the pertinent literature, tracing the development of key concepts, identifying the primary methodologies employed, and highlighting significant findings and their implications. It also identifies gaps in the existing research, underscoring the contribution of our study to the field of spatio-temporal Topic Modeling. Researchers introduced a comprehensive framework for managing, processing, analyzing, and detecting trending topics in streaming data coming from Twitter [7]. Their utilization of a hybrid model selector and their application of deep learning and transfer learning techniques for classifying health-related tweets are noteworthy. The paper presents a methodical approach to topic detection, focused on processing data with sentence granularity, pertinent to the nature of Twitter messages. It engages a variety of techniques including Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), LDA-MALLET, and Biterm Topic Modeling to assess the effectiveness of the proposed framework. These models are used for data dimensionality reduction, clustering of documents, and in-depth analysis of short messages, respectively. Our work parallels this research, particularly in its emphasis on the spatio-temporal aspects of topic modeling. However, our approach is uniquely applied to the domain of tourism, focusing on user-generated content from TripAdvisor in the context of Morocco, rather than health-related Twitter data. Like the study above, our research also grapples with the complexities of conflicting data from different locations and times. We also share a commitment to cleaning and preprocessing data to enhance quality. In addition, both studies underline the importance of visualization to understand the topics' dynamics. A paper presents an innovative two-stage system to detect and track events from tweets [8]. By integrating Latent Dirichlet Allocation (LDA) and a density-contour-based spatio-temporal clustering approach, it manages to create a comprehensive framework for tracking events on Twitter, where events are identified as topics in tweets. The event identification process involves partitioning the geo-tagged tweet stream into temporal windows and running an LDA-based topic discovery step. Subsequently, each tweet is assigned an event label, and density-contour-based spatio-temporal clustering is employed to identify event clusters. Our study resonates with this approach as we also employ LDA for topic modeling in our Schema I. However, in Schema II, we diverge by using swarm intelligence algorithms to form a combined context content vector, hence tracking the dynamic spatiotemporal trends in a more nuanced way. The paper also introduces a novel methodology for ensuring topic continuity through calculating KL-divergences between topics, and a density-contour clustering approach for establishing spatio-temporal continuity. These methodologies bear similarities with our temporal and spatial correlation strategies in the dynamic clustering process. Their work, like ours, emphasizes the significance of spatio-temporal analysis in understanding the dynamics of topic trends, thereby contributing to the broader field of spatio-temporal topic modeling. However, their focus on tracking events in Twitter using density-contour-based clustering, while ours focuses on swarm intelligence, highlights different approaches to the similar challenge of tracking dynamic trends. A novel approach to topic modeling has been presented by authors who developed a multi-objective optimization algorithm based on the swarm intelligence of a bee colony, known as the Multi-Objective Artificial Bee Colony (MOABC) [9]. This method aims to enhance the performance of topic modeling, an area of text analysis that extracts underlying topics from document collections. Traditionally,

Latent Dirichlet Allocation (LDA) has been the most recognized method for topic modeling. LDA models each document as a probabilistic distribution over latent topics, considering a multinomial distribution for the document and the topics, each generated from a Dirichlet distribution with specific parameters. However, the authors argue that there is room for significant improvements in LDA's performance. To address this, the MOABC algorithm has been introduced. The methodology of the MOABC algorithm incorporates several steps: initializing the set of non-dominated solutions, assigning solutions to each bee, executing the main loop of the algorithm, generating modified solutions, replacing original solutions if the modified ones are better, sorting solutions by ranking and crowding, and updating the set of non-dominated solutions until a maximum number of cycles are reached. This approach essentially considers multiple objectives such as coherence, coverage, and perplexity, and each solution represents a set of topics along with their respective weights. The results from the experiments conducted on the Reuters-21578 and TagMyNews datasets indicated that the MOABC approach provides relevant improvements with respect to both LDA and the Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D) [10]. Therefore, the study provided substantial evidence supporting the exploitation of the multi-criteria nature of topic modeling with multi-objective optimization approaches, which marks an important development in the field of topic modeling. A review highlights the applicability of transformers for modeling long-range dependencies across various domains, including NLP [11]. The paper discusses how transformers, which have been successful in NLP, can be adapted for spatio-temporal modeling in different modalities, offering insights that could be applied to spatio-temporal topic modeling in textual data analysis. A work extends traditional language modeling to include spatiotemporal conditions, providing a novel approach for modeling text associated with specific times and places. It aims to capture the neighborhood, periodicity, and hierarchy within spatio-temporal text data, offering insights that are directly applicable to understanding and modeling the dynamics of user-generated content on platforms like TripAdvisor, particularly in how text is generated in response to spatial and temporal contexts. The study develops neural network models for language modeling conditioned on spatio-temporal variables [12].

## 3. METHOD

### 3.1. DATASET

The dataset used for this study is a result of web scraping from the TripAdvisor forum in Morocco. It consists of a total of 29,733 posts, spanning a significant period of more than 15 years from December 2007 to March 2023. Such a wide time frame presents an excellent opportunity for longitudinal study of trends, and the transformation of topics over time, thereby adding a temporal layer to our analysis. Each entry in the dataset includes the post content, the username of the post's author, their location, and the date of the post. The post content is used as the main body of text for topic modeling, and it provides a rich source of diverse perspectives and experiences shared by the forum users. The username of the post's author adds an element of personalization, potentially allowing for the exploration of user-specific topics or trends. The location information of the authors brings a unique perspective to our study. It allows us to understand the geographical distribution of the authors and to assess any potential influences of the authors' home locations on the topics discussed, thereby adding a spatial dimension to our analysis. The date of each post, indicating when the content was shared, is key for the temporal aspect in LDA. This information helps us understand how topics and discussions have evolved over time and could reveal temporal trends or patterns in the data. Collectively, this dataset offers a wealth of information for conducting a comprehensive spatio-temporal topic modeling study. Its size and depth make it suitable for testing and validating our proposed schemas, while its diversity ensures that the findings of our research are representative and applicable to a wide range of situations in the realm of tourism [13].

### 3.2. A MULTI-LAYERED APPLICATION OF LDA AND SWARM-ENHANCED CLUSTERING

In the first schema of our methodology (Fig. 1.), we approach the spatio-temporal topic modeling task through a sequential and layered implementation of Latent Dirichlet Allocation (LDA), followed by the independent application of two different swarm intelligence algorithms for clustering.

Stage 1: LDA Implementation:

- Textual LDA: The first stage of this schema involves applying LDA to user threads, with each thread treated as a separate document. By applying LDA, we extract the underlying topics or themes from the text content, each represented by a set of topic probabilities. This allows us to understand the main themes and topics discussed in the user threads, providing a detailed thematic analysis of the text data.
- Temporal LDA: Next, we apply LDA to the time of publication of each thread, treating each timestamp as a temporal document. This allows us to uncover temporal themes, potentially revealing patterns or trends in discussions over time. By analyzing the distribution of topics over time, we can identify how certain topics gain or lose prominence, reflecting temporal dynamics in user discussions.
- Spatial LDA: The third application of LDA is performed on the author's location data associated with each post. This step uncovers spatial themes, indicative of geographically influenced discussions

or trends. By examining how topics vary across different geographical locations, we can gain insights into region-specific interests and trends.

Each of these applications of LDA produces a topic distribution for each document (thread). These distributions are then vectorized to create a unified representation of the textual, temporal, and spatial themes for each thread. This unified representation is crucial for capturing the multifaceted nature of the data, integrating text, time, and location into a comprehensive feature set [14].

Stage 2: Autoencoder Processing

The resulting vectors from the LDA stages serve as input to an autoencoder. Autoencoders are neural networks trained to recreate their input data, thereby learning compressed, meaningful representations of the input data in their hidden layers. The autoencoder learns to encode the high-dimensional input vectors into a lower-dimensional space, capturing the most salient features of the data. This compressed representation is then decoded back to reconstruct the original input, ensuring that the encoded features retain the essential information from the original vectors [14].

Stage 3: Swarm-Enhanced Clustering

Finally, the output from the autoencoder is subjected to clustering. We independently apply two different swarm intelligence algorithms—Genetic Algorithms [15] and Artificial Bee Colony [16] —to benchmark their performance in identifying and forming distinct clusters. These algorithms enhance the capability of traditional K-Means clustering by leveraging their respective exploratory capabilities.

- Genetic Algorithm (GA): This algorithm optimizes the clustering process by iteratively improving the cluster centroids based on selection, mutation, and crossover processes. The GA starts by vectorizing topics and defining K-Means and GA parameters. It then generates a random initial population and iteratively improves it based on the silhouette coefficient until the optimal population is found. This evolutionary approach ensures robust exploration of the solution space, enhancing the clustering quality [15].
- Artificial Bee Colony (ABC): This algorithm mimics the foraging behavior of bees to find the optimal clustering by exploring and exploiting the solution space effectively. The ABC initializes by vectorizing topics and defining ABC setup parameters. It generates a random population and updates employed onlooker, and scout bees iteratively based on the silhouette coefficient until the optimal population is reached. This bio-inspired approach balances exploration and exploitation, leading to effective clustering outcomes [16].

By integrating LDA with swarm intelligence algorithms, our schema (Fig. 1.) represents an integrated, layered approach to topic modeling. Text, time, and location are independently analyzed but ultimately unified to inform a clustering process optimized through the independent application of swarm intelligence algorithms. This comprehensive approach not only captures the multi-dimensional nature of the data but also leverages advanced clustering techniques to produce meaningful and well-defined topic clusters.

### 3.3. A UNIFIED LDA APPROACH WITH BENCHMARKING OF SWARM-ENHANCED CLUSTERING TECHNIQUES

In the second schema of our methodology (Fig. 2), we address the spatio-temporal topic modeling task through a unified implementation of Latent Dirichlet Allocation (LDA), followed by the independent application of two different swarm intelligence algorithms for clustering. Unlike the first schema, where text, time, and location were independently analyzed, the second schema commences with the concatenation of these elements into a single vector for each document (thread). Each element (text of the post, location of the user, and time of publication) is treated as part of a unified, comprehensive document. This approach allows us to maintain the contextual linkages between these elements, fostering an integrated analysis that inherently reflects their interconnections [17].

Stage 1: Data Concatenation

Each element—text of the post, location of the user, and time of publication—is treated as part of a unified, comprehensive document. This approach allows us to maintain the contextual linkages between these elements, fostering an integrated analysis that inherently reflects their interconnections. By concatenating the text posts, timestamps, and location data into a single vector, we create a multifaceted representation of each document.

Stage 2: LDA on Concatenated Features

Upon constructing these comprehensive vectors, LDA is applied to extract topics. Given the incorporation of textual, temporal, and spatial elements in each vector, the derived topics are inherently spatio-temporal, reflecting themes that capture the interplay between the content of discussions, when they occurred, and where the participants were located. This unified application of LDA allows for a holistic analysis, where the interconnectedness of the different data aspects is preserved and leveraged.

Stage 3: Autoencoder Processing

The resulting topic distributions from the LDA stage serve as input to an autoencoder. Autoencoders are neural networks designed to learn efficient codings of the input data, providing compressed, meaningful representations that retain essential information. This step further refines the representation of the spatio-temporal topics, ensuring that the most salient features are captured and utilized in the subsequent clustering process.

Stage 4: Swarm-Enhanced Clustering

After topic extraction and autoencoder processing, we employ the K-Means clustering algorithm to group these topics. To assess the performance of different optimization strategies for this clustering process, we independently apply two swarm intelligence algorithms—Genetic Algorithms [15] and Artificial Bee Colony [16]. Each of these algorithms is benchmarked against standard K-Means clustering, evaluating their respective capabilities to form distinct and meaningful clusters.

By unifying the analysis of text, time, and location, and benchmarking different swarm-enhanced clustering techniques, the second schema (Schema II) offers an integrated approach to spatio-temporal topic modeling.

It provides valuable comparative insights into the effectiveness of different clustering strategies, demonstrating the benefits of combining LDA with advanced swarm intelligence algorithms for comprehensive and meaningful topic extraction and clustering.

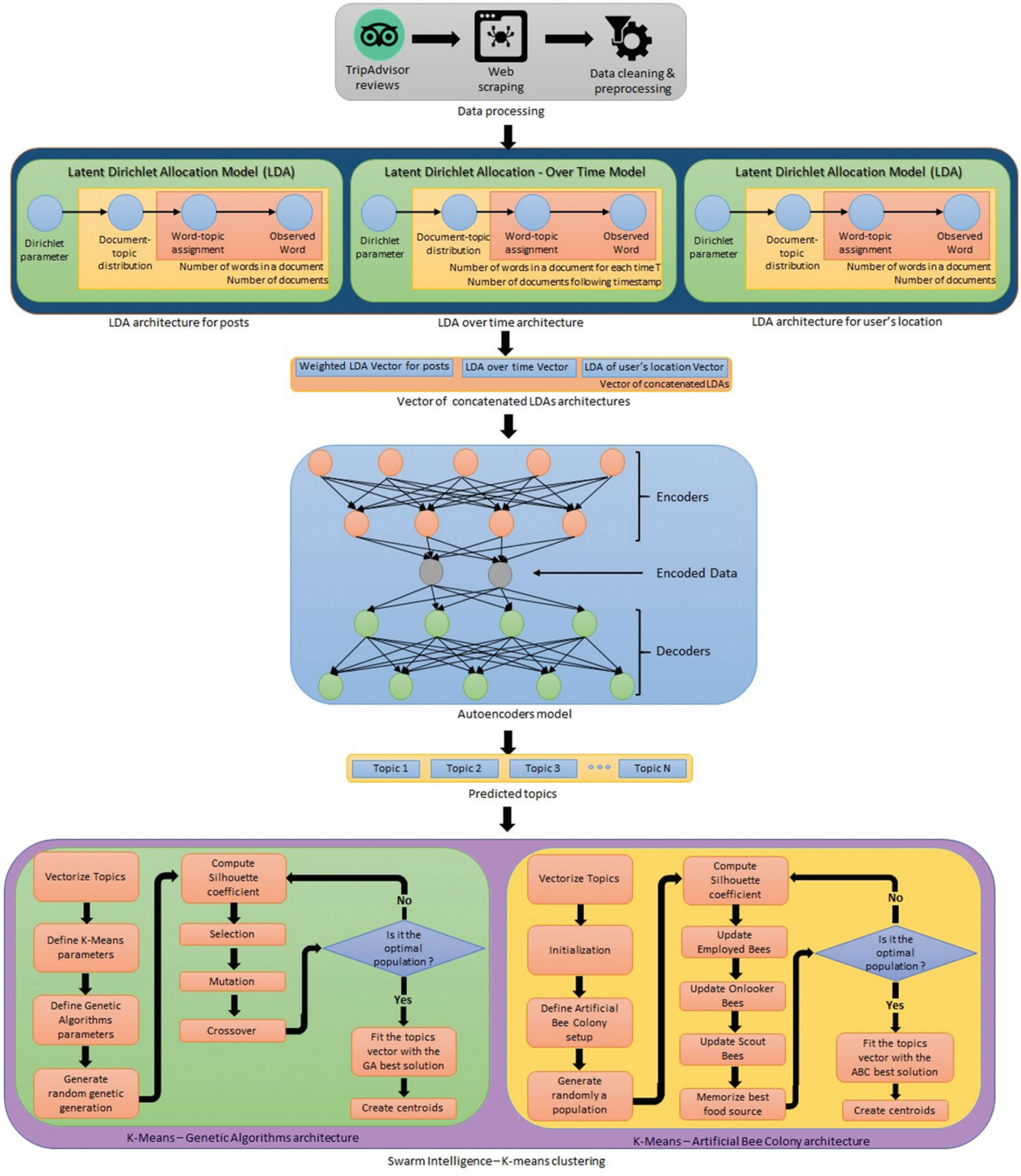



**Fig. 1.** MultLayered LDA approach with swarm intelligence clustering

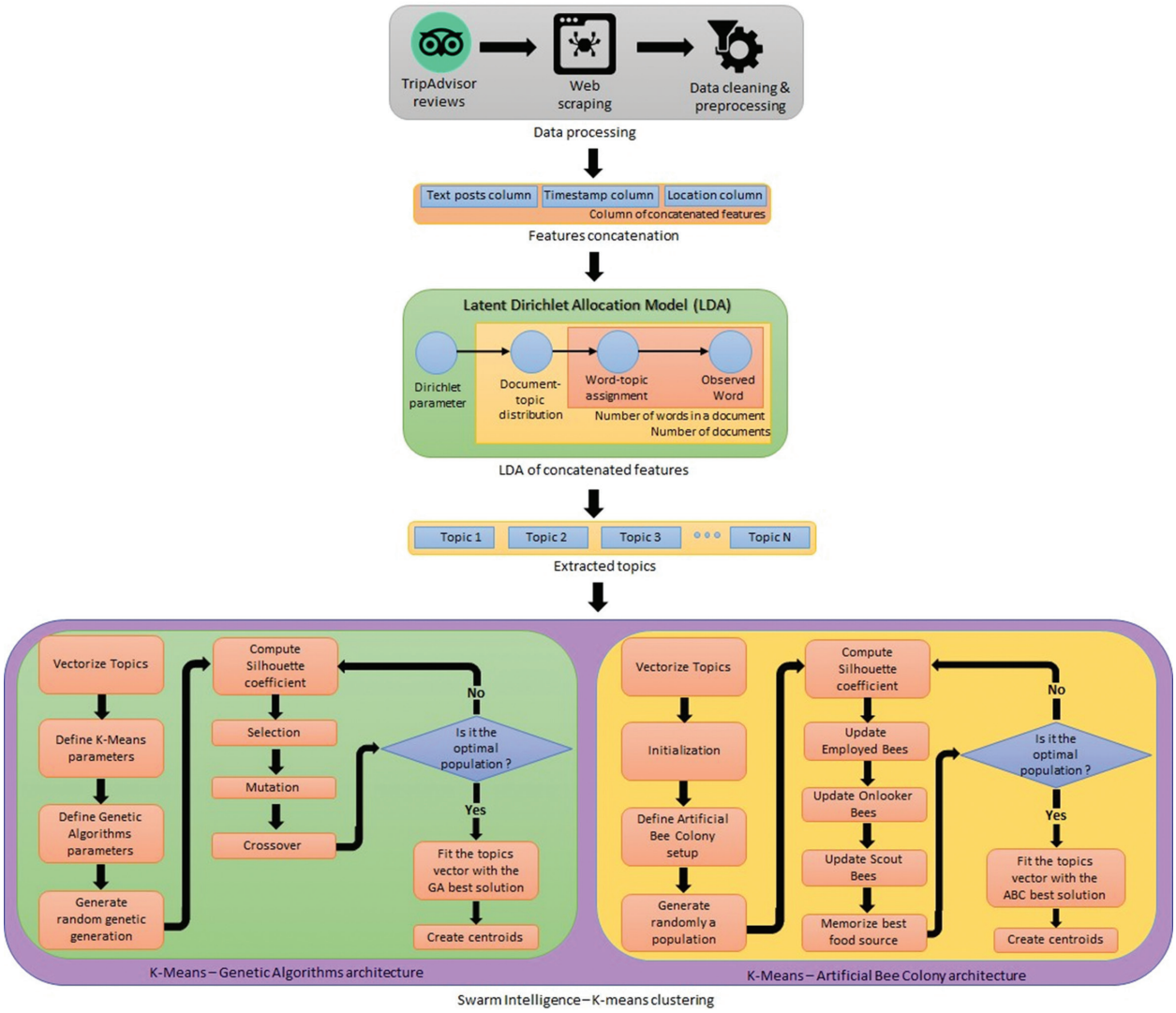



**Fig. 2.** Unified LDA approach with swarm intelligence clustering

## 4. RESULTS AND DISCUSSION

### 4.1. RESULTS

In this study, we conducted a comprehensive comparison of two different methods applied for Spatio-Temporal Topic Modeling, namely, Latent Dirichlet Allocation (LDA) and a hybrid approach that combines K-Means with Swarm Intelligence algorithms. These methods were applied to data derived from TripAdvisor, divided into three vectors: posts, the time of posts, and the location of the post's author. The performance of both methods was evaluated through several metrics, leading to compelling findings. The results from the two schemas provide interesting insights into the performance of different approaches to Spatio-Temporal Topic Modeling. Various metrics, including Coherence [18], Perplexity [19], and Topic Diversity [20] are used respectively for evaluating topic quality, and Silhouette Score [21], Davies-Bouldin Score [22], and Calinkski-Harabasz Score [23] to evaluate clustering performance. The various scores on the tables serve as metrics for the quality and performance of the applied Spatio-Temporal Topic Modeling approach. For all K-Means setups (i.e. Traditional K-Means and Swarm Intelligence K-Means), we set K to predict 20 clusters. The table below shows K-Means clustering with a loop over gamma values.

**Table 1.** K-Means clustering scores on Gamma values evaluation

| Gamma Values | Silhouette Score | Davies-Bouldin Index | Calinski-Harabasz Index |
|---|---|---|---|
| 5 | 0.08 | 2.52 | 1445 |
| 10 | 0.23 | 1.54 | 4428 |
| 15 | 0.25 | 1.45 | 5023 |
| 20 | 0.26 | 1.20 | 4872 |
| 25 | 0.26 | 1.23 | 4834 |
| 30 | 0.25 | 1.46 | 4893 |
| 35 | 0.25 | 1.29 | 4915 |
| 40 | 0.26 | 1.45 | 4953 |
| 45 | 0.25 | 1.21 | 4949 |
| 50 | 0.25 | 1.46 | 4965 |

In Fig. 1., the Silhouette score, which provides a measure of how similar an object is to its own cluster

compared to other clusters, improves as the gamma value increases, with the highest score being 0.26 at a gamma value of 20. The Davies-Bouldin Score, indicative of intra-cluster similarity, decreases as the gamma value increases, suggesting that a higher gamma value results in more distinct clusters. Similarly, the Calinkski-Harabasz Score, indicative of the degree of separation between clusters, increased as the gamma value increased, suggesting more well-separated clusters with an increased gamma value.

**Table 2.** Multi-layered LDA topics quality

| Model | Coherence | Perplexity | Topic Diversity |
|---|---|---|---|
| LDA for posts | 0.53 | -7.79 | 0.70 |
| LDA over time | 0.52 | -7.64 | 0.68 |
| LDA using author's location | 0.51 | -8.08 | 0.71 |

Examining the second table, we can see the performance of the LDA models for posts, LDA over time, and LDA using the author's location. The coherence score indicates that the topics generated were relevant and meaningful. This is particularly noticeable in the weighted LDA for posts, where the coherence score was the highest at 0.53. The Perplexity scores suggest a reasonable predictive performance of the models, with LDA using the author's location having the lowest Perplexity score of -8.08, indicating a better model. Lastly, the Topic Diversity values suggest a good spread of words across the identified topics, with LDA using the author's location demonstrating the highest Topic Diversity at 0.71. This suggests that this model was more successful in ensuring a broader range of topics.

