A comparative analysis of chronic obstructive pulmonary disease using machine learning and deep learning

Ramadoss Ramalingam[1], Vimala Chinnaiyan[2]
Department of Electronics and Communication Engineering, SRM Institute of Science and Technology, Kattankulathur, India

Article Info

Article history:

Received Oct 20, 2021
Revised Aug 3, 2022
Accepted Sep 6, 2022

Keywords:

Chronic obstructive pulmonary disease
Convolutional neural network
Deep learning
Lung computed tomography image
Machine learning

ABSTRACT

Chronic obstructive pulmonary disease (COPD) is a general clinical issue in numerous countries considered the fifth reason for inability and the third reason for mortality on a global scale within 2021. From recent reviews, a deep convolutional neural network (CNN) is used in the primary analysis of the deadly COPD, which uses the computed tomography (CT) images procured from the deep learning tools. Detection and analysis of COPD using several image processing techniques, deep learning models, and machine learning models are notable contributions to this review. This research aims to cover the detailed findings on pulmonary diseases or lung diseases, their causes, and symptoms, which will help treat infections with high performance and a swift response. The articles selected have more than 80% accuracy and are tabulated and analyzed for sensitivity, specificity, and area under the curve (AUC) using different methodologies. This research focuses on the various tools and techniques used in COPD analysis and eventually provides an overview of COPD with coronavirus disease 2019 (COVID-19) symptoms.

Corresponding Author:

Vimala Chinnaiyan
Department of Electronics and Communication Engineering, SRM Institute of Science and Technology
Kattankulathur, 603203, Tamil Nadu, India
Email: vimalac@srmist.edu.in

1. INTRODUCTION

Pulmonary diseases, otherwise referred to as lung diseases or respiratory diseases, affect any part of the respiratory system. Like cancer and diabetes, it is conceivably the most notable in general illnesses. Smoking, infections, genes, and other forms of air pollution, such as radon or breathing in second-hand smoke, are the leading causes of lung diseases [1]. Generally, the lungs expand and relax a thousand-fold each day for oxygen intake and carbon dioxide removal [2]. Pulmonary disease is solely not related to bronchial asthma or coughing. Chronic obstructive pulmonary disease (COPD), sinusitis, lung cancer, chronic pneumonia, and now coronavirus disease 2019 (COVID-19) are all examples of pulmonary diseases [3]. This study is focused on image processing and other cutting-edge methodologies in the diagnosis of COPD. One of the most common respiratory diseases is COPD.

Figure 1 depicts the differences between healthy airways, chronic bronchitis, emphysema, and COPD symptoms and causes. COPD allows less intake of air than usual in the lungs. The symptoms of COPD include difficulty in breathing, coughing, dyspnea, sputum production, wheezing, and chest tightness [4]. The effects of COPD lead to defects in the lungs and brain due to hypoxia. Rational thinking and memory lapses may affect a patient having COPD. COPD is classified into two types: chronic bronchitis and emphysema. Chronic bronchitis incorporates a prolonged cough with secretion, whereas emphysema harms the lungs over time [4], [5]. The majority of COPD patients have a combination of these two diseases discussed earlier. Therefore, COPD is also a preventable and treatable illness.