**Table 3.** Multi-layered LDA clustering scores

| | Models | | |
|---|---|---|---|
| Metrics | K-Means | K-Means – Genetic Algorithms | K-Means – Artificial Bee Colony |
| Silhouette Score | 0.26 | 0.26 | 0.35 |
| Davies-Bouldin Index | 1.20 | 1.24 | 0.81 |
| Calinski-Harabasz index | 4872 | 5047 | 44495 |

In examining the Silhouette Score, both the K-Means and Genetic Algorithms - K-Means methods achieved an identical value of 0.26, indicating a satisfactory degree of cluster compactness and separation, albeit not exceptional. In stark contrast, the application of the Artificial Bee Colony - K-Means approach demonstrated a substantial improvement, with a Silhouette Score of 0.35, reflecting superior cluster cohesion and separation. Further evaluation using the Davies-Bouldin Score showed a slight increase from 1.20 for K-Means to 1.24 when Genetic Algorithms were integrated, representing a moderate degree of cluster separation. However, the application of the Artificial Bee Colony - K-Means algorithm resulted in a score of 0.81, markedly lower than the aforementioned methods, signifying the provision of more distinct and well-separated clusters. With regard to the Calinski-Harabasz Score, the K-Means method yielded a score of 4872, experiencing a marginal improvement to 5047 when Genetic Algorithms were incorporated. Yet, when contrasted with these results, the Artificial Bee Colony - K-Means approach far surpassed both with a remarkable score of 44495. This substantial enhancement signifies that this method generates clusters that are not only denser but also more clearly separated.

These results suggest that using weighted LDA with optimized gamma values in Spatio-Temporal Topic Modeling can lead to significant improvements in topic relevance, predictability, and diversity in the tourism industry [24], which can be beneficial for gaining insights and trends from user-generated content. The following plots show the UMap projection of the predicted clusters using the swarm intelligence approaches:

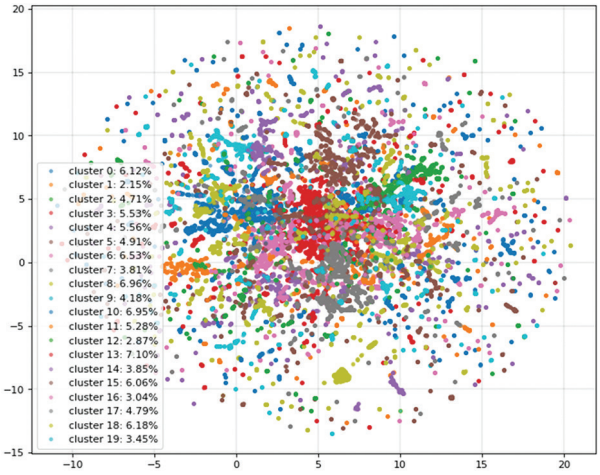



**Fig. 3.** MultiLayered LDA – Genetic Algorithms K-Means UMAP Projection

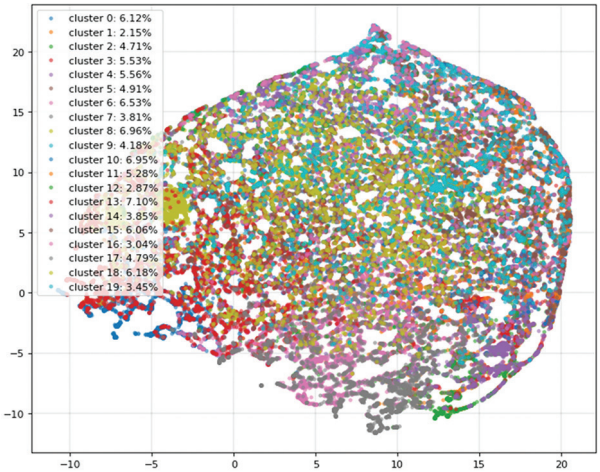



**Fig. 4.** MultiLayered LDA – Artificial Bee Colony K-Means UMAP Projection

The UMAP projection of the twenty predicted clusters offers illuminating insights into the relationships and dependencies within and across clusters. Particularly, the Artificial Bee Colony - K-Means (ABC-K-Means)

model (Fig. 4.) demonstrates a significant degree of interrelatedness and dependence among clusters. The proximity of some clusters and the absence of discernible boundaries in others suggest possible correlations and mutual influences between the topics within these groups. This pattern of clustering indicates that the ABC-K-Means model is adept at recognizing and incorporating the inherent relationships and dependencies in the data, thereby producing clusters that capture the multidimensional structure of the dataset. In contrast, the Genetic Algorithms - K-Means model (Fig. 3.) reveals a notably different pattern. The clusters in this projection appear to be more dispersed and independent, with clear demarcations separating the individual clusters. This spread signifies a higher degree of randomness in the distribution of topics across clusters [25]. The lack of apparent relationships or dependencies between clusters may suggest that this model tends to view each topic as an independent entity, leading to a more scattered and separated clustering [26]. It is therefore inferred that the Genetic Algorithms - K-Means model may be more suitable for datasets where topics are distinct and unrelated.

The following section presents an in-depth analysis and interpretation of the outcomes derived from the second method implemented in this study. The discussion that follows examines the effectiveness and the distinctiveness of this approach, drawing upon various evaluation metrics to assess the quality, coherence, and diversity of the generated topics

**Table 4.** Unified LDA topics quality

| Model | Coherence | Perplexity | Topic Diversity |
|---|---|---|---|
| LDA for concatened features (Posts + Timestamps + Locations) | 0.48 | -7.49 | 0.49 |

In Method II, the LDA model's performance, in terms of Coherence and Perplexity scores, is slightly less optimal compared to Method I, with scores of 0.48 and -7.49, respectively. However, the Topic Diversity score is closely aligned with that of Method I, implying a robust distribution of words across the identified topics.

**Table 5.** Unified LDA clustering scores

| | Models | | |
|---|---|---|---|
| Metrics | K-Means | K-Means – Genetic Algorithms | K-Means – Artificial Bee Colony |
| Silhouette Score | 0.25 | 0.78 | 0.72 |
| Davies-Bouldin Index | 1.27 | 0.6 | 0.61 |
| Calinski-Harabasz index | 5286 | 133744 | 122475 |

Traditional clustering methodologies like K-Means attain a marginally lower Silhouette Score of 0.25 compared to Method I. However, the incorporation of Genetic Algorithms and Artificial Bee Colony techniques substantially amplifies these scores, signaling the formation of better-defined and distinct clusters. Significantly, the Genetic Algorithms strategy achieved the highest Calinski-Harabasz Score, indicative of highly dense and well-separated clusters. To summarize, while both methods showcase impressive performances, their approaches to handling data diverge. Method I, deploying separate LDAs, presents slightly superior coherence and lower perplexity but lags in clustering metric performance. In contrast, Method II, despite demonstrating lower coherence and increased perplexity, stands out significantly in terms of its ability to form clusters, particularly when swarm intelligence algorithms are applied. The UMAP projections of the three clustering methodologies - K-Means, Genetic Algorithms-KMeans (GA-K-Means), and Artificial Bee Colony-K-Means (ABC-K-Means) - provide a visual insight into their effectiveness in data segregation. These projections, presented below, highlight the discernible differences between the approaches.

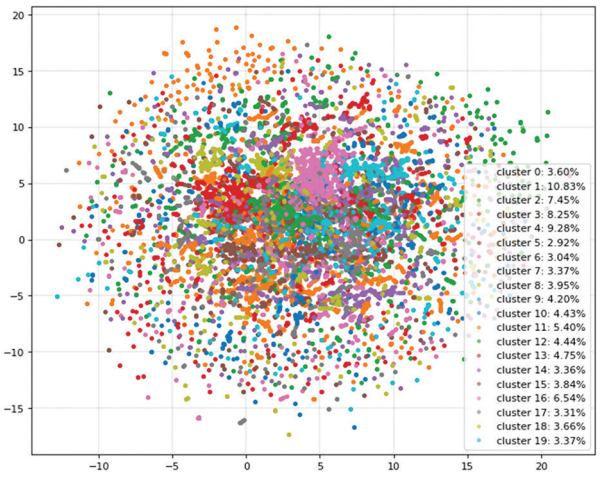



**Fig. 5.** Unified LDA –K-Means UMAP Projection

For the K-Means method (Fig. 5.), the clusters appear to be dispersed seemingly at random, with overlapping boundaries that suggest a lack of strong relationships between clusters.

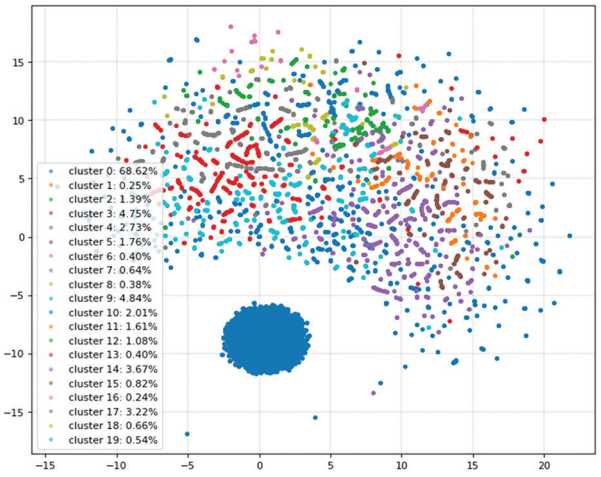



**Fig. 6.** Unified LDA – Genetic Algorithms K-Means UMAP Projection

In contrast, the UMAP projection for the GA-K-Means (Fig. 6.) method shows distinct clusters, suggesting a more structured segregation of data.

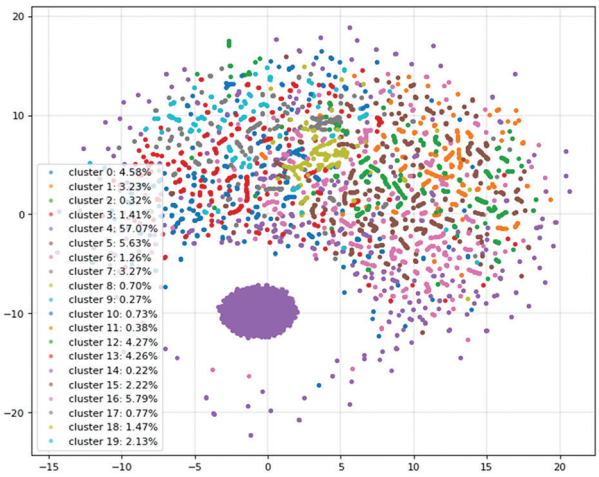



**Fig. 7.** Unified LDA – Artificial Bee Colony K-Means UMAP Projection

The ABC-K-Means method's projection (Fig. 7.) reveals a striking feature: a substantial cluster comprising a large number of samples, clearly differentiated from the remaining clusters. This dominant cluster's presence suggests potential relationships or dependencies between this prominent cluster and the other smaller clusters.

In the dynamic domain of the tourism industry, understanding the patterns and themes in tourists' discussions can provide crucial insights for decision-makers [27]. Word Cloud visualizations offer a potent tool to discern these patterns by prominently displaying the dominant terms within a text corpus. When applied to the clusters identified in our study, Word Clouds [28] can reveal the principal themes of each cluster, allowing us to ascertain the primary topics of discussion, temporal trends, and geographic references.

Considering the volume of data and the constraints of space in a research paper, presenting all 20 clusters for each methodology might not be feasible. Instead, a representative selection that conveys the distinctness and thematic richness of these clusters can effectively serve the same purpose.

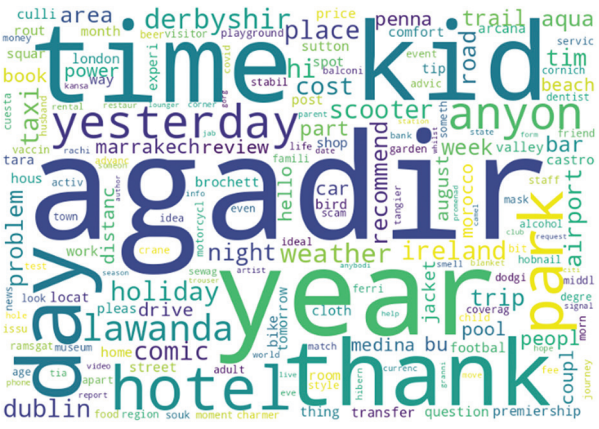




**Fig. 8.** WordCloud of Unified LDA-Swarm Intelligence (Cluster 4, Cluster 0)

For instance, consider the two representative clusters visualized above (Fig. 8.). They provide a snapshot of how each cluster encapsulates a theme interlacing location, time, and related discussions. Despite being just a fraction of the total clusters, these examples sufficiently showcase how each cluster signifies a unique theme, thereby illustrating the capabilities of this spatiotemporal topic modeling approach.

### 4.2. DISCUSSION

Our findings make a substantial contribution to the burgeoning field of Spatio-Temporal Topic Modeling, particularly demonstrating the strengths and potentialities of our two proposed schemas in efficiently extracting pertinent information from an extensive corpus of user-generated content. Each schema boasts unique attributes and yields notable advancements within this domain. When evaluated through the lens of our findings, Schema I, which independently applies LDA to the text, time, and location, produces comparatively higher Coherence scores. This indicates that the topics discerned through this method are not only more pertinent but also imbued with deeper meaning. This outcome resonates with prior research that accentuates the pivotal role of context in comprehending user-generated content [29]. Nonetheless, the enhanced clustering scores that emerge when Genetic Algorithms and Artificial Bee Colony are integrated underscore the value of utilizing swarm intelligence algorithms to navigate large data clusters. In contrast, (Fig. 2.), despite a slightly lower coherence and higher perplexity, delivers outstanding results in forming distinct and homogenous clusters. The algorithm's proficiency is particularly noticeable when swarm intelligence methodologies are incorporated. This indicates that a comprehensive overview of the generated topics can be obtained when context, in terms of location and time, is amalgamated with text into a single vector, leading to better-separated and more homogeneous clusters [30]. These findings have practical implications that stretch beyond the sphere of academic interest, specifically, within the tourism sector, ascertaining pat-

terns and trends in tourist behaviors and preferences is pivotal for strategic decision-making. Our models are engineered to extrapolate nuanced insights from the enormous and continually expanding corpus of user-generated content available on platforms such as TripAdvisor. In the context of Morocco's tourism sector, the Schema II model demonstrated substantial relevance. The word maps generated from this model gave a comprehensive overview of the discussion trends, incorporating time and location variables. The visually displayed results illustrate the relational dynamics within the tourism sector. The insights drawn from these maps can be leveraged by Moroccan tourism authorities to understand the temporal and spatial patterns of discussions and to strategize their services accordingly. The results of our research contribute to the expanding body of literature exploring advanced methodologies for topic modeling and clustering. The study underscores the significance of innovative methods in analyzing user-generated content and the potential of these approaches in extracting actionable insights across various sectors, including tourism.

## 5. CONCLUSION

In conclusion, our study introduces two novel schemas for Spatio-Temporal Topic Modeling using data from TripAdvisor's Morocco forum. We found distinct advantages in each schema, providing key insights into topic modeling. Schema I, which applies Latent Dirichlet Allocation (LDA) independently to text, time, and location, offered more coherent and predictable topics. Schema II, which integrates context and content into one vector, excelled in forming distinct clusters, especially with swarm intelligence algorithms like Genetic Algorithms.

These findings bear significant implications for tourism, enabling decision-makers to identify trends, enhance services, and create strategic plans based on tourist behavior. However, we acknowledge our study's limitations, including a reliance on TripAdvisor data and a fixed temporal scope (2007-2023). Future research will seek to integrate more data sources, apply advanced topic modeling techniques, and expand the temporal range to validate our schemas further.

## 6. REFERENCES


[1] S. Easton, N. Wise, "Online portrayals of volunteer tourism in Nepal: Exploring the communicated disparities between promotional and user-generated content", Worldwide Hospitality and Tourism Themes, Vol. 7, No. 2, 2015, pp. 141-158.

[2] Y. Liu et al. "Spatio-temporal topic models for check-in data", Proceedings of the IEEE International Conference on Data Mining, Atlantic City, NJ, USA, 14-17 November 2015, pp. 889-894.

[3] W. Min et al. "Multimodal spatio-temporal theme modeling for landmark analysis", IEEE MultiMedia, Vol. 21, No. 2, 2014, pp. 20-29.

[4] X. Luna et al. "Predictive spatio-temporal models for spatially sparse environmental data", Statistica Sinica, Vol. 15, 2005, pp. 547-568.

[5] J. Chen et al. "Local topic detection using word embedding from spatio-temporal social media", Proceedings of the International Conference on Neural Information Processing, 2019, pp. 629-641.

[6] K. Zhao et al. "Topic exploration in spatio-temporal document collections", Proceedings of the International Conference on Management of Data, San Francisco, CA, USA, June 2016.

[7] M. Asghari, D. Sierra-Sosa, A. S. Elmaghraby, "A topic modeling framework for spatio-temporal information management", Information Processing & Management, Vol. 57, No. 6, 2020, p. 102340.

[8] Y. Zhang, C. F. Eick, "Tracking events in Twitter by combining an LDA-based approach and a density-contour clustering approach", International Journal of Semantic Computing, Vol. 13, No. 1, 2019, pp. 87-110.

[9] C. González-Santos, M. A. Vega-Rodríguez, C. J. Pérez, "Addressing topic modeling with a multiobjective optimization approach based on swarm intelligence", Knowledge-Based Systems, Vol. 225, 2021, p. 107113.

[10] Q. Zhang, H. Li, "MOEA/D: A multiobjective evolutionary algorithm based on decomposition", IEEE Transactions on Evolutionary Computation, Vol. 11, No. 6, 2007, pp. 712-731.

[11] E. Shabaninia et al. "Transformers in action recognition: A review on temporal modeling", abs/2302.01921, 2022.

[12] J. Diaz et al. "Spatio-temporal conditioned language models", Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 25-30 July 2020.

[13] D. Maier et al. "Applying LDA topic modeling in communication research: Toward a valid and reliable methodology", Communication Methods and Measures, Vol. 12, No. 2-3, 2018, pp. 93-118.

[14] W. Wang et al. "Topic-guided variational auto-encoders for text generation", arXiv:1903.07137, 2019.

[15] S. Forrest, "Genetic algorithms", ACM Computing Surveys, Vol. 28, No. 1, 1996, pp. 77-80.

[16] D. Karaboga, B. Akay, "A comparative study of artificial bee colony algorithm", Applied Mathematics and Computation, Vol. 214, No. 1, 2009, pp. 108-132.

[17] Q. Gu, Z. Li, J. Han, "Linear discriminant dimensionality reduction", Proceedings of Machine Learning and Knowledge Discovery in Databases: European Conference, Athens, Greece, 5-9 September 2011, pp. 549-564.

[18] K. Stevens, P. Kegelmeyer, D. Andrzejewski, D. Buttler, "Exploring topic coherence over many models and many topics", Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, Jeju Island, Korea, July 2012, pp. 952-961.

[19] W. Zhao et al. "A heuristic approach to determine an appropriate number of topics in topic modeling", BMC Bioinformatics, Vol. 16, Suppl. 13, 2015.

[20] F. Nan, R. Ding, R. Nallapati, B. Xiang, "Topic modeling with Wasserstein autoencoders", arXiv:1907.12374, 2019.

[21] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis", Journal of Computational and Applied Mathematics, Vol. 20, 1987, pp. 53-65.

[22] D. L. Davies, D. W. Bouldin, "A cluster separation measure", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 2, 1979, pp. 224-227.

[23] T. Calinski, J. Harabasz, "A dendrite method for cluster analysis", Communications in Statistics - Simulation and Computation, Vol. 3, No. 1, 1974, pp. 1-27.

[24] Q. Li, S. Li, S. Zhang, J. Hu, J. Hu, "A review of text corpus-based tourism big data mining", Applied Sciences, Vol. 9, No. 16, 2019, p. 3300.

[25] W. Li, Y. Feng, D. Li, Z. Yu, "Micro-blog topic detection method based on BTM topic model and K-means clustering algorithm", Automatic Control and Computer Sciences, Vol. 50, 2016, pp. 271-277.

[26] M. S. Handcock, A. E. Raftery, J. M. Tantrum, "Model-based clustering for social networks", Journal of the Royal Statistical Society: Series A (Statistics in Society), Vol. 170, No. 2, 2007, pp. 301-354.

[27] Z. Doborjeh, N. Hemmington, M. Doborjeh, N. Kasabov, "Artificial intelligence: a systematic review of methods and applications in hospitality and tourism", International Journal of Contemporary Hospitality Management, Vol. 34, No. 3, 2022, pp. 1154-1176.

[28] F. Heimerl, S. Lohmann, S. Lange, T. Ertl, "Word Cloud Explorer: Text analytics based on word clouds", Proceedings of the 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6-9 January 2014, pp. 1833-1842.

[29] L. Shifman, "Online entertainment | Cross-cultural comparisons of user-generated content: An analytical framework", International Journal of Communication, Vol. 10, 2016, p. 20.

[30] A. El-Kishky, Y. Song, C. Wang, C. Voss, J. Han, "Scalable topical phrase mining from text corpora", arXiv:1406.6312, 2014.

# TelMedAI: A Framework for Patient Speech Recognition and Conversion into Desired Language Towards Telemedicine System

Original Scientific Paper

**MrudulaOwk***
GITAM University,
Department of CSE, GITAM School of Technology
Rushikonda, Visakhapatnam -530045
mowk@gitam.edu

**Pusarla Sindhu**
GITAM University,
Department of CSE, GITAM School of Technology
Rushikonda, Visakhapatnam -530045
spusarla@gitam.edu

**Deepthi Godavarthi**
VIT-AP University,
School of Computer Science and Engineering(SCOPE)
Amaravati, Andhra Pradesh, India- 522237
deepthi.g@vitap.ac.in

**T. Krishna Mohana**
Department of ECE Aditya College of Engineering
Krishnamohana_ece@acoe.edu.in

*Corresponding author

***Abstract*** *– Telemedicine is the practice of technology-enabled remote communication between patient and doctor. This phenomenon in healthcare has the potential to make services affordable and save time and money. Besides telemedicine allows care givers and family members to join conversations with doctors. Indian government initiated the National Telemedicine Network (NTN) to serve remote areas in healthcare by integrating existing healthcare facilities.Literature has revealed that existing works lack in an integrated approach for patient speech translation in language-independent fashion and automatic detection of disease and symptoms based on speech.There is a need for an automated system using Artificial Intelligence (AI) to recognize patient's speech and identify symptoms based on given audio description. We proposed a framework known as* ***TelMedAI*** *which is designed to recognize patient speech to comprehend disease symptoms besides converting the speech text into desired language. The framework is useful for realizing a telemedicine system. Speech to Speech (STS) module takes the patient's audio content into English audio. STS module exploits the Bi-LSTM model with an encoder, decoder and attention mechanism for translation. Then Google Speech API is used to convert English audio into English text. Then the framework exploits Natural Language Processing (NLP) to improve the quality of text. Afterwards, the disease and symptoms miner module eventually recognizes a list of diseases and corresponding symptoms. We proposed an algorithm known as Learning based Disease and Symptom Recognition from Patient Speech (LbDSRPS). This algorithm has the functionality to develop* ***TelMedAI*** *which helps doctors in telemedicine. Our empirical study has revealed that* ***TelMedAI*** *takes technology-driven telemedicine research forward significantly. The highest accuracy achieved by the proposed framework is 68.13% which is much better than the baseline LSTM model used for voice translation.*

***Keywords***: *Telemedicine System, Patient Speech Recognition, Deep Learning, Artificial Intelligence, Multi-Lingual Text Conversion*

*Received: November 20, 2023; Received in revised form: April 30, 2024; Accepted: April 30, 2024*

## 1. INTRODUCTION

Research on the realization of telemedicine systems with technology-enhanced approaches is most relevant in the contemporary world. The recent COVID-19 pandemic has emphasized the importance of telemedicine. Telemedicine is the system which enables people of all walks of life to gain access to healthcare services just by making a phone call. This phenomenon has many significant advantages such as saving time, effort, and money besides getting medical advice without time and geographical restrictions. In order to develop a telemedicine system, it is important to have efficient translation services. The rationale behind this is that the patient may speak any language and doctor needs to understand what the patient says while doctor does not know the language of the patient. Therefore, telemedicine has many implementation challenges as

explored in [1]. Nevertheless, there has been research endeavours to realize technology-driven telemedicine systems as found in the literature.

Machine learning models are explored in [2] and [3] towards understanding patient diseases and symptoms based on the description given by persons. Speech recognition research carried out as found in [4-8] provide the basis for understanding the importance of patient speech recognition in telemedicine besides how the existing methods could achieve it. Voice recognition in an IoT-integrated home automation system is investigated in [9] while a similar kind of approach is used in [10] as part of remote patient monitoring. Researchers explored telemedicine with novel technologies as discussed in [1, 11-13]. The realization of telemedicine has many challenges as discussed in [11]. One significant challenge is the voice-to-voice conversion from source to target language. Artificial intelligence and its usage in telemedicine is studied in [12]. Though there are innovative technologies, it is observed in [1] that telemedicine implementation is still challenging due to several complexities involved. Patient engagement early and disease diagnosis through the telemedicine system is investigated in [14]. From the literature, it is found that there is need for a system that can translate patient voices into English and mine diseases and symptoms to help doctors realise a telemedicine system. Towards this end, we proposed an AI enabled system for patient speech translation into desired target language and analyse patient's health symptoms and diseases. The novelty of the proposed system is it is designed to be language-independent and helps in analysing and identifying patient's disease and symptoms to assist doctors. The proposed system plays crucial role in realizing a technology-driven telemedicine system. Our contributions to this paper are as follows. We proposed a framework known as **TelMedAI** which is designed to recognize patient speech to comprehend disease symptoms besides converting the speech text into desired language and recognize the patient's disease and symptoms to develop the telemedicine system. We proposed an algorithm known as learning-based Disease and Symptom Recognition from Patient Speech (LbDSRPS) to develop**TelMedAI** which helps doctors in telemedicine. We built an application for evaluating **TelMedAI** and the underlying algorithm. The utility of bi-LSTM model used in the STS module is evaluated and the results are compared with baseline LSTM model.The remainder of the paper is structured as follows. Section 2 reviews the literature on existing research efforts using machine learning in the field of telemedicine. Section 3 presents the proposed framework along with its modus operandi in detail. Section 4 presents experimental results while section 5 concludes our work and provide limitations of the current work and directions for future research.

## 2. RELATED WORK

This section reviews the literature on existing research endeavours on the usage of techniques towards telemedicine. Zahia et al. [2] opines that pressure injuries burden healthcare systems. Non-invasive imaging, including Deep Learning, aids accurate assessment, but limited data hinders progress. Their investigation has revealed the need for non-invasive healthcare services that could be beneficial to general public. Dargan et al. [3] explored deep learning to show versatility and progress across various fields. Challenges include optimizing hierarchies and maintaining databases. The insights of their research reveal that deep learning models are very useful in computer vision applications. Stanovov et al. [4] proposed a cloud-based speech-controlled wheelchair system, emphasizing the benefits of multiple cloud speech recognition APIs for accuracy. Recognition of speech could help in solving many real world problems such as directing navigation of a wheel chair. Sescleiferet al. [5] focused on online crowdsourcing which aids in efficient and effective perceptual speech assessments, particularly beneficial for cleft palate surgeries.Speech perception plays crucial role in certain applications such as telemedicine where patient's speech needs to be understood accurately. Ivan et al. [9]. An IoT-enabled voice recognition system is explored. Kho et al. [11] observed that telemedicine faces implementation challenges due to inadequate attention to change management. A process-oriented approach is recommended. From their work is understood that there are gaps in existing telemedicine systems. One such gap is lack of usage of AI. Olivia et al. [12] investigated on advancements in digital technologies that offer transformative potential for ophthalmology, notably amid the COVID-19 pandemic, yet challenges persist. In their research, there is inference that improving end-to-end communications in healthcare considering every possibility helps in developing more robust applications. Spachos et al. [15] observed that the pandemic spurred changes in interactions with objects. Voice-activated IoT devices have healthcare potential, yet security concerns persist. Voice based communications in telemedicine system needs technology adoption to leverage seamless telemedicine services.

Lesso et al. [16] opined that telehealth in respiratory care lacks efficient cough detection. Proposed Hu moments-based system offers high sensitivity and specificity. However, it is observed that there is need for enhancing the system to have generalized system for all kinds of diseases. Erdene et al. [17] observed that mobile health technologies aid in early stroke detection through continuous monitoring with diverse sensor-equipped devices. Such systems are useful for alerting people of health issues. However, the availability of telemedicine can expect possibilities in rendering healthcare services. Jamshidiet al. [18] stated that COVID-19's global impact necessitates AI-driven solutions for diagnosis and treatment acceleration, in-

tegrating varied data sources for effective platforms. Covid pandemic also necessitated and reinforced the need for technology driven approach where people can have health services over phone. Panganibanet et al. [19] a deep learning model is evaluated for diagnosis of a patient's disease.Since deep learning models are found good at learning audio signals, using them in development of a telemedicine system could have a positive impact. Wang et al. [20] observed that COVID-19 intensified the role of robotics in healthcare. Despite advancements, challenges like cost and accessibility persist. Such challenges can be overcome with an efficient telemedicine system for some of the health issues. Dharmale et al. [21] exploited phonetic system in healthcare industry for leveraging speech recognition system. Choutri et al. [22] focused on human-drone interaction where speech is recognized of multiple languages. Kerwagen et al. [23] investigated on diagnostic management along with speech recognition and usability in healthcare. Shindel et al. [24] studied healthcare applications integrated with blockchain technology. Javaid et al. [25] explored machine learning and its utility in healthcare domain besides providing valuable insights. From the literature, it is found that there is a need for a system that can translate patient voice into English and mine diseases and symptoms for helping doctors realise a telemedicine system.

## 3. PROPOSED FRAMEWORK

We proposed a framework known as **TelMedAI** which is designed to recognize patient speech to comprehend disease symptoms besides converting the speech text into desired language. The framework is useful for realizing a telemedicine system. Overview of **TelMedAI** is shown in Fig.1. The source speaker (patient) may speak over the telephone in any language. If that language is not English, the proposed framework helps in converting a voice in different language to a voice in English. This conversion is known as Speech to Speech (STS) conversion. The STS module in the proposed system exploits the deep learning model LSTM for converting patient speech (voice) into English (voice). Fig. 2 illustrates how the STS conversion module does it. Once the patient's voice is converted to voice in English, it is one of the outcomes of the system which can be listened by the doctor directly. Further **TelMedAI** has provision for converting speech in English to English text. This conversion is carried out by speech to text conversion module of the framework.

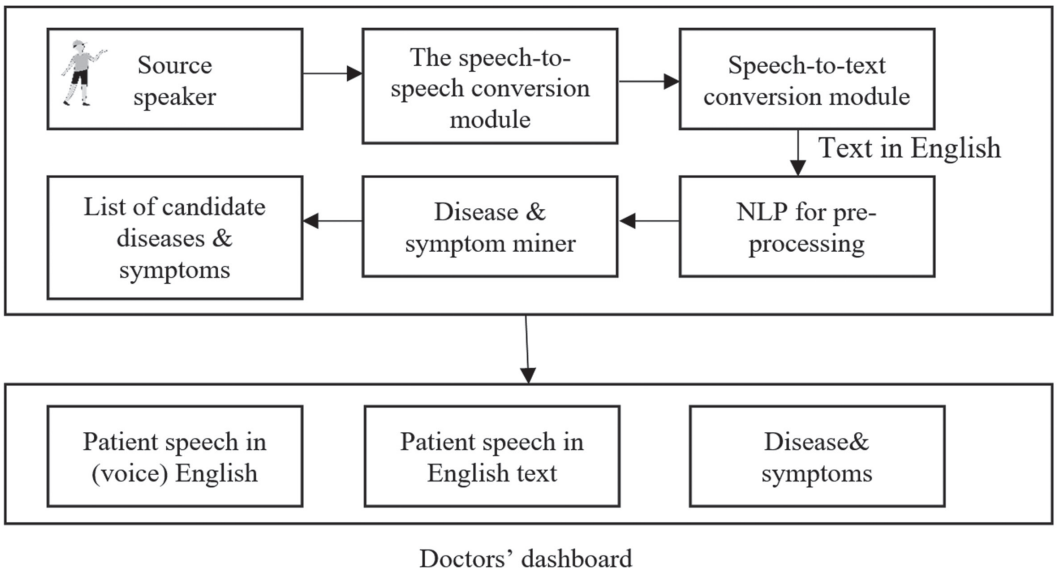



**Fig. 1.** Overview of the proposed framework known as TelMedAI useful for telemedicine

Text in English is another desired output useful to doctor in the telemedicine system. Afterwards, the text in English is subjected to pre-processing using NLP followed by text mining to discover disease and symptoms found in the patient's speech. The rationale and motivation behind three outcomes in doctor's dashboard is that it will enable doctor to revisit and correlate findings. Therefore, the third outcome of the system useful for doctor is the identified disease and its symptoms. Since voice input is translated into English, only textual data is analysed for identification of disease and symptoms. The STS conversion module makes use of bi-directional LSTM used as an encoder and decoder network. We preferred the LSTM model as order of the inputs and outputs is important. In other words, LSTM is well known for its ability to function in the temporal domain. Using the Fourier transform, the given patient's voice is converted to a Mel spectrogram of speech in the original language. This high level representation of the patient's voice is given to the Bi-LSTM model to convert data into linear spectrograms of speech in English (the target language). The spectrogram of speech of English is subjected to Griffin Lim Vocoder to convert the high-level representation to target voice file.

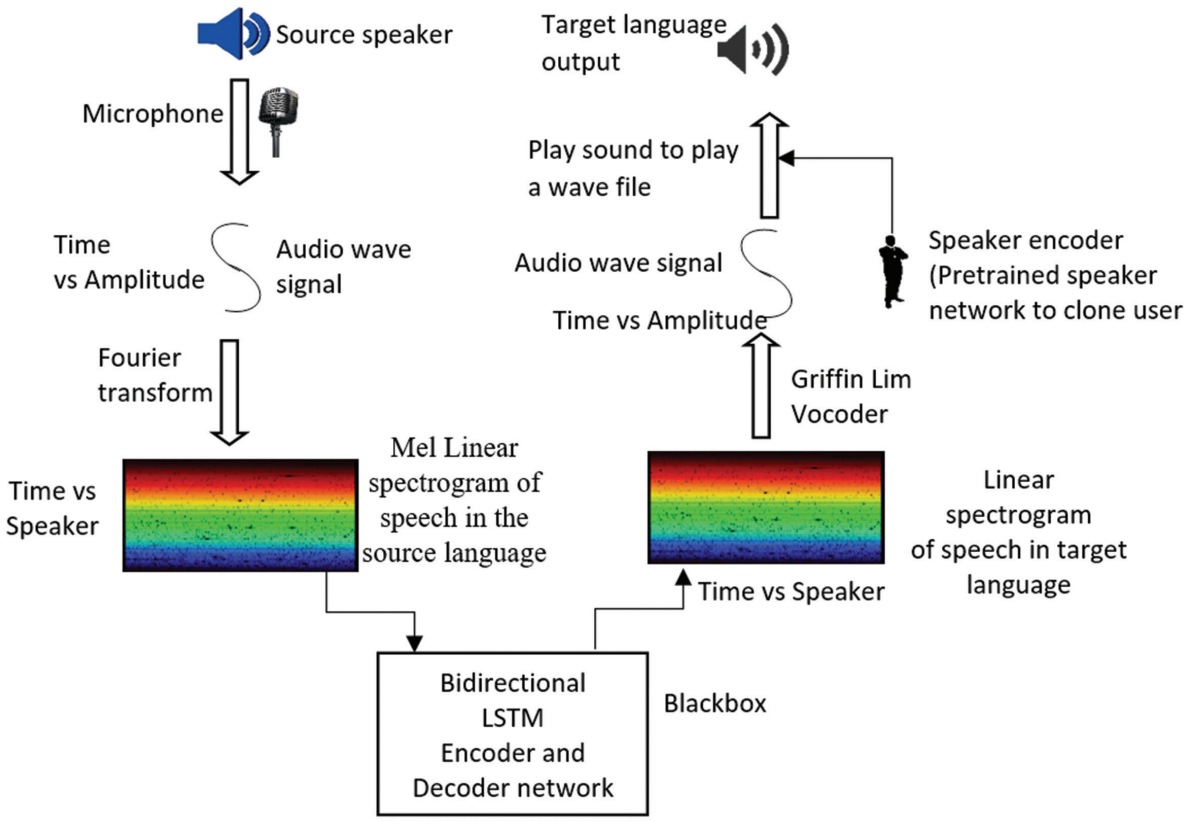



**Fig. 2.** Functional flow of STS conversion module

Once the source voice is converted to the target voice file, speech to text conversion module comes into picture. This module makes use of Google Speech API to translate voice in English to text in English. Before elaborating further process involved in the TelMedAI, we describe the encoder and decoder functionality involved in the STS conversion module. The basic encoder and decoder functionality of LSTM is shown in Fig. 3.

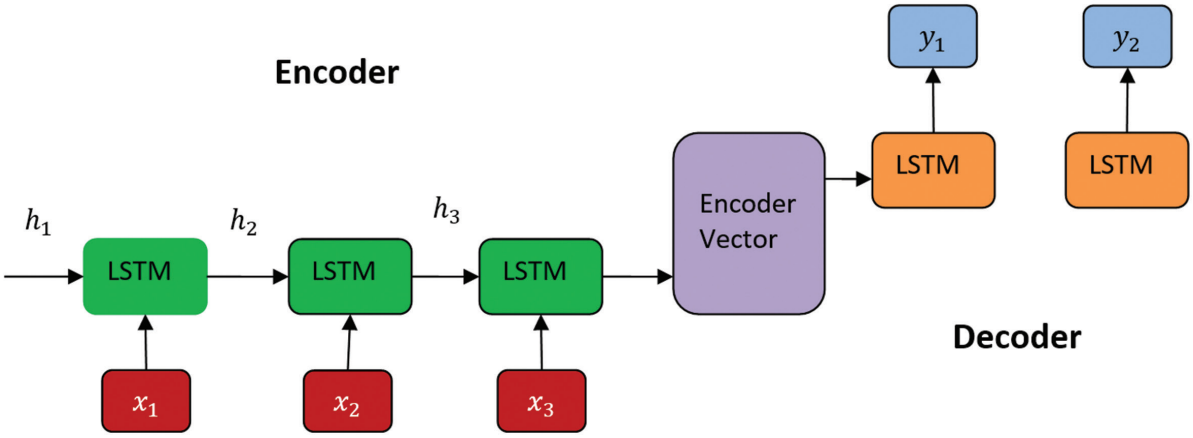



**Fig. 3.** LSTM based encoder and decoder network

The encoder is made up of three recurrent LSTM units while the decoder is made up of two recurrent LSTM units. Each unit acts on the given input and propagates output to the next LSTM unit. The encoder computes hidden states as expressed in Eq. 1 where $h_t$ denotes hidden state in given time step, $W$ refers to weight matrix.

$$h_t = f(W^{hh}h_{t-1} + W^{hx}x_t) \quad (1)$$

It involves the multiplication of weights associated with the previous hidden state and the input vector. Concerning decoder, computation of the hidden state is done as expressed in Eq. 2 where $h_t$ denotes hidden state in given time step, $W$ refers to weight matrix.

$$h_t = f(W^{hh}h_{t-1}) \quad (2)$$

Computation of final output in the absence of any activation function is carried out as expressed in Eq. 3 where $y$ refers to output, $W$ indicates weight matrix and t refers to given time step.

$$y_t = Softmax(W^s h_t) \quad (3)$$

It reflects output in the form of a multiplication matrix derived from a weight matrix associated with hidden vector in the given time step. An important consideration is the attention mechanism which has the potential to improve the network as it lies between encoder and decoder. The attention mechanism gets rid of any possible misalignment between the encoder and decoder. The Bi-LSTM-based block box shown in Fig. 2 plays an important role in converting patient's voice from one language to another language (STS). This process is further elaborated as illustrated in Fig. 4. Stacking LSTM units is done to develop the encoder which produces outputs form Mel spectrograms of the patient's voice in native language. Then the output is concatenated. Afterwards, an attention mechanism is used to optimize outcomes to be given to the decoder in order to avoid error. Then a stacked LSTM unit is used as decoder followed by a dense layer and Griffin Lim Vocoder to generate Mel spectrograms of target language. Here Griffin Lim Vocoder algorithm is used to generate target spectrograms. The attention mechanism used in the proposed system is multi-head attention as it is run many times instead of computing attention once.

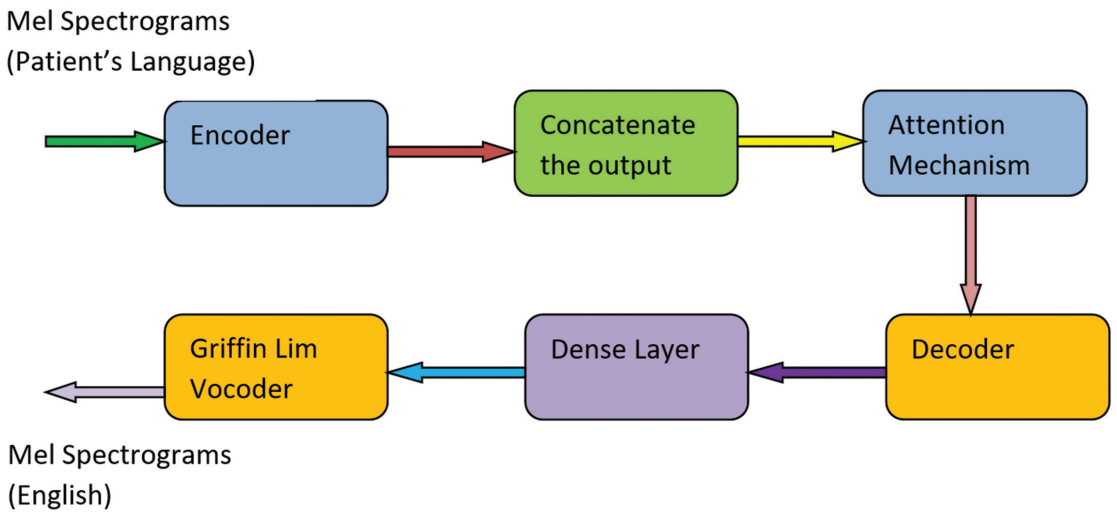



**Fig. 4.** Technical details of the pipeline involved in the STS module

The dense layer involved in the network is itself a well-connected neural network. It is used to generate outcomes as expressed in Eq. 4.

$$Output = activation\ (d\ ot(input,\ kernel) + bias) \quad (4)$$

It makes use of weight kernel matrix and input tensor to have a dot product which is nothing but second LSTM unit's output in the decoder of the network. Here the bias is optional and set to zero. MSE is used as a loss function to minimize error in computations. We also used Adam optimizer to improve accuracy of the model. Now let us get back to the framework **TelMedAI** presented in Fig. 1. After converting from English speech to text using Google Speech API, the resultant text is subjected to NLP for pre-processing. It is used to get rid of meaningless words in the speech. Afterwards, the disease and symptom miner is responsible for discovering disease (s) and corresponding symptoms. This module provides the desired convenience to doctors as it provides list of candidate diseases and corresponding symptoms. Our implementation of the disease and symptom miner module is influenced by the work of [21] where more technical information about how the diseases and the symptoms are identified from text in English. As presented in Algorithm 1, it takes patient speech as input. It is the voice of patient describing about his/her disease to seek doctor's advice through telemedicine system. After completion of the processing, the proposed algorithm results in three outputs that are useful to doctor in disease diagnosis and treatment. They are known as translated English voice files, English text and recognized disease and symptoms. The given patient voice is used to generate a Mel spectrogram. This is the high-level representation of patient's voice information sent to doctor. Then the algorithm makes use of Bi-LSTM with encoder, decoder along with attention mechanism to convert Mel spectrogram of source language voice into the Mel spectrogram of target language that is English.

Algorithm 1: Learning based disease and symptom recognition from patient speech

**Algorithm**: Learning-based Disease and Symptom Recognition from Patient Speech

**Input**

Patient speech (voice file)ov

**Outputs**

Translated English voice file $tv$

English text $tt$

recognized disease and symptoms $R$

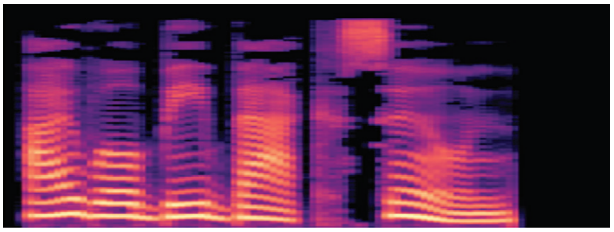
1. Begin
2. $S \leftarrow$GenerateMelSpectogram(ov)
3. Build Bi-LSTM model $m$

4. Compile the model *m*
5. *T*←*GenerateTargetMelSpectogram*(*m, S*)
6. *tv*←*GenerateEnglishVoiceFile* (*T, Griffin Lim Vocoder algorithm*)
7. *tt*←*GoogleSpeechAPI*(*tv*)
8. *tt'*←*NLPTechniques*()
9. *R*←*DiseaseAndSymptomMiner*(*tt'*)
10. Display *tv*
11. Display *tt*
12. Display *R*13.
13. End

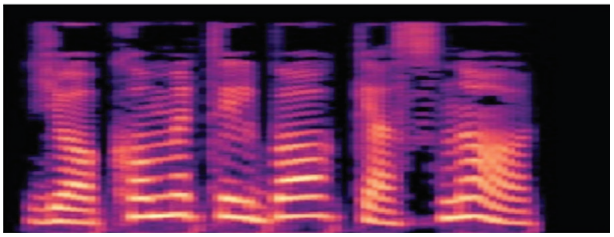
The target Mel spectrogram is given to the Griffin Lim Vocoder algorithm which generates English voice file. This process is known as STS (speech to speech) conversion. Once STS is completed, its resultant English voice file is converted to English text using Google Speech API. The resultant English text contains information provided by the patient to doctor but translated to English. This text is subjected to NLP techniques such as stop word removal, stemming and lemmatization to improve quality of the text. Afterwards, the disease and symptoms miner module is used to discover disease and corresponding symptoms from English text.

## 4. EXPERIMENTAL RESULTS

We implemented our framework TelMedAI using Python language, machine learning library and Google Speech API. This section presents experimental results in terms of patient voice converting to Mel spectrograms or source and target (English) languages. It also provides the accuracy of deep learning model proposed in this paper and compare the same with baseline LSTM model. Fig. 5 presents Mel spectrogram in source and target languages.



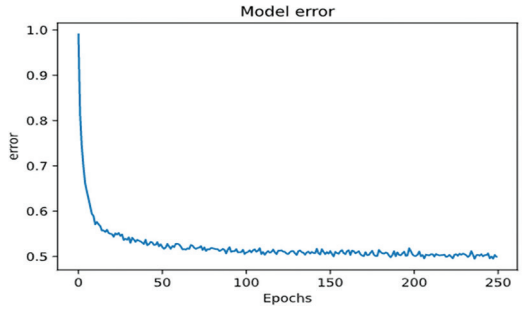
(a)



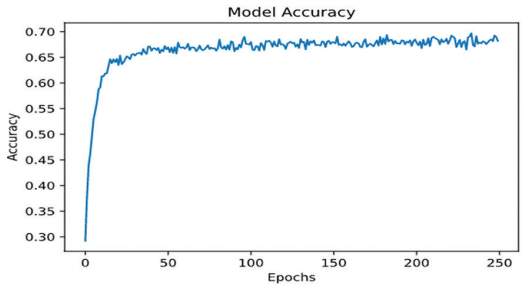
(b)

**Fig. 5.** Mel Spectrograms (a) for English sentence "I am suffering from fever" and (b) equivalent in Swedish

Patient's speech saying "Jag harbour" in Swedish language is converted to source language Mel Spectrogram and target language (English) Mel Spectrogram. English equivalent of the patient's speech is "I am suffering from fever".




(a)




(b)

**Fig. 6.** Performance in terms of (a) error and (b) accuracy of the Bi-LSTM with attention in the proposed system

The model performance in terms of error and accuracy is visualized against number of epochs as shown in Fig. 6. It is the result of Bi-LSTM with attention in the proposed system. Table 1 shows performance comparison.

**Table 1.** Shows performance comparison

| Model | Accuracy |
|---|---|
| Baseline LSTM | 62.24 |
| Bi-LSTM with Attention (proposed system) | 68.13 |

Accuracy of the Bi-LSTM model with the attention mechanism used in the proposed framework is compared with that of a baseline LSTM model.

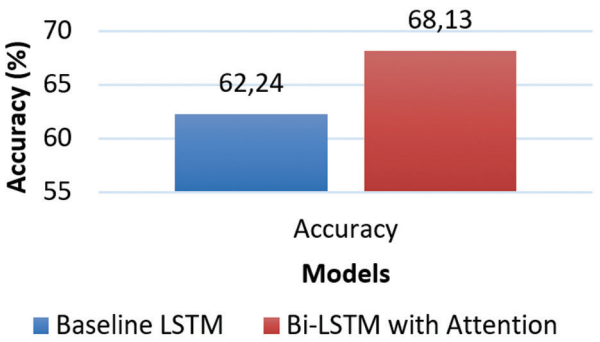



**Fig. 7.** Performance comparison between the proposed and existing models

Performance of the deep learning model used in the proposed system (Bi-LSTM with attention) is compared with the baseline LSTM model, as shown in Fig. 7, in terms of accuracy in translating the patient's source language voice to English voice. The proposed model in the paper outperformed the baseline model. Highest accuracy achieved by the proposed framework is 68.13%. Considering the difficulty in STS conversion, this accuracy is significant. However, the proposed system has several limitations that need to be overcome in future. First, it is implemented as a preliminary system and it is no way perfect to be used in a telemedicine system without further improvement. Second, the proposed system is yet a laboratory study which takes prerecorded patient voice as input. However, it needs to be improved to consider live voice of patients. Third, as of now, the system is yet to be integrated with a telephone line to capture patient's voice and evaluate the functionality. Fourth, the proposed system has to be improved further to deploy in a healthcare unit where doctors can take patient calls to understand their health issues, diagnose and give treatment.

## 5. CONCLUSION AND FUTURE WORK

In this paper, we proposed a framework known as **TelMedAI** which is designed to recognize patient speech to comprehend disease symptoms besides converting the speech text into desired language. The framework is useful for realizing a telemedicine system. Speech to Speech (STS) module takes patient's audio content into English audio. STS module exploits the Bi-LSTM model with encoder, decoder and attention mechanism for translation. The deep learning model is used to convert source language Mel Spectrogram into English Mel Spectrogram. The target Mel spectrogram is given to Griffin Lim Vocoder algorithm which generates English voice file. Then Google Speech API is used to convert English audio into English text. Then the framework exploits Natural Language Processing (NLP) to improve quality of text. Afterwards disease and symptoms miner module eventually recognizes a list of diseases and corresponding symptoms. We proposed an algorithm known as Learning based Disease and Symptom Recognition from Patient Speech (LbDSRPS). This algorithm has the functionality to develop **TelMedAI** which helps doctors in telemedicine.

Our empirical study has revealed that TelMedAI takes technology-driven telemedicine research forward significantly. As of now, our system is tested with patients' voice files. However, to develop a complete telemedicine system, there is need for much work to be one. In future, we intend to improve our system in two phases. In the first phase, we evaluate it with live patient's voice and in the second phase we deploy it in a healthcare unit for live testing and further improvement. Improving accuracy is also to be considered in future.

## 6. REFERENCES


[1] A. Khodadad-Saryazdi, "Exploring the telemedicine implementation challenges through the process innovation approach: A case study research in the French healthcare sector", Technovation, Vol. 107, 2021, p. 102273.

[2] S. Zahia, G. Zapirain, M. Begoña, S. Xavier, A. González, P. J. Kim, A. Elmaghraby, "Pressure injury image analysis with machine learning techniques: A systematic review on previous and possible future methods", Artificial Intelligence in Medicine, Vol. 102, No. C, 2020.

[3] S. Dargan, M. Kumar, M. R. Ayyagari, G. Kumar, "A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning", Archives of Computational Methods in Engineering, Vol. 27, 2019, pp. 1071-1092.

[4] A. Koložvari, R. Stojanović, A. Zupan, E. Semenkin, V. Stanovov, D. Kofjač, A. Škraba, "Speech-Recognition Cloud Harvesting for Improving the Navigation of Cyber-Physical Wheelchairs for Disabled Persons", Microprocessors and Microsystems, Vol. 69, No. C, 2019, pp. 179-187.

[5] A. M. Sescleifer, C. A. Francoisse, A. Y. Lin, "Systematic review: online crowdsourcing to assess perceptual speech outcomes", Journal of Surgical Research, Vol. 232, 2018, pp. 351-364.

[6] B. F. Zaidi, S. A. Selouani, M. Boudraa, M. S. Yakoub, "Deep neural network architectures for dysarthric speech analysis and recognition", Neural Computing and Applications, Vol. 33, 2021, pp. 9089-9108.

[7] P. T. Bradley, J. Patterson, "Attitudes to the Implementation of Speech and Language Therapist Led Low Risk Two Week Wait Clinic in the UK: A Survey Exploration Using Normalization Process Theory", Journal of Voice, Vol. 38, No. 1, 2021, pp. 86-95.

[8] C. Millar, L. B. Carey, A. E. Hill, S. Attrill, M. I. Avgoulas, E. Drakopoulos, C. A. Sutton, "Global citizenship and social justice among speech-language pathologists: A scoping review", Journal of Communication Disorders, Vol. 103, 2023, pp. 1-11.

[9] I. F. Míguez, P. F. Lamas, T. M. F. Caramés, "Design, Implementation, and Practical Evaluation of a Voice Recognition Based IoT Home Automation

System for Low- Resource Languages and Resource-Constrained Edge IoT Devices: A System for Galician and Mobile Opportunistic Scenarios", IEEE Access, Vol. 11, 2023, pp. 63623-63649.

[10] M. Talal et al. "Smart Home-based IoT for Real-time and Secure Remote Health Monitoring of Triage and Priority System using Body Sensors: Multi-driven Systematic Review", Journal of Medical Systems, Vol. 43, No. 3, 2019.

[11] J. Kho, N. Gillespie, M. K. Melinda, "A systematic scoping review of change management practices used for telemedicine service implementations", BMC Health Services Research, Vol. 20, No. 1, 2020.

[12] J.-P. O. Li et al. "Digital technology, tele-medicine and artificial intelligence in ophthalmology: A global perspective", Progress in Retinal and Eye Research, Vol. 82, 2021, p. 100900.

[13] A. H. Talal, U. Jaanimägi, K. Davis, J. Bailey, B. M. Bauer, A. Dharia, S. S. Dickerson, "Facilitating engagement of persons with opioid use disorder in treatment for hepatitis C virus infection via telemedicine: Stories of onsite case managers", Journal of Substance Abuse Treatment, Vol. 127, 2021.

[14] I. Kaur, T. Behl, L. Aleya, H. Rahman, A. Kumar, S. Arora, I. J. Bulbul, "Artificial intelligence as a fundamental tool in management of infectious diseases and its current implementation in COVID-19 pandemic", Environmental Science and Pollution Research, Vol. 28, 2021, pp. 40515-40532.

[15] P. Spachos, S. Gregori, M. J. Deen, "Voice Activated IoT Devices for Healthcare: Design Challenges and Emerging Applications", IEEE Transactions on Circuits and Systems II: Express Briefs, Vol. 69, No. 7, 2022, pp. 3101-3107.

[16] J. M. Alvarez, C. H. Barcelo, P. Lesso, P. C. Higuera, "Robust Detection of Audio-Cough Events using local Hu moments", IEEE Journal of Biomedical and Health Informatics, Vol. 23, No. 1, 2019, pp. 184-196.

[17] B. O. Erdene, J. L. Saver, "Automatic Acute Stroke Symptom Detection and Emergency Medical Systems Alerting by Mobile Health Technologies: A Review", Journal of Stroke and Cerebrovascular Diseases, Vol. 30, No. 7, 2021, p. 105826.

[18] M. B. Jamshidi et al. "Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment", IEEE Access, Vol. 8, 2020, pp. 109581-109595.

[19] E. B. Panganiban, A. C. Paglinawan, W. Y. Chung, G. L. S. Paa, "ECG diagnostic support system (EDSS): A deep learning neural network based classification system for detecting ECG abnormal rhythms from a low-powered wearable biosensors", Sensing and Bio-Sensing Research, Vol. 31, 2021, p. 100398.

[20] X. V. Wang, L. Wang, "A literature survey of the robotic technologies during the COVID-19 pandemic", Journal of Manufacturing Systems, Vol. 60, 2021, pp. 823-836.

[21] G. Dharmale, D. D. Patil, T. Ganguly, N. Shekapure, "Effective speech recognition for healthcare industry using the phonetic system", Journal of Autonomous Intelligence, Vol. 7, No. 5, 2024, pp. 1-14.

[22] K. Choutri, M. Lagha, S. Meshoul, M. Batouche, "A Multi-Lingual Speech Recognition-Based Framework to Human-Drone Interaction", Electronics, Vol. 11, No. 12, 2022, p. 1829.

[23] F. Kerwagen, K. F. Fuchs, M. Ullrich, A. Schulze, "Usability of a mHealth Solution using Speech Recognition for Point of care Diagnostic Management", Journal of Medical Systems, Vol. 47, 2023, pp. 1-10.

[24] R. Shinde, S. Patil, K. Kotecha, V. Potdar, G. Selvachandr, "Securing AI-based healthcare systems using blockchain technology: A state-of-the-art systematic literature review and future", Transactions on Emerging Telecommunications Technologies, Vol. 35, No. 1, 2024, p. e4884.

[25] M. Javaid, A. Haleem, R. P. Singh, R. Suman, R. Shanay, "Significance of machine learning in healthcare: Features, pillars and applications", International Journal of Intelligent Networks, Vol. 3, 2022, pp. 58-73.

# Antenna Array Diagnosis in the Presence of Unknown Mutual Coupling using Optimization Technique

Original Scientific Paper

**Oluwole John Famoriji***
Department of Electrical and Electronic Engineering Technology,
University of Johannesburg, P.O. Box 524, Auckland Park,
Johannesburg 2006, South Africa.
famoriji@mail.ustc.edu.cn

**Thokozani Shongwe**
Department of Electrical and Electronic Engineering Technology,
University of Johannesburg, P.O. Box 524, Auckland Park,
Johannesburg 2006, South Africa.
tshongwe@uj.ac.za

*Corresponding author

***Abstract*** *– Antenna array diagnosis is an important operation in communication systems, whenever element (s) failure in the array that worsening the projected radiation pattern occur. There are various diagnostic techniques found in literature that employ compressive sensing. Conversely, the techniques are based on easy formulation of array factor with no incorporation of mutual coupling existing between the radiators. This article shows how this deficiency lead to defective and bad diagnosis when there is presence of mutual coupling using port-level coupling matrix and average embedded antenna pattern. Furthermore, the element excitations are optimized to reduce the effect of mutual coherence of system measurement matrix, causing reduced measurements required for effective fault detection. Numerical simulation and experimental results demonstrate how the incorporation of mutual coupling generates an adequate and reliable array diagnosis, which are not found in literature. For instance, when fault number is set at 5, and SNR equals 10 dB, the smallest measurements needed for the diagnosis, which is the most effective diagnosis, are achieved when the optimized excitations are used. In conclusion, the implementation of the developed framework using measurement probe in space, shows enough results towards the practical deployment for antenna systems in wireless communication system.*

***Keywords****: Antenna arrays, array signal processing, array diagnosis, optimization techniques, SNR, mutual coupling*

*Received: Received: February 15, 2024; Received in revised form: March 30, 2024; Accepted: April 19, 2024*

## 1. INTRODUCTION

Antenna array diagnosis is an important research topic that finds application in civilian and military. Present and upcoming technology use larger number of elements in the active arrays. For instance, large number of elements is employed in massive MIMO (multiple input multiple output), full MIMO systems, and telecommunication devices that employ sophisticated arrays. Hence, the demand for a reliable antenna array diagnosis is an inevitable task to rectify the distorted radiation characteristics because of element (*s*) failure [1-6]. In addition, fault diagnosis is important in 5G wireless systems, where a very large number of elements are needed to satisfy the required reconfigurability and high radiation behaviour [2]. However, the more the number of elements in beamforming configuration, the more the probability of failed element (s). Therefore, an effective and highly reliable fault diagnosis method remains important, because replacement operation and manual dismantling take a lot of time, costly, and not feasible in satellite communications [2, 4]. Compressive sensing (*CS*) technique has been adapted to fault diagnosis in antenna arrays, because the number of failed radiators is assumed and always smaller than the total number of radiators in the antenna array.

Some array diagnosis algorithms, such as the backward transformation method (BTM) [5] and matrix method (MM) [6], efficiently identify the locations

and excitations of faulty elements using discrete Fourier transform (*DFT*) and matrix inversion techniques, respectively. However, these approaches are highly susceptible to noise and require a minimum number of sampling points. Specifically, for *MM*, the sampling points should equal or exceed the number of elements in the array to prevent ill-conditioned matrices during solving. Additionally, diagnostic methods employing intelligent optimization algorithms, such as genetic algorithms [7] and artificial neural networks [8], are computationally intensive.

Hence, the primary challenge in array diagnosis currently revolves around selecting an appropriate method to swiftly identify the faulty elements within the array. Additionally, this solution must exhibit low sensitivity to noise and provide flexibility in the number of sampling points utilized. The matrix pencil method (MPM) was originally introduced for estimating the parameters of complex exponential and attenuation exponential signals [8, 9]. However, a drawback of MPM is its limitation in handling the continuous distribution of synthesised element locations [7-10], rendering it unsuitable for array synthesis featuring elements positioned at fixed grid coordinates.

Furthermore, there are different array diagnostic methods that employed CS [3], [8-18], even with validation with experiments [18]. The array diagnosis is demonstrated in most approaches in literature employing recovery sparse solution from small number of measurements to show the situation (healthy of faulty) of elements. Some methods advised array diagnosis employing measured data taken from a point with different excitations [18], [19]. Conversely, each technique uses easy array factor dependent far-field model. Simplicity is offered, but non-ideal negligibility is contentious towards array diagnosis in practical sense, specifically when inter-element spacing is in smaller. Some recent literature modelled multipath channel to the fixed probe when there are faults [20], [21], but there is no work, to the best of authors' knowledge, that considers the mutual coupling (MC) impacts [22] in fault identification with the employment of a fixed receiver probe for measurements.

This article demonstrates why MC should be considered in fault diagnosis, and if not accounted for causes poor diagnosis. It is demonstrated how the proposed fault diagnosis method that employs a fixed probe and excitations optimization achieves optimal performance simply even while MC is considered. Furthermore, numerical experiment is demonstrated and the results involved AEAP (average embedded antenna pattern) and PLCM (port-level coupling matrix) methods; implying the proposed method provides effective and reliable array diagnosis. In addition, two MC modeling methods are presented in this article, they can be applied based on array patterns and available data about the antenna array to the user.

## 2. SYSTEM MODEL

This section provides the analysis of the faulty array at far-field, and the proposed fault detection approach.

### 2.1. FAULTY ANTENNA ARRAY AT FAR FIELD

The Based on the linearity feature of Maxwell's equations; that EM field originated from antenna array is formulated as a linear excitations juxtaposition of elements in the array, $E(r)=\sum_{j=1}^{N} \alpha_j (r) y_j$, here, $E(r)$ denotes the EM wave at point $r$, while $y_j$ are the excitations, $\alpha_j(r)$ are the resulted combination of the coefficients that consist the information regarding the EM surrounding of the array element and the measurement setup, $N$ is the number of elements in the array. Since the aim of this article is to conduct fault diagnosis with inter-element MC, then the computational steps that provide the system model can be outlined as follows.

#### 2.1.1. How is the Element' Fault Modeled?

On this issue, the excitation $y_j$ is replaced by $y_j \delta_j$, $\delta_j \in C$ show the state of fault of the element. For instance, $\delta_j=1$ means there is no fault, while $\delta_j=0$ means a deceased element.

#### 2.1.2. MC Modeling

In simple term, if coupling is neglected (from element pattern isolation technique), $\alpha_j$ is made up of parameters that determine the gain of the element, and phase due to distance between the measurement point and location of the element. Conversely, if the $\alpha_j$ parameters are calculated using active element patterns [23], then the effects of MC are incorporated fully compared to the $N$ full wave EM simulations of all measurement. Many more techniques are found in literature, which is involve AEAP [24] and PLCM methods [25-29]. Then, the measured field is expressed as [12]

$$E(r) = \sum_{j=1}^{N} y_j \alpha_j(r) \delta_j, \tag{1}$$

while the parameter expression of $\alpha_j$ is a function of the particular model employed for MC impacts.

### 2.2. PROPOSED FAULT DETECTION TECHNIQUE

By critical examination of Eq. (1), it can be observed that for a fixed measurement location, the $\alpha_j$ parameters are unchanged, hence the '*r*' argument is dropped. Consequently, a M measurements vector is built, $\tilde{x}$, and by relation to the excitation, we have $\tilde{x}=\sum_{i=1}^{N} y_i^{(i)} \alpha_i \delta_i$, where $y_i^{(i)}$ represents element i excitation for measurement *j*, which formulate the following model [12]

$$\tilde{x} = \begin{bmatrix} y^{(1)} \\ \vdots \\ y^{(M)} \end{bmatrix} \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_N \end{bmatrix} = \underbrace{Y diag(\alpha)}_{D} \delta, \tag{2}$$

where $y^{(j)} \in C^{1\times N}$ has the excitations of the element for the $j^{th}$ measurement, leading to total excitation matrix, $Y \in C^{M\times N}$, and $\delta \in C^{N\times 1}$ denotes the vector of the state of fault.

Generally, in fault diagnosis, it is usually assumed that reference measurements, such as $x^{(R)}$, are accessible for a particular state of elements, $\delta^{(R)}$, which corresponds to array without fault. Another assumption here is the sparsity of the number of failed radiators as to the reference, so, following the formulation of vector from the differential measurement, $x=x^{(R)}-\tilde{x}$, the resulting problem requiring solution for the diagnosis of faults becomes [12]

$$\min_{p}\|p\|_0, s.t. \|x - Dp\| < \mu, p = \delta^{(R)} - \delta, \qquad (3)$$

$\mu$ is a term that shares proportionality to standard deviation of the measurement noise.

The sparse recovery requires that $D$ (sensing matrix) exhibits low mutual coherence [27, 28]. For mutual coherence, we optimized the excitation matrix $Y$ via the approximation of a Grassmannian matrix by alternating method [27], which exhibits performance improvement in comparison to randomized $Y$ [29]. D is a diagonal matrix that multiplies $Y$, the mutual coherence of matrix $D$ equals that of the $Y$, hence the optimization of the mutual coherence heedless to the kind of MC model employed. The main knowledge acquired is the constancy of the linear coupled forward model for various excitations. Note, this is impossible for multiple points measurements. Using the standard techniques to transform the unconstrained nonconvex optimization problem [15, 16, 30] we have [12]

$$\min_{p}\|x - Dp\|_2^2 + \gamma \|p\|_z^z, 0 < p < 1, \qquad (4)$$

$\gamma$ is the hyperparameter (empirical). The problem is resolved via the iterative reweighted $l_1$ minimization [30-37] implemented using alternating direction method of multipliers (ADMM) [38].

## 3. NUMERICAL SIMULATIONS, RESULTS, AND DISCUSSION

This section demonstrate and verifies the effectiveness of the proposed diagnosis method. 100 elements constituted of WR90 open-ended waveguide array working at 27 GHz as depicted in Fig. 1 [5]. The radiation characteristics at no fault at $\phi=0^0$ is given in Fig. 2. The array aperture size is 24.88 x 12.18 mm$^2$ and the spacing between elements in both $x$ and $y$ axis is $\lambda$ and 0.5$\lambda$, respectively. The Ansys HFSS v.19 software was used for the simulations.

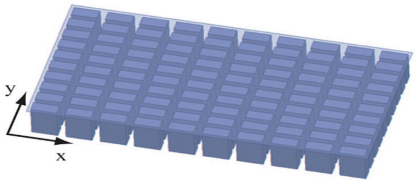



**Fig. 1.** Open ended waveguide array used for diagnostic demonstration [5].

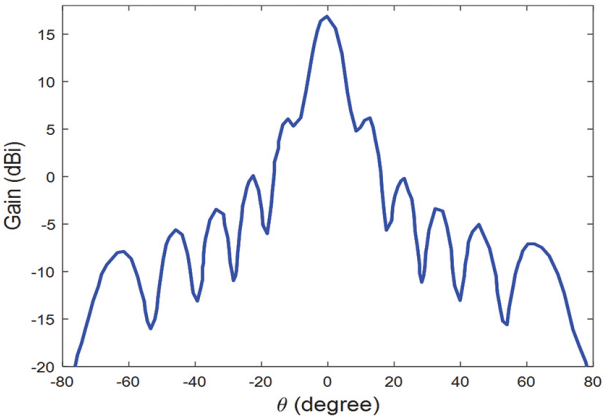



**Fig. 2.** Simulated radiation pattern of antenna array without element failure at principal plane $\phi=0^0$[5].

The waveguide is excited using a 50 Ω generator impedance and designed on the substrate ASTRA MT77 with relative permittivity $\varepsilon_r$ = 3 and loss factor tan $\delta$ = 0.0017 to have an impedance of 50 Ω. Also, the physical overlapping of elements at smaller value spacing between elements, such as $d$=0.45$\lambda$, and broadband radiation pattern used in wireless networks is ensured.

Next is to present the results of the diagnosis using the complex E-field ($E_y$) measured data obtained at fixed point via excitations optimization. The measurements of field of the faulty and healthy arrays (i.e. the forward models of $\alpha_j$ in Eqn. (1)) of isolated pattern method (using the array factor method [33]), average embedded pattern method (using Eqn. (11) of [23]), and coupling matrix method (using Eqn. (A9) of [25]) were simulated using Matlab Antenna Toolbox. It is important to state that the MC model employed in this article is not restricted, any MC model can be employed by appropriate computation of $\alpha_j$ based on the proposed scheme (as in Eqn.) (2).

The point of measurement is fixed at a spherical angular point ($\theta_0$, $\phi_0$)=(0,0), $r$=1000$\lambda$ (in z-direction). Using these measurements, the solution of the fault is gotten by the iterative reweighted $\ell$1 minimization. The hyperparameters $\gamma$ of Eqn. (4) is given as $\gamma$=0.25$\|D^H x\|_\infty$ and $z$=0.5 ($z$ quasi-norm) for each fault diagnosis result. The $\gamma$ is an empirical value obtained from a grid search from 0.1$\|D^H x\|\infty$ to $\|D^H x\|_\infty$. The upper limit $\|D^H x\|_\infty$ is explained in [27]. The phase and amplitude of excitations are set at [0, 2$\pi$], [0, 1], respectively. Both randomized and optimized excitations are quantified into six-bit phase and amplitude, while we obtain randomized excitations via multinomial probability distribution.

The solution that is recovered is designed into binary numbers using the threshold of the actual part by 1/2, i.e. for nth antenna $\delta_n$=0. Fault recovery is successful when the accurate faulty elements reconstruction and corresponding position is attained. It is important to state that thresholding action is unnecessary when dealing with non-binary faults. Rate of successful recovery (RSR) metrics is employed to present the results,

and the realizations percentage leading to successful recovery. The results are computed using 600 Monte-Carlo simulations with random fault positions. The findings originated from the proposed method are outlined below.

### 3.1. ENHANCEMENT IN FAULT DETECTION USING EXCITATIONS OPTIMIZATION

The proposed method incorporates the impact of MC. This implies there should be similar improvement in performance when excitations optimization with respect to random excitations is used. This is validated via performance analysis of fault diagnosis with coupling matrix model, without and with excitation matrix Y optimization at 10 dB signal-to-noise ratio (SNR). The excitation matrix Y optimization is confirmed to be useful for coupling matrix method as depicted in Fig. 3. When the average pattern method is used as the MC model, similar results were attained. This is an important result as all the merits of the optimized element excitation method continuously apply when the impacts of MC are considered. Similar findings were observed at higher SNRs, and there is an improvement generally in the RSR for a particular number of faults.

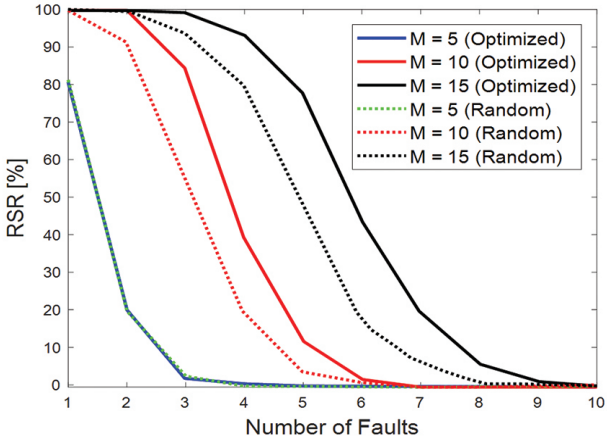
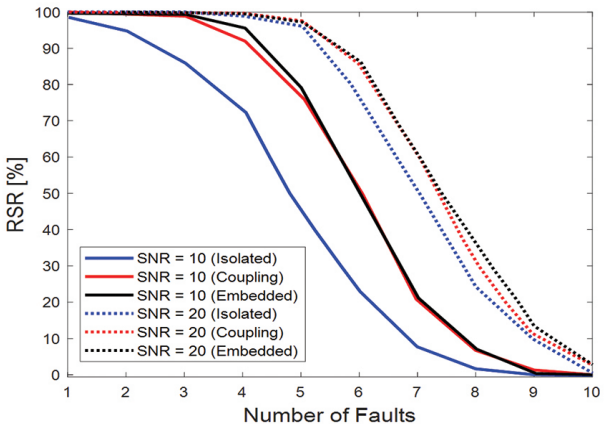


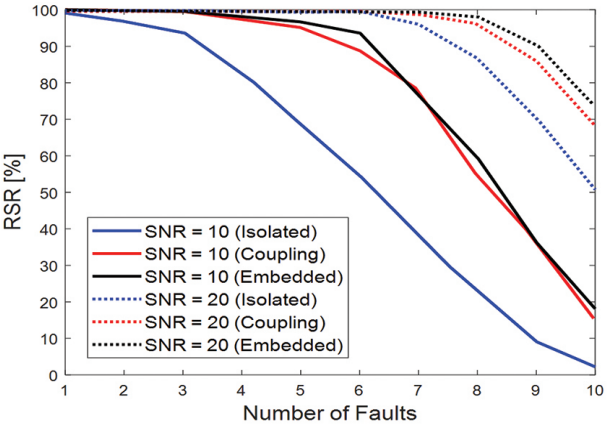

**Fig. 3.** Plot of RSR for coupling matrix method using iterative reweighted $\ell 1$ minimization, at $N$=100, $d$=0.45$\lambda$, and SNR=10 dB

### 3.2. ENHANCEMENT IN FAULT DETECTION WITH MUTUAL COUPLING

The RSR percentage versus the faults number via various coupling models with iterative reweighted $\ell 1$ minimization and excitations optimization for 10 dB and 20 dB, respectively, is depicted in Fig. 4. Fig. 4 are for two $M$, and it can be observed that evident reliability improvement in fault diagnosis is observed when MC impacts are considered in both measurement cases. For instance, when SNR is 10 dB with 15 measurements, $M$, and 3 faults, RSR of 100 is attained in the presence of MC, as against the 85 RSR without MC. In addition, when SNR is 10 dB with 20 measurements, $M$, and 4 faults, 98 RSR is attained in the presence of MC as against the 83 RSR without MC.




(a)




(b)

**Fig. 4.** Plot of RSR using iterative reweighted $\ell 1$ minimization, at $N$=100, $d$=0.45$\lambda$, and excitations optimization for 2 measurements, $M$, (a) $M$=15 (b) $M$=20.

Furthermore, the RSR accuracy difference when the MC is considered is higher than when low SNR measurements are used. For instance, comparing the performance of high and low SNRs in Fig. 4, the high SNR line shoot out more than that of low SNRs. For high SNR, an improved RSR is observed when the presence of MC is considered, even at bigger number of faults. For instance, when SNR is 20 dB, and measurements, $M$ is 20 including 10 faults; isolated pattern method attains 50 percent RSR, while coupling models attain about 71% RSR.

### 3.3. VARYING THE DISTANCE BETWEEN ELEMENTS: ANALYSIS

When the antenna elements are closer, the impact of MC is more [34], it is expected to have all the methods converge to close performance at higher spacing between elements. As this can be demonstrated in the results, it is valid based on least measurements $M$ needed correct diagnosis of faults. Tables I to IV depict the measurements number, $M$ required to achieve 90% RSR and 98% RSR when array diagnosis is conducted for various spacings between elements with the developed mod-

els and random/optimized excitations. Fault number is set at 5, and SNR equals 10 dB. The smallest measurements M needed for the diagnosis, which is the most effective diagnosis, are achieved when the optimized excitations are used (as per Table 1 & 3).

Generally, the AEAP models and coupling matrix with excitations optimization exhibit better options of finding fault in this scenario. Conversely, it becomes important to state that the coupling matrix method is of the assumption that the array impedance matrix is readily available for the user. When it is not available, the average embedded pattern model is recommended for practical application with large array aperture, because it requires only the response of the embedded pattern of the element at the center, easy to achieve.

**Table 1**. The required number of measurements to achieve 98% RSR with varying spacing $d$ and 5 number of faults for optimized excitation

| Spacing, $d$ | 0.45$\lambda$ | 0.95$\lambda$ | 1.45$\lambda$ | 1.95$\lambda$ |
|---|---|---|---|---|
| Embedded Pattern | 25 | 19 | 19 | 19 |
| Coupling Matrix | 23 | 20 | 21 | 19 |
| Isolated | 0 | 32 | 40 | 19 |

**Table 2.** The required number of measurements to achieve 98% RSR with varying spacing $d$ and 5 number of faults for random excitation

| Spacing, $d$ | 0.45$\lambda$ | 0.95$\lambda$ | 1.45$\lambda$ | 1.95$\lambda$ |
|---|---|---|---|---|
| Embedded Pattern | 28 | 25 | 24 | 25 |
| Coupling Matrix | 31 | 24 | 27 | 23 |
| Isolated | 0 | 35 | 48 | 25 |

**Table 3.** The required number of measurements to achieve 90% RSR with varying spacing $d$ and 5 number of faults for optimized excitation

| Spacing, $d$ | 0.45$\lambda$ | 0.95$\lambda$ | 1.45$\lambda$ | 1.95$\lambda$ |
|---|---|---|---|---|
| Embedded Pattern | 18 | 17 | 17 | 17 |
| Coupling Matrix | 18 | 17 | 17 | 17 |
| Isolated | 38 | 22 | 27 | 17 |

**Table 4.** The required number of measurements to achieve 90% RSR with varying spacing $d$ and 5 number of faults for random excitation

| Spacing, $d$ | 0.45$\lambda$ | 0.95$\lambda$ | 1.45$\lambda$ | 1.95$\lambda$ |
|---|---|---|---|---|
| Embedded Pattern | 22 | 20 | 20 | 19 |
| Coupling Matrix | 24 | 20 | 22 | 21 |
| Isolated | 45 | 25 | 31 | 20 |

## 4. CONCLUSION

In this paper, two methods of adding the impact of MC for effective and more reliable antenna array fault diagnosis via a notable CS method that employs fixed probe based measurements, and excitations optimization. Based on the accessible data about the array, and corresponding characteristics, both methods can be employed. For a fairly large array, AEAP method is an appropriate forward model, and coupling matrix method is appropriate for a particular antenna group. The superiority of the forward models, which incorporate the impact of MC, have been demonstrated and the shortcoming of a forward model without the MC influences incorporation has been presented. For instance, when fault number is set at 5, and SNR equals 10 dB, the smallest measurements needed for the diagnosis, which is the most effective diagnosis, are achieved when the optimized excitations are used. The proposed technique is verified and demonstrated to be highly correct, and reliable in fault finding in antenna arrays where the impacts of MC are cannot be ignored. Finally, the proposed technique is more practical and recommended for identification of faults in antenna arrays.

## 5. REFERENCES


[1] O. M. Bucci, M. D. Migliore, G. Panariello, P. Sgambato, "Accurate diagnosis of conformal arrays from near-field data using the matrix method", IEEE Transactions on Antennas and Propagation, Vol. 53, No. 3, 2005, pp. 1114-1120.

[2] A. Massa, P. Rocca, G. Oliveri, "Compressive sensing in electromagnetics—A review." IEEE Antennas and Propagation Magazine, Vol. 57, 2015, pp. 224-238.

[3] M. D. Migliore, "A compressed sensing approach for array diagnosis from a small set of near-field measurements", IEEE Transactions on Antennas and Propagation, Vol. 59, 2011, pp. 2127-2133.

[4] A. Massa, M. Bertolli, G. Gottardi, A. Hannan, D. Marcantonio, G. Oliver, A. Polo, F. Robol, P. Rocca, F. Viani, "Compressive sensing as applied to antenna arrays: Synthesis, diagnosis, and processing", Proceedings of the IEEE International Symposium on Circuits and Systems, Florence, Italy, 27-30 May 2018, pp. 27-30.

[5] J. J. Lee, E. M. Ferren, D. P. Woollen, K. M. Lee, "Near-field probe used as a diagnostic tool to locate defective elements in an array antenna", IEEE Transactions on Antennas and Propagation, Vol. 36, No. 6, 1988, pp. 884-889.

[6] O. M. Bucci, M. D. Migliore, G. Panariello, P. Sgambato, "Accurate diagnosis of conformal arrays from near-field data using the matrix method", IEEE Transactions on Antennas and Propagation, Vol. 53, No. 3, 2005, pp. 1114-1120.

[7] R. Iglesias, F. Ares, M. Fernandez-Delgado, J. A. Rodriguez, J. Bregains, S. Barro, "Element failure de-

tection in linear antenna arrays using case-based reasoning", IEEE Antennas and Propagation Magazine, Vol. 50, No. 4, 2008, pp. 198-204.

[8] X. Wang, K. Konno, Q. Chen, "Diagnosis of array antennas based on phaseless near-field data using artificial neural network", IEEE Antennas and Propagation Magazine, Vol. 69, No. 7, 2021, pp. 3840-3848.

[9] O. J. Famoriji, T. Shongwe, "An effective antenna array diagnosis method via multivalued neural network inverse modeling approach", Advanced Electromagnetics, Vol. 10, No. 3, 2021, pp. 58-70.

[10] O. J. Famoriji, T. Shongwe, "A Recovery Performance Study of Compressive Sensing Methods on Antenna Array Diagnosis from Near-Field Measurement Data", Applied Computational Electromagnetics Journal, Vol. 36, No. 8, 2021, pp. 1-7.

[11] A. F. Morabito, A. R. Lagana, T. Isernia, "Optimizing power transmission in given target areas in the presence of protection requirements", IEEE Antennas and Wireless Propagation Letters, Vol. 14, 2014, pp. 44-47.

[12] A. Fadamiro, A. Semomhe, O. J. Famoriji, F. Lin, "A multiple element calibration algorithm for active phased array antenna", IEEE Journal of Multiscale and Multiphysics, Vol. 4, 2019, pp. 163-170.

[13] M. D. Migliore, "Array diagnosis from far-field data using the theory of random partial Fourier matrices", IEEE Antennas and Wireless Propagation Letters, Vol. 12, 2013, pp. 745-748.

[14] T. Ince, G. Ogucu, "Array failure diagnosis using nonconvex compressed sensing", IEEE Antennas and Wireless Propagation Letters, Vol. 15, 2015, pp. 992-995.

[15] O. Famoriji, Z. Zhang, A. Fadamiro, Md. Ali, R. Zakariyya, F. Lin, "Planar array diagnostic tool for millimeter wave communication systems", Electronics, Vol. 7, No. 12, 2018, pp. 1-23.

[16] B. Fuchs, L. Le Coq, M. D. Migliore, "Fast antenna array diagnosis from a small number of far-field measurements", IEEE Transactions on Antennas and Propagation, Vol. 64, No. 6, 2016, pp. 2227-2235.

[17] O. J. Famoriji, A. Fadamiro, Z. Khan, Z. Zhang, F. Lin, "Active antenna array diagnosis from far-field measurements", Proceedings of the International Conference on Integrated Circuits, Technologies and Applications, Beijing, China, 21-23 November 2018, pp. 1-4.

[18] C. Xiong, G. Xiao, Y. Hou, M. Hameed, "A compressed sensing based element failure diagnosis method for phased array antenna during beam steering", IEEE Antennas and Wireless Propagation Letters, Vol. 18, No. 9, 2019, pp. 1756-1760.

[19] W. Li, W. Deng, Q. Yang, X. Zhang, J. Zhang, M. D. Migliore, "A Hybrid Method for Array Diagnosis Using Random Perturbation-Convex Local Minimizer", Proceedings of the IEEE Radar Conference, Boston, MA, USA, 22-26 April 2019, pp. 1-5.

[20] S. Costanzo, A. Borgia, G. Di Massa, D. Pinchera, M. D. Migliore, "Radar array diagnosis from undersampled data using a compressed sensing/sparse recovery technique", Journal of Electrical and Computer Engineering, Vol. 2013, 2013, pp. 324-331.

[21] R. Sun, W. Wang, L. Chen, G. Wei, W. Zhang, "Blind diagnosis for millimeter-wave large-scale antenna systems", IEEE Communications Letters, Vol. 25, No. 7, 2021, pp. 2390-2394.

[22] G. Medina, A. S. Jida, S. Pulipali, R. Talwar, N. A. J, T. Y. Al-Naffouri, A. Madanayake, M. E. Eltayeb, "Millimeter-Wave Antenna Array Diagnosis with Partial Channel State Information", Proceedings of the ICC IEEE International Conference on Communications, Montreal, QC, Canada, 14-23 June 2021, pp. 1-5.

[23] I. Gupta, A. Ksienski, "Effect of mutual coupling on the performance of adaptive arrays", IEEE Transactions on Antennas and Propagation, Vol. 31, No. 5, 1983, pp. 785-791.

[24] Y. Zhang, H. Zhao, "Failure diagnosis of a uniform linear array in the presence of mutual coupling", IEEE Antennas and Wireless Propagation Letters, Vol. 14, 2015, pp. 1010-1013.

[25] A. Ludwig, "Mutual coupling, gain and directivity of an array of two identical antennas", IEEE Transactions on Antennas and Propagation, Vol. 24, No. 6, 1976, pp. 837-841.

[26] D. Kelley, W. Stutzman, "Array antenna pattern modeling methods that include mutual coupling

effects", IEEE Transactions on Antennas and Propagation, Vol. 41, No. 12, 1993, pp. 1625-1632.

[27] T. Su, H. Ling, "On modeling mutual coupling in antenna arrays using the coupling matrix", Microwave and Optical Technology Letters, Vol. 28, 2001, pp. 231-237.

[28] B. Clerckx, C. Craeye, D. Vanhoenacker-Janvier, C. Oestges, "Impact of Antenna Coupling on 2 x 2 MIMO Communications", IEEE Transactions on Vehicular Technology, Vol. 56, No. 3, 2007, pp. 1009-1018.

[29] C. Craeye, D. Gonzalez-Ovejero, "A review on array mutual coupling analysis", Radio Science - RADIO SCI, Vol. 46, 2011, pp. 23-33.

[30] M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing", $1_{st}$ Edition, Springer, 2010.

[31] E. Candes, J. Romberg, "Sparsity and incoherence in compressive sampling", Inverse Problems, Vol. 23, No. 3, 2007, p. 969.

[32] E. J. Candes, T. Tao, "Near-optimal signal recovery from random projections: Universal encoding strategies?", IEEE Transactions on Information Theory, Vol. 52, No. 12, 2006, pp. 5406-5425.

[33] Q. Lyu, Z. Lin, Y. She, C. Zhang, "A comparison of typical p minimization algorithms", Neurocomputing, Vol. 119, 2013, pp. 413-424.

[34] S. Foucart, M.-J. Lai, "Sparsest solutions of underdetermined linear systems via q-minimization for lq-minimization for $q \leqslant 1$", Applied and Computational Harmonic Analysis, Vol. 26, 2009, No. 3, pp. 395-407.

[35] S. Boyd et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers", Foundations and Trends in Machine learning, Vol. 3, No. 1, 2011, pp. 1-122.

[36] R. C. Hansen, "Phased Array Antennas", John Wiley & Sons, 2009.

[37] O. J. Famoriji, T. Shongwe, "Electromagnetic machine learning for estimation and mitigation of mutual coupling in strongly coupled arrays", ICT Express, Vol. 9, No. 1, 2023, pp. 8-15.

[38] O. J. Famoriji, T. Shongwe, "Antenna array diagnosis via smart sensing of electromagnetics with learnable data acquisition and processing", Proceedings of the AEIT International Annual Conference, Rome, Italy, 3-5 October 2023.

# Performance Investigation of 17 Level Reduced Switch Count Multilevel Inverter

Original Scientific Paper

**Murugesan Manivel***
Department of Electrical and Electronics Engineering
Karpagam Institute of Technology, Coimbatore, Tamil Nadu
murugesan.kec@gmail.com

**Sivaranjani Subramani**
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu
sivaranjanis@skcet.ac.in

*Corresponding author

**Lakshmanan Palani**
Department of Electrical and Electronics Engineering
Narasaraopeta Engineering College, Andhra Pradesh
lakchandp@gmail.com

**Kesavan Tamilselvan**
Department of Electrical and Electronics Engineering
Easwari Engineering College, Tamil Nadu
t.kesavan87@gmail.com

***Abstract*** *– The primary objective of this paper is to introduce a 17-level multilevel inverter with only eight power switches and three diodes that can be suitable for electric vehicle applications. This setup utilizes three distinct unequal DC sources to create the 17 levels of output voltage waveforms. The modes of operations of the proposed topology have been discussed in detail. The calculation of conduction losses, switching losses, efficiency, total standing voltage, and cost function per level for the suggested inverter has been elaborated. Due to more semiconductor power switches, diodes, capacitors, driver circuits, and DC sources in typical inverters, switching losses, costs, and harmonic distortion are increased. The nearest-level control technique has been utilized to control the switching elements of the recommended configuration. The performance comparison of various multilevel inverter topologies has also been discussed. The suggested multilevel inverter provides a higher efficiency of 98.60%, cost function of 3.6 for a weight coefficient of 1, improved power quality, and higher reliability. A reduction in the power switches significantly reduces the convolution of switching circuitry. The total harmonic distortion produced by this inverter is 3.49%, which comes under the IEEE standard of 5%.*

***Keywords***: *Multilevel Inverter, Electric Vehicle, Nearest Level Control, Total Harmonic Distortion, Pulse Width Modulation*

*Received: November 25, 2023; Received in revised form: March 22, 2024; Accepted: April 9, 2024*

## 1. INTRODUCTION

Multilevel inverters are generally used in smart grids, battery-powered electric vehicles (EVs), and FACTS devices because of attributes including higher power quality, enhanced adaptability, and lessened dv/dt stress. Due to their capacity to generate high terminal output voltages employing low and medium voltage components, they are highly suited for the aforementioned purposes [1, 2]. Multilevel inverters (MLIs) are generally classified as one of the three types: diode clamped or neutral point clamped multilevel inverters (NPCMLI or DCMLI), flying capacitor multilevel inverters (FCMLI), and cascaded H-bridge multilevel inverters (CHBMLI). Diodes are mainly used in this inverter for clamping voltage, so it is called DCMLI, which demands more diodes and power components when the output waveform levels rise, which complicates circuit control. By using the redundancy in switching states offered by clamping capacitors, the voltage balancing problem caused by diodes in DCMLI can be resolved in FCMLI. Capacitors are used as clamping components, so it is called FCMLI; however, higher passive circuit components would decrease circuit reliability [3-6]. Compared to the other two MLIs, a MLI that consists of serially connected or cascaded H-bridges, called CHBMLI, is enormously modular and easier to control. Furthermore, it eliminates the necessity for voltage balancing circuits by using devoid DC sources. However, in the typical CHBMLI, each additional DC source enhances the need for four power electronic switches [7]. A unique multilevel inverter arrangement has been pre-

sented with n+5 power switches and 'n' independent sources for high levels. The standing voltage across the power switches in the polarity changer increases significantly as a result of the elevated levels in the terminal voltage waveform. Power switches with the highest voltage and the largest heat sinks must be used [8, 9]. The overwhelming majority of structures that rely on autonomous DC sources for producing levels in output waveforms are based on the idea of using switched DC input voltages to add level and polarity flipping using H-Bridge. The number of elements rises since the total blocking output across the H-Bride is four times the DC input voltage [10-12]. A framework that incorporates a four-quadrant power switch design, enabling both balanced and unbalanced switching circuit operation. The configuration features a large number of power switches, which increase losses [13]. MLIs have gained popularity in a variety of low- to high-voltage, high-power applications because of their unique features. Improved power quality, less electromagnetic interference, decreased voltage stress, and semiconductor device loss are a few merits of MLIs. Three key issues with MLIs are high precision control, an increase in power semiconductor switches and capacitors, and voltage balancing [14-17]. MLIs have not yet been extensively employed in low-power electric transportation, even though they have been used in traction drives in a number of experiments. These arrangements are seen in low-power automobiles like electric buses and passenger EVs, as well as high-power electric trains and ships. They use a standard two-level inverter because of its low DC-link voltage, simplicity in design, and convenience of use [18-22]. Eleven switches, three capacitors, three diodes, and one source of seven-level SCMLI are implemented. The capacitors in this configuration are self-balanced. However, the overall circuitry complexity is higher due to more switches, diodes, and capacitors [23]. Six switches, two diodes, and two sources of 5 level MLI have been introduced for renewable energy applications [24]. Seven switches, three diodes, and three sources of 15 level MLI have been discussed [25]. Ten switches, three diodes, and three capacitors with 13 level MLI have been introduced [26]. Eight switches, two sources 11 level inverter have been employed [27].

This work provides a new reduced switch count multilevel inverter (NRSCMLI) that combines a multi-conversion cell and H-bridge. The proposed MLI is very compact and requires only 8 switches, 3 diodes, and 3 unequal DC voltage sources. The nearest level control (NLC) technique is designed for this inverter. A comparative study has been done based on the number of switches, diodes, driver circuits, capacitors, and DC sources, total blocking voltage per unit, and cost function per level for 17 levels of various existing MLIs. The suggested configuration's key design feature has fewer components for seventeen levels when compared to conventional MLIs.

## 2. CONVENTIONAL MULTILEVEL INVERTERS

NPCMLI, or DCMLI, has been used to create a multilevel output voltage, which is shown in Fig. 1a. To achieve the needed output levels, power switches are placed in series, with diodes functioning as clamping elements. NPCMLI has minimised expenses and is more adjustable since it uses diodes for clamping. To accomplish a waveform with'm' levels in the output voltage or current, (m-1) DC-link capacitors are required. Each phase desires 2(m−1) power semi-conductor switches and (m−1) (m−2) diodes. There is an unbalanced stress on switches throughout the operation of NPCMLI. It requires more diodes and power switches. Moreover, this NPCMLI elevates the inverter power while needing one central and one high DC-link voltage [2, 3]. The next design, FCMLI, is seen in Figu. 1b. It employs flying capacitors in place of clamped diodes and delivers a clean output waveform while being safer and easier to operate. Because capacitors are used as clamping elements in the FCMLI, it is also simple to regulate. Every switch in this MLI shares the same amount of power dissipation.

Higher switching frequencies and balanced clamping capacitor voltages at high DC-link voltages are essential for the FCMLI. Since capacitors are more costly than diodes, this inverter is more complex and expensive than NPCMLI [4]. Alternatively, CHBMLI topology is one type of MLI that involves multiple symmetrical DC sources with H-bridges, as shown in Fig. 1c. The same circuit with unequal sources is specified as a hybrid MLI, as shown in Fig. 1d. Each H-bridge comprises four power switches [15]. Fig. 1e and Fig. 1f display the new cascaded MLI, which encompasses two stages; the first stage is two switches with equal or unequal DC sources, which can be prolonged effortlessly to create more levels. The second stage is the H-bridge, which converts the stepped direct current waveform into an alternate current waveform. Figure 1g shows a circuit diagram of a modified reduced switch count MLI, which consists of a multi-conversion cell and an H-bridge. A multi-conversion cell has only 4 switches and 4 power diodes with 4 equal DC sources. A multi-conversion cell is serially connected to an H-bridge for AC conversion. Figure 1h shows the circuit diagram of a modified hybrid multilevel inverter with unequal DC sources [25].

## 3. NEW REDUCED SWITCH COUNT MULTILEVEL INVERTER

The layout of the recommended structure, comprising four power switches and three unequal voltage sources ($1V_{DC}$, $2V_{DC}$, and $5V_{DC}$) is illustrated in Fig. 2. The unequal voltage sources are created by 1:1, 1:2, 1:5 transformers and DC-DC converters. This proposed structure produces a unidirectional stepped 17 level output voltage or current waveform. It is further connected to the H bridge circuit to create an alternate output waveform. Fig. 3 displays the overall construction of the proposed NRSCMLI. The significant benefit of the recommended MLI is that it has 17 levels while using only 8 power switches. The operation of the proposed multilevel is;

Mode 1: Devices $S_5$, $S_6$, $S_7$ & $S_8$ conduct, and it creates zero output voltage.

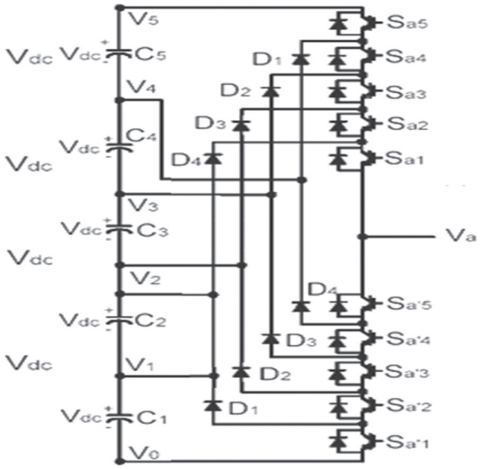
Mode 2: Switches $S_1$, $S_5$, $S_8$, $D_2$ & $D_3$ conduct and generate $1V_{dc}$=28V, it supplies to the load.

Mode 3: Switching elements $S_2$, $S_5$, $S_8$, $D_1$ and $D_3$ conduct and producing $2V_{dc}$=56V, it supplies the load.

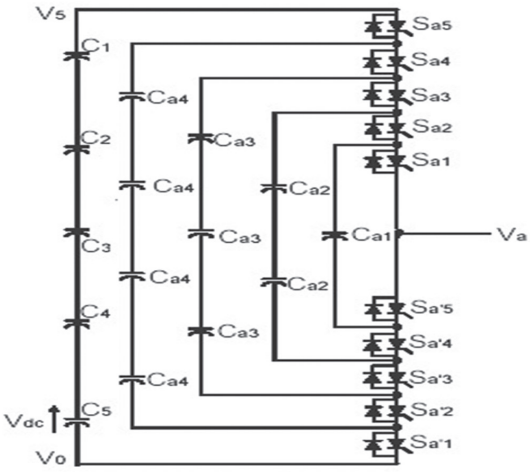
Mode 4: Power switches S1, S2, S5, S8, D3 & D4 conduct and generate $1V_{dc}+2V_{dc}=3V_{dc}$.

Mode 5: In this operating mode, switches S4 and D1 conduct and produce $5V_{dc}-1V_{dc}=4V_{dc}$=112V to load.

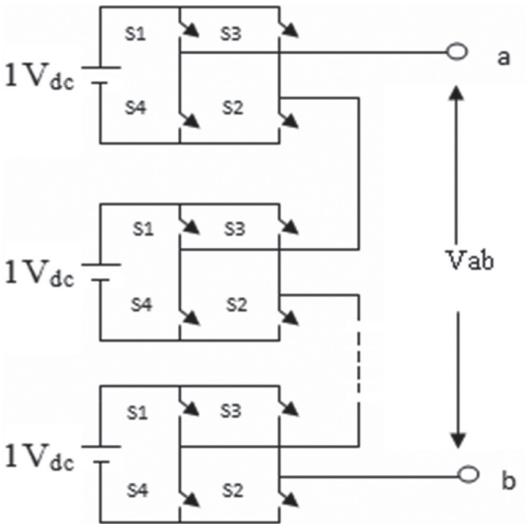
Mode 6: Device $S_3$ only conducts along with diodes $D_1$ and $D_2$ and produces $1V_{dc}+5V_{dc}=5V_{dc}$=140V.



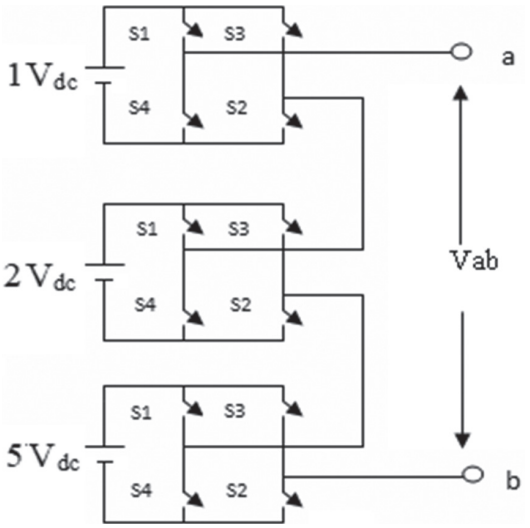

(a)



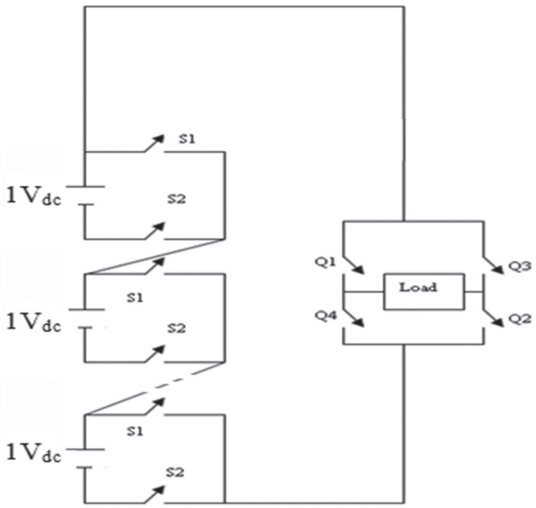

(b)



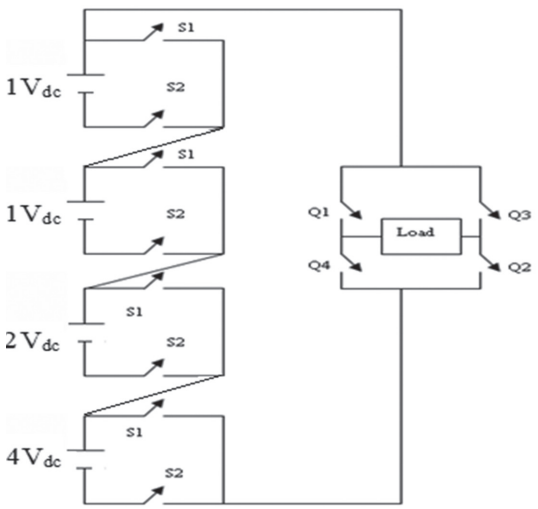

(c)




(d)




(e)




(f)

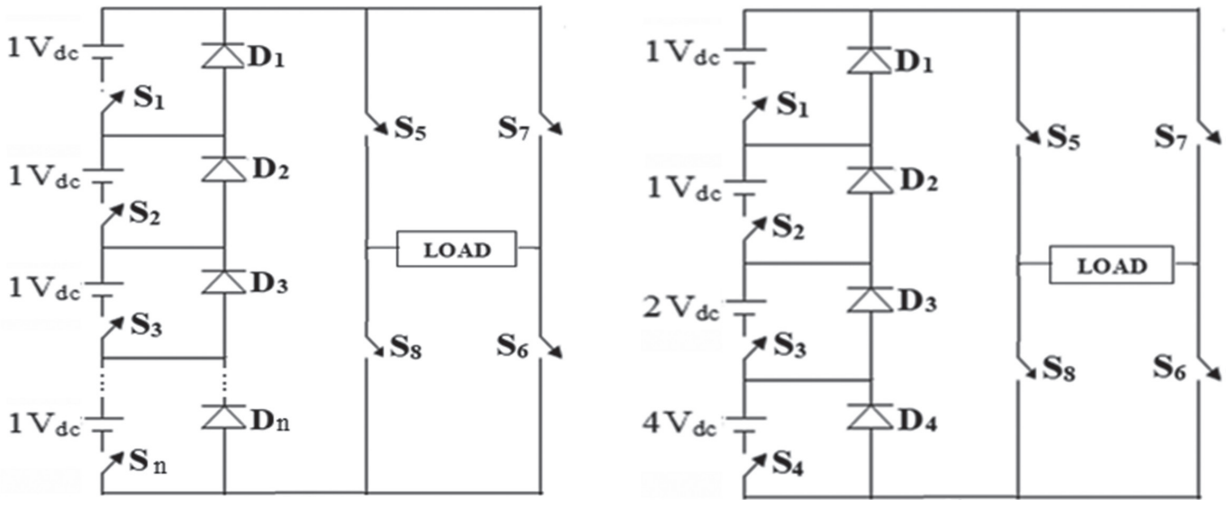



**Fig. 1.** a) Diode Clamped MLI, b) Flying Capacitor MLI, c) Cascaded H bridge MLI, d) Hybrid MLI, e) New Cascaded MLI, f) New Hybrid MLI, g) Modified Cascaded MLI, h) Modified Hybrid MLI. [2-4, 15, 25]

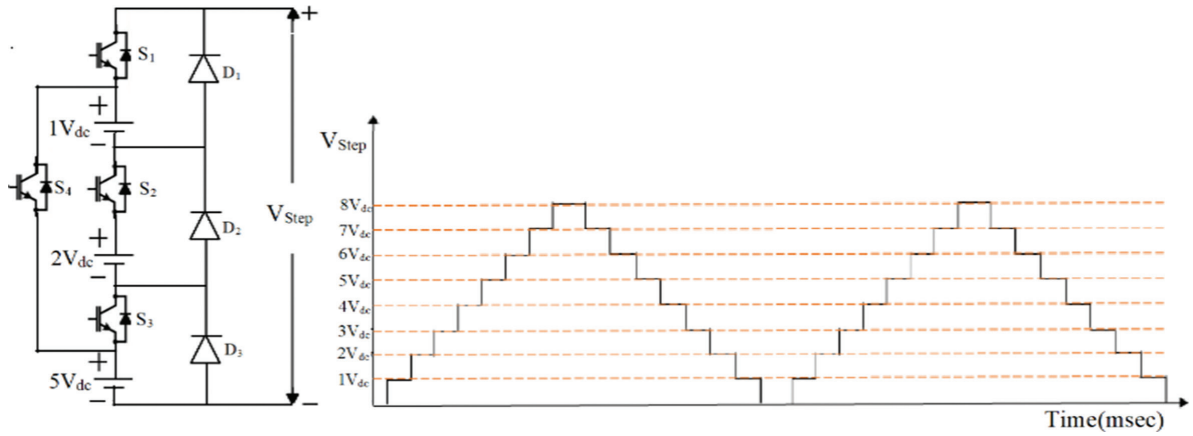



**Fig. 2.** Multi-Conversion Cell for New Reduced Switched Count multilevel inverter

Mode 7: Switches $S_{1'}$ $_{S3}$, and $D_2$ conduct and generates $1V_{dc}+5V_{dc}=6V_{dc}=168V$.

Mode 8: Elements $S_2$, $S_3$, and $D_1$ conduct and give $2V_{dc}+5V_{dc}=7V_{dc}=196V$ to the load.

Mode 9: Switches $S_1$, $S_2$, $S_3$ conduct, and the voltage across the terminal is $1V_{dc}+2V_{dc}+5V_{dc}=8V_{dc}=224V$.

Mode 10: Switches $S_2$, $S_3$, and $D_1$ conduct and create $2V_{dc}+5V_{dc}=7V_{dc}=196V$, supplies to the load.

Mode 11: Switches $S_1$, $S_3$ and $D_2$ conduct, and the load output voltage is $1V_{dc}+5V_{dc}=6V_{dc}=168V$.

Mode 12: Switches $S_3$, $D_1$ and $D_2$ conduct, and voltage across the load is $5V_{dc}=140V$.

Mode 13: Switches $S_4$ and $D_1$ conduct, and the output voltage is $5V_{dc}$-$1V_{dc}=4V_{dc}=112V$.

Mode 14: Power switches $S_1$, $S_2$, $S_5$, & $S_8$ and $D_3$ conduct, and the developed load voltage is $1V_{dc}+2V_{dc}=3V_{dc}=84V$.

Mode 15: Devices $S_2$, $S_5$, $S_8$, $D_1$ and $D_3$ conduct, and the developed output voltage is $2V_{dc}=56V$.

Mode 16: Switches $S_1$, $S_5$, $S_8$, $D_2$ and $D_3$ conduct, and the circuit generates $1V_{dc}=28V$ for the load.

Mode 17: In this mode, $S_5$, $S_6$, $S_7$ & $S_8$ conduct, and it produces zero voltage. Mode 1 to 17 give a 17 level output voltage for a positive half cycle. The same modes are repeated to generate a negative half cycle also but switches $S_6$ and $S_7$ conduct instead of switches $S_5$ and $S_8$. Fig. 6 and Table 1 display the modes of operation of the suggested MLI.

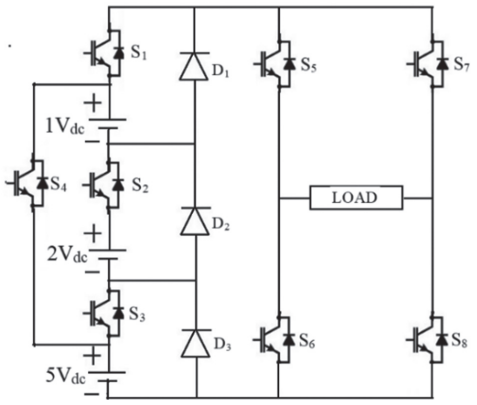



**Fig. 3.** Proposed New Reduced Switched Count multilevel inverter

## 4. ESTIMATION OF LOSS AND EFFICIENCY

This part describes the theoretical calculation of losses used to determine the efficiency of the recommended inverter. Conduction and switching losses are the most significant losses, and their computation relies on the assumption that the load is completely resistive and that the voltage at the inverter's output terminal resembles a staircase output waveform. When MOSFET power switches in a MLI are activated and conduct electricity, conduction loss occurs. The overall loss during conduction in the suggested inverter architecture is calculated for each MOSFET power switch. The same has been calculated for the polarity transformation unit separately. Power switches are employed as a polarity conversion unit in the proposed inverter, and load current in a single phase is observed with respect to neutral. In this case, the loss due to the switch conduction period during the basic cycle quarter is calculated.

$$P_{CON} = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} {I_R}^2(t)\, R_{ON,}Tdt \tag{1}$$

The transistor's on-state resistance and load current are symbolized as $R_{on}T$ and $I_R(t)$ respectively. By engaging a high number of auxiliary units, the projected system delivers 17 levels of output voltage and current in the load side. As a result, the load current is stated as

$$I_R = I_{P,}sinwt \tag{2}$$

The average loss of a single-phase system during conduction is estimated using (1) and (2) as,

$$P_{CON1} = \frac{4}{\pi}\int_0^{\frac{\pi}{2}} {I_{P,}}^2 sin^2\, wtR_{ON,}Tdt \tag{3}$$

The multi-conversion unit's MOSFET switches are active during the basic cycle. The losses of multi-conversion for the '$\pi$' cycle period is as

$$P_{CON2} = \frac{1}{\pi}\int_0^{\pi} {I_{P,}}^2 sin^2\, wtR_{ON,}Tdt \tag{4}$$

The loss during conduction is computed using the following equation for the single multi-conversion unit in the proposed MLI.

$$P_{CON,TOTAL} = P_{CON1} + P_{CON2} \tag{5}$$

In MOSFET power switches, switching loss occurs by the overlap of voltage and current during the ON/OFF or vice versa transition. The following formula is used to estimate the energy loss of MOSFET power switches during their on and off times.

$$E_{ON} = \frac{V_{ON}\times I}{6}\, T_{ON} \tag{6}$$

where $V_{ON}$ - on-state voltage, I - current after switching on, $T_{ON}$ - turn on time.

$$E_{OFF} = \frac{V_{OFF}\times I}{6}\, T_{OFF} \tag{7}$$

$V_{OFF}$ - off-state voltage of MOSFET, $I_{OFF}$ - current passing through the MOSFET before it turns on, and $T_{OFF}$ - turn-off time. Using equations (6) and (7), the switching loss in the predicted MLI is calculated separately for each power switch in each unit. After half of the fundamental cycle, the MOSFET power switches in the polarity conversion unit are turned on and off, and the switching loss is calculated as follows.

$$P_{SW1} = 2\times f\times (E_{ON} + E_{OFF}) = \frac{1}{3}\times f\times I\times (V_{ON}T_{ON} + V_{OFF}T_{OFF}) \tag{8}$$

The fundamental switching frequency is f. Likewise, the switching loss of a multi-conversion unit is considered as throughout the half time. It is specified by,

$$P_{SW2} = 2\times f\times (E_{ON} + E_{OFF})f\times I(V_{ON}T_{ON}V_{OFF}T_{OFF}) \tag{9}$$

For a full cycle, the overall switching loss can be inferred as

$$\mathrm{P_{SW,TOTAL}} = \mathrm{f}\times(\mathrm{P_{SW1}} + \mathrm{P_{SW2}}) \tag{10}$$

The power loss of the NRSCMLI is computed by using (11)

$$P_{LOSS,,TOTAL} = P_{CON,T} + P_{SW,T} \tag{11}$$

The efficiency is calculated by using the following formulae

$$Efficiency = \frac{P_{OUT}}{P_{out}+P_{LOSS,T}}\times 100\% \tag{12}$$

The calculations show that the efficiency of the 17-level NRSCMLI is 98.60%, its conduction losses are 10.24W, its switching losses are 0.36W, its input power is 756.7W, its output power is 746.1W.

## 5. NEAREST LEVEL CONTROL(NLC) FOR 17 LEVEL NEW REDUCED SWITCH COUNT MULTILEVEL INVERTER

NLC is a common approach used in modular multilevel converters. Low switching frequency control technology known as NLC or the round approach is adaptable and simple to use. The semiconductor switches of the converter are activated by the appropriate gating signals produced by the NLC approach. Take the basic sinusoidal function with unit magnitude multiplied by the modulation index as a starting point. Allow the resulting signal to pass through the inverter's total positive levels (N) after that. Employ the round function next to obtain the necessary levels. The round function serves effectively to decrease switching losses by causing just one commutation to occur between two voltage steps, as shown in Fig. 4. Considering a Vref, the modulation index 'm' the nearest voltage level that can be estimated as follows:

$$V_{ref}=\mu\left(\frac{L-1}{2}\right)Vsin(wt) = V_m\sin(wt)$$
$$m = \frac{V_m}{\left(\frac{L-1}{2}\right)V},\ V_m = Vround\left(\frac{V_{ref}}{V}\right) \tag{13}$$

$V_{ref}$ - reference sinusoidal voltage
$L$ - number of levels
$m$ - modulation index
$V_m$ – peak ac voltage
$V_{dc}$ – voltage difference between two levels

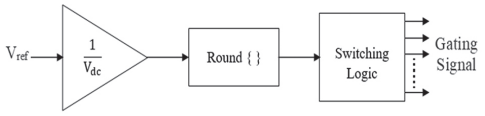



**Fig. 4.** NLC control method

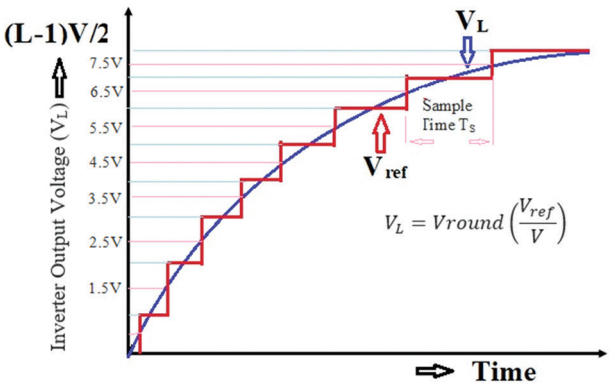



**Fig. 5.** NLC technique synthesis for 17 level

Fig. 5 indicates the identification of the nearest voltage level and the principle of the NLC method for the projected 17 level inverter.

## 6. RESULTS AND DISCUSSIONS

Fig. 7 shows the simulation diagram of NRSCMLI, Fig. 8 depicts the switching pulses of NRSCMLI; and Fig. 9 shows the MATLAB simulated output voltage waveform of the 17-level proposed MLI for a resistive load of 10 kW. The magnitude of the voltage is 224V, and the current is 44 A. Fig. 10 shows the output voltage and current waveform for a resistive (R) and inductive (L) load of 11 KVAR. The magnitude of the current is 50 A and the nature of the current waveform is sinusoidal. Fig. 11 shows the output voltage and current waveform for parallel resistive and inductive loads. From this figure, it clearly indicates that the nature of the output current waveform is sinusoidal for R and RL loads. From the Fast Fourier Transform (FFT) analysis window, it is understood that while the levels are raised, harmonics and THD are minimized. Fig. 12 shows that for the 17-level inverter, the voltage THD value is 3.49%. A simulation result of NRSCMLI is validated by the hardware prototype. Four MOSFETs (IRF250) are connected to form a multi conversion cell, and it is connected to an H-bridge, which comprises four MOSFETs. 3 unequal DC sources are attained by using specially designed 3 separate transformers in the ratios of 1:1, 1:2 and 1:5. Input for the all transformer is 28V. The AC output of transformers is converted into DC voltages by using rectifiers. A PIC 16F877 microcontroller is employed as the primary processor, it delivers gate trigger signals.

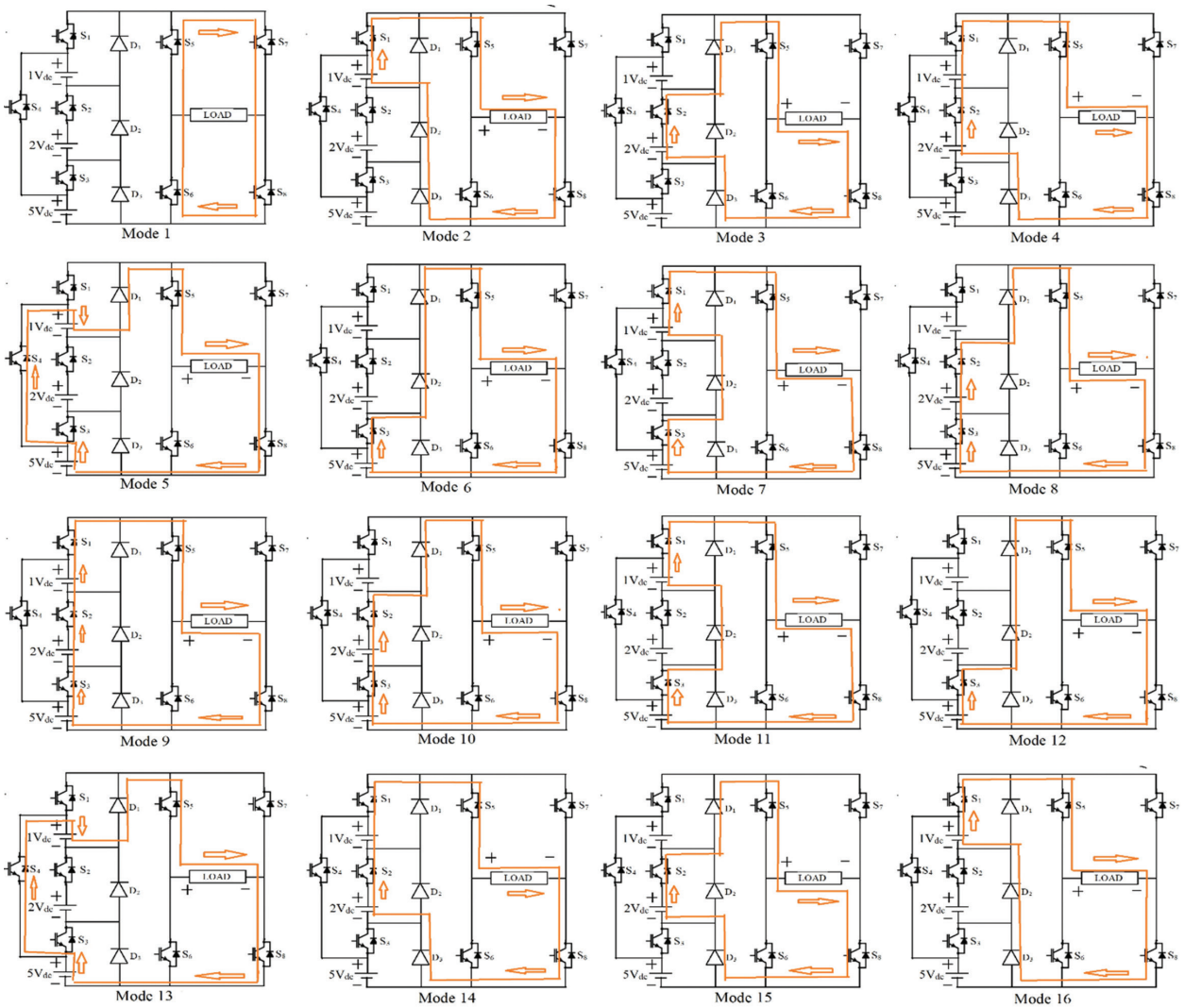



**Fig. 6.** Operation of New Reduced Switched Count multilevel inverter

**Table 1.** Operation of Proposed Multilevel Inverter

| | Modes | Load Current Path | | | | | | | | | | | Output Voltage(V) | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | D1 | D2 | D3 | | |
| Positive Half Cycle | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | OFF | OFF | OFF | 0Vdc | 0V |
| | 2 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | OFF | ON | ON | 1Vdc | +28V |
| | 3 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | ON | OFF | ON | 2Vdc | +56V |
| | 4 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | OFF | OFF | ON | 3Vdc | +84V |
| | 5 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | ON | OFF | OFF | 4Vdc | +112V |
| | 6 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | ON | ON | OFF | 5Vdc | +140V |
| | 7 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | OFF | ON | OFF | 6Vdc | +168V |
| | 8 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | ON | OFF | OFF | 7Vdc | +196V |
| | 9 | 1 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | OFF | OFF | OFF | 8Vdc | +224V |
| | 10 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | ON | OFF | OFF | 7Vdc | +196V |
| | 11 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | OFF | ON | OFF | 6Vdc | +168V |
| | 12 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 | ON | ON | OFF | 5Vdc | +140V |
| | 13 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | ON | OFF | OFF | 4Vdc | +112V |
| | 14 | 1 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | OFF | OFF | ON | 3Vdc | +84V |
| | 15 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 1 | ON | OFF | ON | 2Vdc | +56V |
| | 17 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | OFF | ON | ON | 1Vdc | +28V |
| | 18 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | OFF | OFF | OFF | 0Vdc | 0V |
| Negative Half Cycle | 19 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | OFF | OFF | OFF | 0Vdc | 0V |
| | 20 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | OFF | ON | ON | -1Vdc | -28V |
| | 21 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | ON | OFF | ON | -2Vdc | -56V |
| | 22 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | OFF | OFF | ON | -3Vdc | -84V |
| | 23 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | ON | OFF | OFF | -4Vdc | -112V |
| | 24 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | ON | ON | OFF | -5Vdc | -140V |
| | 25 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | OFF | ON | OFF | -6Vdc | -168V |
| | 26 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | ON | OFF | OFF | -7Vdc | -196V |
| | 27 | 1 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | OFF | OFF | OFF | -8Vdc | -224V |
| | 28 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | 0 | ON | OFF | OFF | -7Vdc | -196V |
| | 29 | 1 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | OFF | ON | OFF | -6Vdc | -168V |
| | 30 | 0 | 0 | 1 | 0 | 0 | 1 | 1 | 0 | ON | ON | OFF | -5Vdc | -140V |
| | 31 | 0 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | ON | OFF | OFF | -4Vdc | -112V |
| | 32 | 1 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | OFF | OFF | ON | -3Vdc | -84V |
| | 33 | 0 | 1 | 0 | 0 | 0 | 1 | 1 | 0 | ON | OFF | ON | -2Vdc | -56V |
| | 34 | 1 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | OFF | ON | ON | -1Vdc | -28V |
| | 35 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | OFF | OFF | OFF | 0Vdc | 0V |

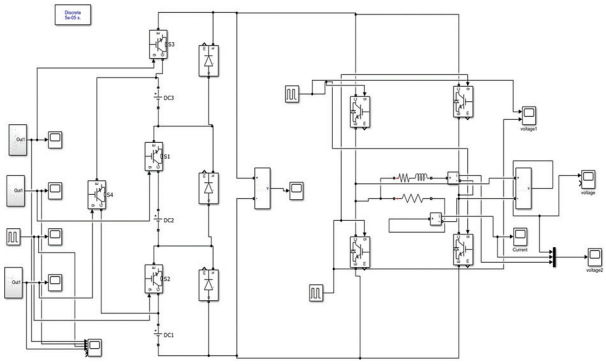


**Fig. 7.** Simulation Model of NRSCMLI

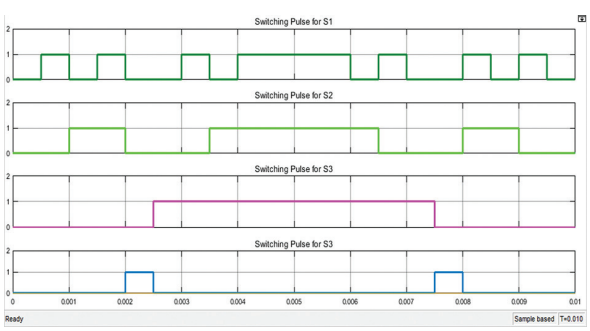



**Fig. 8.** Switching Pulses of NRSCMLI

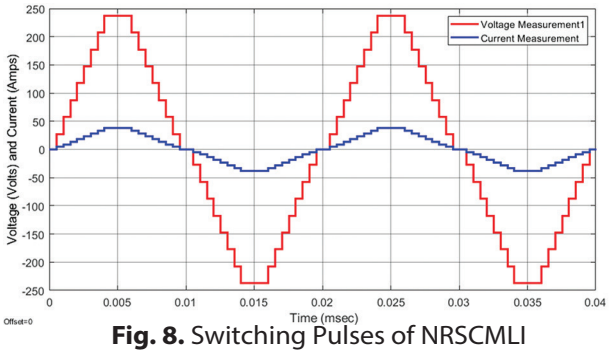



**Fig. 8.** Switching Pulses of NRSCMLI

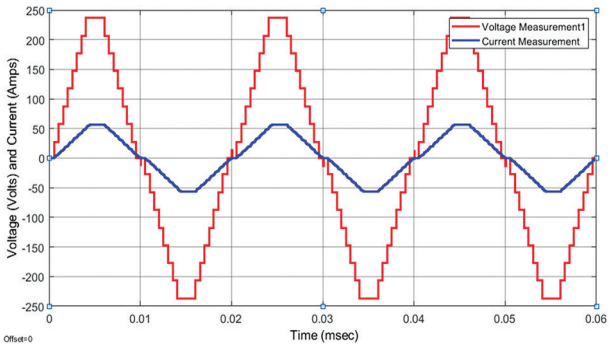



**Fig. 9.** 17-level output voltage and current of NRSCMLI for R load

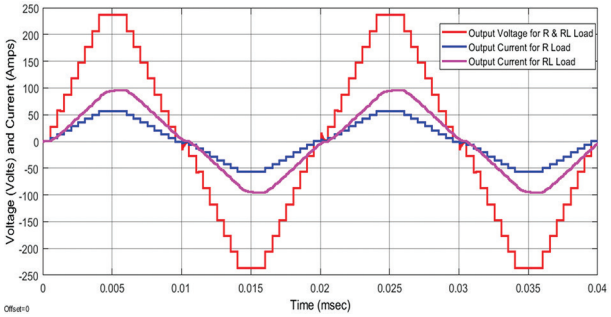



**Fig. 11.** Output voltage and current of NRSCMLI for parallel R and RL Load

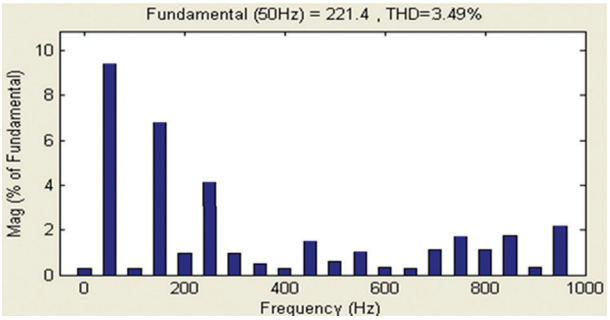



**Fig. 12.** FFT analysis

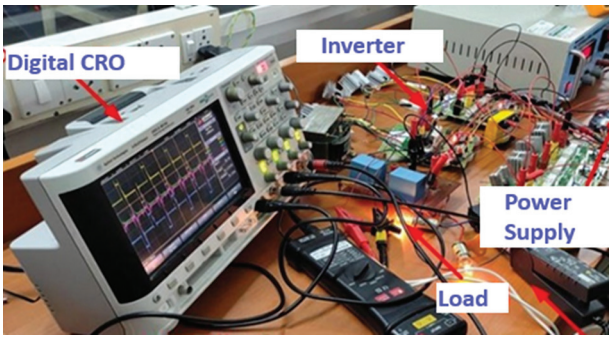



**Fig. 13.** Hardware output voltage waveform of new reduced switch count MLI

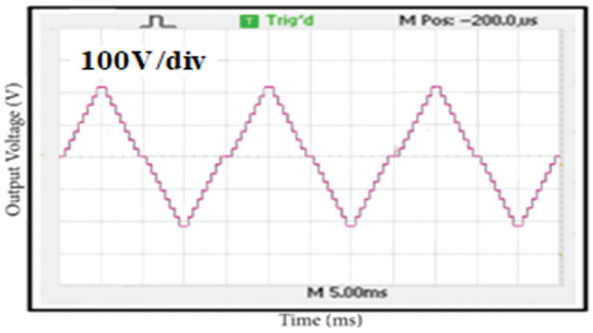



**Fig. 14.** Hardware output voltage waveform of new reduced switch count MLI

The microcontroller's control signal controls the MOSFET gate terminal's on and off states. Fig. 13 illustrates the hardware arrangement for the suggested NRSCMLI. Fig. 14 exhibits the hardware output voltage of the projected NRSCMLI. The output is measured for a resistive load of 1 kW. The hardware output current waveform is shown in Fig. 15. The magnitude of the load current is 2A and it is evident that the waveforms of the voltage and current are almost sinusoidal. Seventeen level MLI produces an output voltage of 224 volts at a frequency of 50 Hz. The efficiency was estimated using data from a MOSFET switch with variable output power. FFT analysis from Fig. 16 displays that the voltage THD is 3.51%.

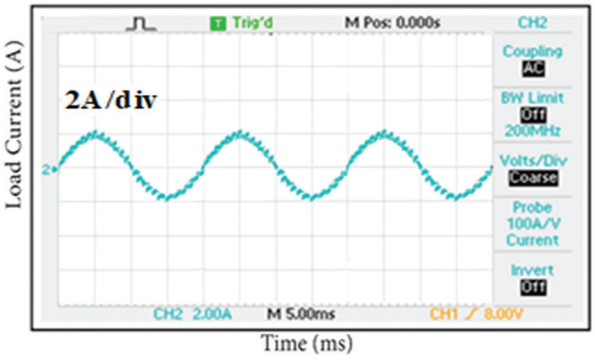



**Fig. 15.** Hardware output current waveform of new reduced switch count MLI

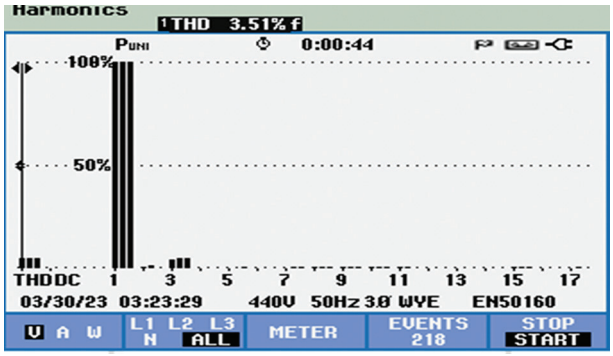



**Fig. 16.** Hardware FFT analysis of 17-level new reduced switch count MLI

## 7. CALCULATION OF TSV AND COST FUNCTION

From Fig. 1, components of conventional MLIs are calculated. Total standing voltage and cost functions for the MLIs are calculated as follows:

$$\text{Total Standing Voltage} = \sum_{i=1}^{n} MBV_{Si} \quad (14)$$

$$CF = (N_S + N_{gd} + N_D + N_C + \alpha TSV_{PU}) \times N_{DC} \quad (15)$$

$N_S$ – Number of Switches
$N_{gd}$ – Number of gate drives
$N_D$ – Number of Diodes
$N_C$ – Number of Capacitors
$TSV$ – Total Standing Voltage
$\alpha$ – Weight co-efficient
$N_{DC}$ – Number of Sources

### 7.1. DIODE CLAMPED MLI (DCMLI)

$N_S$=$N_{gd}$=16, $N_D$=14, $N_C$=8, $\alpha$=1, $N_{DC}$=1& $TSV_{PU}$ is 32
Cost Function = (16+16+14+8+32)*1= 86
$CF$/Level = $CF$/(Number of Levels) = 5.1

### 7.2. FLYING CAPACITOR MLI (FCMLI)

$N_S$=$N_{gd}$=16, $N_D$=0, $N_C$=18, $\alpha$=1, $N_{DC}$=1 & $TSV_{PU}$ is 32

Cost Function = (16+16+0+18+32)*1= 82

$CF$/Level = $CF$/(Number of Levels) = 4.8

### 7.3. CASCADED H-BRIDGE MLI (CHBMLI)

$N_S$=$N_{gd}$=32, $N_D$=$N_C$=0, $\alpha$=1, $N_{DC}$=8 & $TSV_{PU}$ is 4

Cost Function = (32+32+0+0+4)*8= 544

$CF$/Level = $CF$/(Number of Levels) = 32

### 7.4. CASCADED H-BRIDGE HYBRID MLI (CHBHMLI)

$N_S$=$N_{gd}$=12, $N_D$=$N_C$=0, $\alpha$=1, $N_{DC}$=3 & $TSV_{PU}$ is 6

Cost Function = (12+12+0+0+6)*3= 90

$CF$/Level = $CF$/(Number of Levels) = 5.3

### 7.5. NEW CASCADED MLI (NCMLI)

$N_S$=20, $N_{gd}$=18, $N_D$=$N_C$=0, $\alpha$=1, $N_{DC}$=8 & $TSV_{PU}$ is 11

Cost Function = (20+18+0+0+6)*8= 392

$CF$/Level = $CF$/(Number of Levels) = 23.1

### 7.6. NEW CASCADED HYBRID MLI (NCHMLI)

$N_S$=12, $N_{gd}$=10, $N_D$=$N_C$=0, $\alpha$=1, $N_{DC}$=4 & $TSV_{PU}$ is 8.25

Cost Function = (12+10+0+0+8.25)*4= 121

$CF$/Level = $CF$/(Number of Levels) = 7.1

### 7.7. MODIFIED CASCADED MLI (MCMLI)

$N_S$=12, $N_{gd}$=10, $N_D$=$N_C$=0, $\alpha$=1, $N_{DC}$=4 & $TSV_{PU}$ is 8.25

Cost Function = (12+10+0+0+8.25)*4= 121

$CF$/Level = $CF$/(Number of Levels) = 7.1

### 7.8. MODIFIED HYBRID MLI (MCMLI)

$N_S$=8, $N_{gd}$=6, $N_D$=4, $N_C$=0, $\alpha$=1, $N_{DC}$=4 & $TSV_{PU}$ is 6.12

Cost Function = (8+6+4+0+6.12)*4= 96.5

$CF$/Level = $CF$/(Number of Levels) = 5.7

### 7.9. PROPOSED MLI

Total standing voltage and cost functions for the proposed MLI are calculated as follows:

$TSV = MBV_{S1}+MBV_{S2}+MBV_{S3}+MBV_{S4}+MBV_{S5}+MBV_{S6}+MBV_{S7}+MBV_{S8}$

$TSV$=3.5$V_{dc}$+4$V_{dc}$+2.5$V_{dc}$+3.5$V_{dc}$+4$V_{dc}$+4$V_{dc}$+4$V_{dc}$+4$V_{dc}$ = 29.5$V_{dc}$

$TSV_{PU} = TSV/MBV$ = (29.5$V_{dc}$)/(8$V_{dc}$ )=3.6875

$N_S$ =8, $N_{gd}$ = 6, $N_D$ = 3, $N_C$ =0, $\alpha$ =1 and $N_{DC}$ = 3

$CF$=(8+6+3+0+1×3.6875)×3= 62.0625

$CF$/Level = $CF$/(Number of Levels) = 62.0625/17=3.65

From the above calculations, the cost function per level is lower when compared to conventional topologies. So, it can be easily configured as three phase system with 24 switches, 20 driver circuits, 9 DC sources, and 9 diodes.

## 8. PERFORMANCE COMPARISON OF MULTILEVEL INVERTERS

The several MLI structures are compared to stages associated with the quantity of power switches employed. Table 2 shows the comparison of various types of MLIs based on the needs of switches, driver circuits, diodes, capacitors, and DC sources. TBVpu, cost function per unit and THD are also compared. Fig. 17 represents the pictorial representations of the comparison of various existing MLIs. Figure 18 depicts the comparison of the efficiency of various existing MLIs, it is clear that the suggested topology produces 98.60%. The NRSCMLI has benefits of 17 level with only 8 power switches when multisource is utilized; THD is 3.51% and efficiency is also high when compared to various multilevel inverters. The proposed MLI 3 DC sources are used to generate 17-levels output, whereas in conventional inverters, a minimum of 4 DC sources are required; otherwise, one or two sources with more capacitors are required to generate the same 17-level output. With the use of more capacitors, voltage balancing issues may arise. From the comparison, it appears that the developed MLI is best suited for electric vehicle applications with low power switches, reduced THD, and more efficiency.

**Table 2.** Performance Comparison of Multilevel Inverters

| S.No. | Name of the Topology | Switches Needed | Driver Circuits Needed | Diodes Needed | Capacitors Needed | DC Sources Needed | $TBV_{pu}$ | $CF$ $\alpha$=0.5 | $CF$ $\alpha$=1 | $THD$ | Efficiency (%) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | DCMLI | 16 | 16 | 14 | 8 | 1 | 32 | 4.1 | 5.1 | 7.53 | 96.81 |
| 2 | FCMLI | 16 | 16 | 0 | 18 | 1 | 32 | 3.9 | 4.8 | 7.12 | 96.52 |
| 3 | CHBMLI | 32 | 32 | 0 | 0 | 8 | 4 | 31.1 | 32 | 7.81 | 95.75 |
| 4 | CHBHMLI | 12 | 12 | 0 | 0 | 3 | 6 | 4.8 | 5.3 | 7.1 | 97.10 |
| 5 | NCMLI | 20 | 18 | 0 | 0 | 8 | 11 | 20.5 | 23.1 | 7.44 | 96.21 |
| 6 | NCHMLI | 12 | 10 | 0 | 0 | 4 | 8.25 | 6.1 | 7.1 | 5.51 | 97.12 |
| 7 | MCMLI | 12 | 10 | 8 | 0 | 8 | 7.5 | 15.9 | 17.6 | 6.26 | 96.88 |
| 8 | MHMLI | 8 | 6 | 4 | 0 | 4 | 6.12 | 5.0 | 5.7 | 4.71 | 98.24 |
| 9 | Proposed MLI | 8 | 6 | 3 | 0 | 3 | 3.68 | 3.3 | 3.6 | 3.49 | 98.60 |

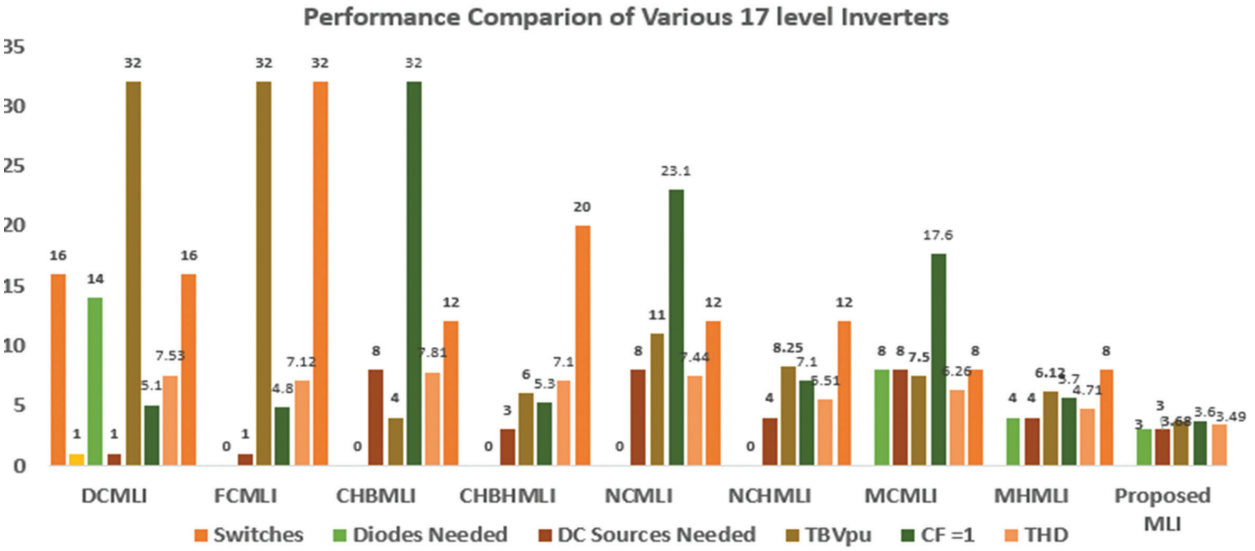



**Fig. 17.** Performance Comparison of various existing 17 level inverters

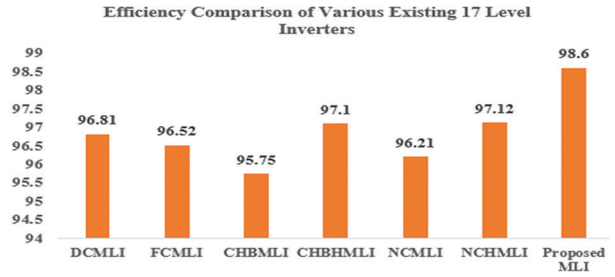



**Fig. 18.** Comparison of Efficiency of various MLIs

## 9. CONCLUSION

The proposed work has been validated on the fundamental structural set-up of the new reduced switch count MLI for electric vehicle applications. The different MLI structures and their circuit operations have been discussed. The mathematical expressions for the calculation of the needed voltage magnitude at the load have been pronounced. In the conservative approach, as the levels increase, the essential power switches also increase. Due to the large count of power switches, the harmonics, total harmonic distortion, switching losses, and implementation costs have also increased. The suggested topology considerably reduces the power switches to 8 for 17 levels, which lowers the losses due to reduced switching and conduction losses and improves the efficiency to 98.60%. The cost function is 3.6 for a weight coefficient of 1. Lower-order harmonics are removed, which leads to an improvement in reduced total harmonic distortion (THD) of 3.49%. Hence, the developed MLI is best suited for electric vehicle applications.

## 10.REFERENCES


[1] A. Bughneda, M. Salem, A. Richelli, D. Ishak, S. Alatai, "Review of Multilevel Inverters for PV Energy System Applications", Energies, Vol. 14, No. 6, 2021, p. 1585.

[2] J. Rodriguez, J. S. Lai, F. Z. Peng, "Multilevel Inverters: a Survey of Topologies, Controls and Applications", IEEE Transactions on Industrial Electronics, Vol. 49, No. 4, 2002, pp. 724-738.

[3] D. A. B. Zambra, C. Rech, J. R. Pinheiro, "Comparison of Neutral-Point-Clamped, Symmetrical and Hybrid Asymmetrical Multilevel Inverters", IEEE Transactions on Industrial Electronics, Vol. 57, No. 7, 2010, pp. 2297-2306.

[4] J. Zhang, S. Xu Z. Din, X. Hu, "Hybrid Multilevel Converters: Topologies, Evolutions and Verifications", Energies, Vol. 2, 2019, p. 615.

[5] D. Ruiz-Caballero, R. M. Ramos-Astudillo, S. A. Mussa, M. L. Heldwein, "Symmetrical Hybrid Multilevel DC–AC Converters with Reduced Number of Insulated DC Supplies", IEEE Transactions on Industrial Electronics, Vol. 57, No. 7, 2010, pp. 2307-2314.

[6] Y. Hinge, H. Koizumi, "A Single-Phase Multilevel Inverter using Switched Series/Parallel DC Voltage Sources", IEEE Transactions on Industrial Electronics, Vol. 57, No. 8, 2010, pp. 2643-2650.

[7] N. S. Hasan, N. Rosmin, D. A. A. Osman, A. H. Mustaamal, "Reviews on Multilevel Converter and Modulation Techniques", Renewable and Sustainable Energy Reviews, Vol. 80, 2017, pp. 163'174.

[8] J. Fang, Z. Li, S. M. Goet, "Multilevel Converters with Symmetrical Half-Bridge Submodules and Sensorless Voltage Balance", IEEE Transactions on Industrial Electronics, Vol. 36, No. 1, 2020, pp. 447-458.

[9] H. Nasiri Avanaki, R. Barzegarkhoo, E. Zamiri, Y.

Yang, F. Blaabjerg, "Reduced Switch-Count Structure for Symmetric Multilevel Inverters with a Novel Switched-DC-Source Submodule", IET Power Electronics, Vol. 12, No. 2, 2019, pp. 311-321.

[10] S. Alatai, M. Salem, D. Ishak, H. S. Das, M. A. Nazari, A. Bughneda, M. Kamarol, "A Review on State-of-the-Art Power Converters: Bidirectional, Resonant, Multilevel Converters and their Derivatives", Applied Science, Vol. 11, No. 21, 2021, pp. 1-43.

[11] M. Salem, A. Richelli, K. Yahya, M. N. Hamidi, T. Z. Ang, I. Alhamrouni, "A Comprehensive Review on Multilevel Inverters for Grid-Tied System Applications", Energies, Vol. 15, No. 17, 2022, p. 6315.

[12] K. P. Panda, S. S. Lee, G. Panda, "Reduced Switch Cascaded Multilevel Inverter with New Selective Harmonic Elimination Control for the Standalone Renewable Energy System", IEEE Transactions on Industry Applications, Vol. 55, No. 6, 2019, pp. 7561-7574.

[13] F. Richardeau, T. T. L. Pham, "Reliability Calculation of Multilevel Converters: Theory and Applications", IEEE Transactions on Industrial Electronics, Vol. 60, No. 10, 2012, pp. 4225-4233.

[14] P. Wiatr, M. P. Kazmierkowski, "Model Predictive Control of Multilevel Cascaded Converter with Boosting Capability – a Simulation Study", Bulletin of the Polish Academy of Sciences: Technical Sciences, Vol. 64, No. 3, 2016, pp. 581-590.

[15] M. Malinowski, "Cascaded Multilevel Converters in Recent Research and Applications", Bulletin of the Polish Academy of Sciences: Technical Sciences, Vol. 65, No. 5, 2017, pp. 567–576.

[16] J. I. Leon, S. Vazquez, L. G. Franquelo, "Multilevel Converters: Control and Modulation Techniques for their Operation and Industrial Applications", Proceedings of the IEEE, Vol. 105, No. 11, 2017, pp. 2066-2081.

[17] A. Poorfakhraei, M. Narimani, A. Emadi, "A Review of Modulation and Control Techniques for Multilevel Inverters in Traction Applications", IEEE Access, Vol. 9, 2021, pp. 24187-24204.

[18] K. Sadigh, M. Abarzadeh, K. A. Corzine, V. Dargahi, "A New Breed of Optimized Symmetrical and Asymmetrical Cascaded Multilevel Power Converters", IEEE Journal of Emerging and Selected Topics in Power Electronics, Vol. 3, No. 4, 2015, pp. 1160-1170.

[19] H. Tu, H. Feng, S. Srdic, S. Lukic, "Extreme Fast Charging of Electric Vehicles: A Technology Overview", IEEE Transactions on Transportation Electrification, Vol. 5, No. 4, 2019, pp. 861-878.

[20] R. Barzegarkhoo, M. Forouzesh, S. S. Lee, F. Blaabjerg, Y. Siwakoti, "Switched-Capacitor Multilevel Inverters: A Comprehensive Review", IEEE Transactions on Power Electronics, Vol. 37, No. 9, 2022, pp. 11209-11243.

[21] S. T. Meraj, K. Hasan, A. Masaoud "A Novel Configuration of Cross-Switched T-Type (CT-Type) Multilevel Inverter," IEEE Transactions on Power Electronics, Vol.3 5, No. 4, 2019, pp. 3688-3696.

[22] M. N. Hamidi, D. Ishak, M. A. A. M. Zainuri, C. A. Ooi, "Multilevel Inverter with Improved Basic Unit Structure for Symmetric and Asymmetric Source Configuration", IET Power Electronics, Vol. 13, No. 7, 2020, pp. 1445-1455.

[23] N. P. Gopinath, K. Vijayakumar, J. S. Mohd Ali, K. Raghupathi, S. Selvam, "A Triple Boost Seven-Level Common Ground Transformerless Inverter Topology for Grid-Connected Photovoltaic Applications", Energies, Vol. 16, No. 8, 2023, p. 3428.

[24] A. Palani, V. Mahendran, K. Vengadakrishnan, S. Muthusamy, O. P. Mishra, M. R. Maurya, K. K. Sadasivuni, "A Novel Design and Development of Multilevel Inverters for Parallel Operated PMSG-Based Standalone Wind Energy Conversion Systems", Iranian Journal of Science and Technology, Transactions of Electrical Engineering, Vol. 48, 2024, pp. 277-287.

[25] M. Murugesan, K. Lakshmi, "Design and Implementation of Modified Multilevel Inverter FED BLDC Motor for Electric Vehicle Application", Proceedings of the Bulgarian Academy of Sciences, Vol. 76, No. 10, 2023, pp. 15625-1571.

[26] K. V. S. Kumar, A. Dheepanchakkravarthy, "A Single Source Hybrid Nine-Level Multilevel Inverter with Extension Topology", Advances in Electrical and Computer Engineering, Vol. 23, No. 2, 2023, pp. 75-84.

[27] B. Pragathi, S. Kumar, P. R. Kumari, A. R. Singh, "Performance Evaluation of Hybrid Multilevel Inverter with a High-Frequency Switching Technique", Journal of Engineering and Applied Science, Vol. 70, No. 101, 2023, p. 976.

# INTERNATIONAL JOURNAL OF ELECTRICAL AND COMPUTER ENGINEERING SYSTEMS

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

## About this Journal

The International Journal of Electrical and Computer Engineering Systems publishes original research in the form of full papers, case studies, reviews and surveys. It covers theory and application of electrical and computer engineering, synergy of computer systems and computational methods with electrical and electronic systems, as well as interdisciplinary research.

## Topics of interest include, but are not limited to:

- Power systems
- Renewable electricity production
- Power electronics
- Electrical drives
- Industrial electronics
- Communication systems
- Advanced modulation techniques
- RFID devices and systems
- Signal and data processing
- Image processing
- Multimedia systems
- Microelectronics
- Instrumentation and measurement
- Control systems
- Robotics
- Modeling and simulation
- Modern computer architectures
- Computer networks
- Embedded systems
- High-performance computing
- Parallel and distributed computer systems
- Human-computer systems
- Intelligent systems
- Multi-agent and holonic systems
- Real-time systems
- Software engineering
- Internet and web applications and systems
- Applications of computer systems in engineering and related disciplines
- Mathematical models of engineering systems
- Engineering management
- Engineering education

## Paper Submission

Authors are invited to submit original, unpublished research papers that are not being considered by another journal or any other publisher. Manuscripts must be submitted in doc, docx, rtf or pdf format, and limited to 30 one-column double-spaced pages. All figures and tables must be cited and placed in the body of the paper. Provide contact information of all authors and designate the corresponding author who should submit the manuscript to https://ijeces.ferit.hr. The corresponding author is responsible for ensuring that the article's publication has been approved by all coauthors and by the institutions of the authors if required. All enquiries concerning the publication of accepted papers should be sent to ijeces@ferit.hr.

The following information should be included in the submission:

- paper title;
- full name of each author;
- full institutional mailing addresses;
- e-mail addresses of each author;
- abstract (should be self-contained and not exceed 150 words). Introduction should have no subheadings;
- manuscript should contain one to five alphabetically ordered keywords;
- all abbreviations used in the manuscript should be explained by first appearance;
- all acknowledgments should be included at the end of the paper:
- authors are responsible for ensuring that the information in each reference is complete and accurate. All references must be numbered consecutively and citations of references in text should be identified using numbers in square brackets. All references should be cited within the text;
- each figure should be integrated in the text and cited in a consecutive order. Upon acceptance of the paper, each figure should be of high quality in one of the following formats: EPS, WMF, BMP and TIFF;
- corrected proofs must be returned to the publisher within 7 days of receipt.

## Peer Review

All manuscripts are subject to peer review and must meet academic standards. Submissions will be first considered by an editor-in-chief and if not rejected right away, then they will be reviewed by anonymous reviewers. The submitting author will be asked to provide the names of 5 proposed reviewers including their e-mail addresses. The proposed reviewers should be in the research field of the manuscript. They should not be affiliated to the same institution of the manuscript author(s) and should not have had any collaboration with any of the authors during the last 3 years.

## Author Benefits

The corresponding author will be provided with a .pdf file of the article or alternatively one hardcopy of the journal free of charge.

### Units of Measurement

Units of measurement should be presented simply and concisely using System International (SI) units.

## Bibliographic Information

Commenced in 2010.
ISSN: 1847-6996
e-ISSN: 1847-7003

Published: semiannually

## Copyright



## Subscription Information

The annual subscription rate is 50€ for individuals, 25€ for students and 150€ for libraries.

## Postal Address

Faculty of Electrical Engineering,
Computer Science and Information Technology Osijek,
Josip Juraj Strossmayer University of Osijek, Croatia
Kneza Trpimira 2b
31000 Osijek, Croatia

# IJECES Copyright Transfer Form

(Please, read this carefully)

This form is intended for all accepted material submitted to the IJECES journal and must accompany any such material before publication.

**TITLE OF ARTICLE** (hereinafter referred to as "the Work"):

COMPLETE LIST OF AUTHORS:

The undersigned hereby assigns to the IJECES all rights under copyright that may exist in and to the above Work, and any revised or expanded works submitted to the IJECES by the undersigned based on the Work. The undersigned hereby warrants that the Work is original and that he/she is the author of the complete Work and all incorporated parts of the Work. Otherwise he/she warrants that necessary permissions have been obtained for those parts of works originating from other authors or publishers.

Authors retain all proprietary rights in any process or procedure described in the Work. Authors may reproduce or authorize others to reproduce the Work or derivative works for the author's personal use or for company use, provided that the source and the IJECES copyright notice are indicated, the copies are not used in any way that implies IJECES endorsement of a product or service of any author, and the copies themselves are not offered for sale. In the case of a Work performed under a special government contract or grant, the IJECES recognizes that the government has royalty-free permission to reproduce all or portions of the Work, and to authorize others to do so, for official government purposes only, if the contract/grant so requires. For all uses not covered previously, authors must ask for permission from the IJECES to reproduce or authorize the reproduction of the Work or material extracted from the Work. Although authors are permitted to re-use all or portions of the Work in other works, this excludes granting third-party requests for reprinting, republishing, or other types of re-use. The IJECES must handle all such third-party requests. The IJECES distributes its publication by various means and media. It also abstracts and may translate its publications, and articles contained therein, for inclusion in various collections, databases and other publications. The IJECES publisher requires that the consent of the first-named author be sought as a condition to granting reprint or republication rights to others or for permitting use of a Work for promotion or marketing purposes. If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IJECES publisher assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Authors of IJECES journal articles and other material must ensure that their Work meets originality, authorship, author responsibilities and author misconduct requirements. It is the responsibility of the authors, not the IJECES publisher, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it.

- The undersigned represents that he/she has the authority to make and execute this assignment.
- For jointly authored Works, all joint authors should sign, or one of the authors should sign as authorized agent for the others.
- The undersigned agrees to indemnify and hold harmless the IJECES publisher from any damage or expense that may arise in the event of a breach of any of the warranties set forth above.

_____ _____

**Author/Authorized Agent** **Date**

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