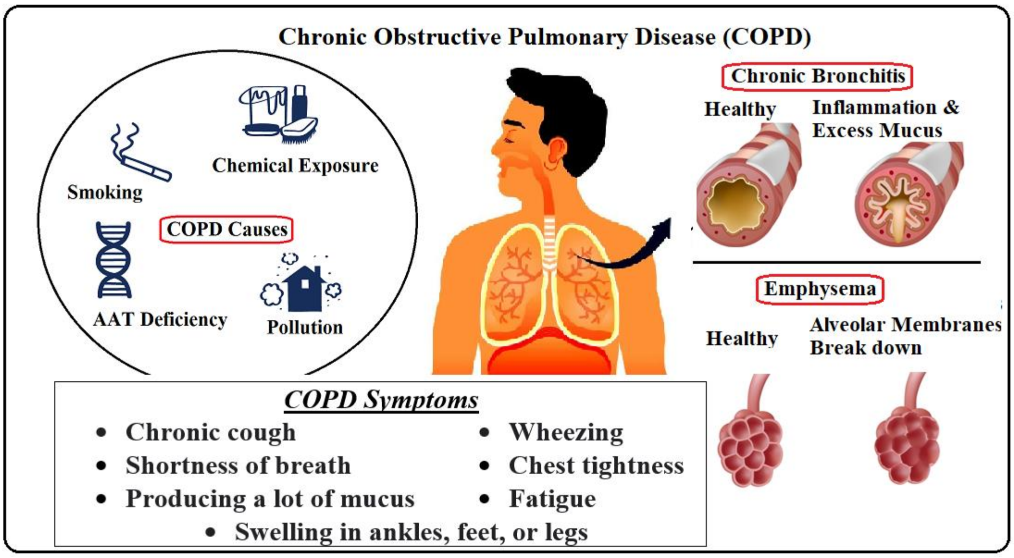

Figure 1. Chronic obstructive pulmonary disease

Different approaches, including machine learning and deep learning (either deep or non-deep) models with convolutional neural network (CNN), were recently developed to recognize and categorize pulmonary illness in medical pictures without difficulty [6], [7]. In theory, the benefits of a deep learning (DL) technique for increasing the precision and efficiency of human COPD detection and enhancing human understanding of COPD subtypes have been demonstrated [8]–[10]. Deep CNN has made significant changes in how both natural and medical images are processed [11]–[14]. In this paper, a concise review of pulmonary diseases, including COPD, is outlined in a detailed fashion. It is acknowledged that this review paper can provide sufficient data concerning pulmonary diseases and the subsequent elucidation regarding their detection and classification using image processing techniques and a deep CNN.

Pulmonary disease is a respiratory disease that affects the lungs as well as the other respiratory organs. The disease is primarily caused as a result of smoking, inhalation of second-hand tobacco, and infection due to asbestos, radon, or air pollution [9], [15], [16]. Image processing is a process that is used to apply a relevant mathematical operation or an algorithm to a digitized image which results in an upgraded image [17], [18].

Medical imaging analysis using software tools is a valuable diagnostic tool for a wide range of diseases, where image information is an essential part of the decision-making process [19]. COPD research that used artificial intelligence (AI) and machine learning (ML) methods was looked up in this study, which looked at the research in the literature [19]. Several types of medical images that are used for analysis for research and diagnostics include magnetic resonance imaging (MRI), chest X-ray [20], positron emission tomography (PET)-computed tomography (CT) scan, ventilation/perfusion scan (V/Q scan) [21], and CT scan, pulmonary angiography, ultrasound, and biopsy. This paper concentrates on and exploits CT images to train the available existing models and the proposed models. All relevant research was gathered and documented in the "Related works" section.

2. RELATED WORKS

CT pictures of respiratory organs are commonly used to assess and analyze the study of COPD [1]. A study included inspiratory, and expiratory CT scans and pulmonary function test (PFT) measured parameters from 69 test subjects, in which a count of 13 healthy subjects and 56 COPD-affected subjects at diverse tiers were selected and analyzed. The leave-m-out cross-validation approach was used by considering m=7, were tested the classifier's performance. The accuracy of the results acquired during this experiment was greater than 84%, which demonstrated its efficiency in identifying the COPD stage using just CT scans without requiring PFT values. This indicates that the image-based COPD detection tools contribute to the professional vision of providing remedial treatment.

The earlier the determination of sickness is investigated and classified, the higher the rate and possibility of relieving the patient [3]. AI and expert systems were successfully used to explore different issues in the clinical field, such as shortening the finding and classification period, time acquirement, and extended capability. Thus, artificial neural networks (ANN) appear to have an effective utilization in COPD classification. A framework was formulated which enabled the doctor to determine the COPD level rapidly and with high performance. This paper utilizes the support vector machines (SVM) and k-nearest neighbors (KNN) strategies to arrange the COPD severity levels. These methodologies are assessed using a test dataset sample from the Prajna Health Care Medical Hospital. The exploratory outcome demonstrates the effectiveness of these strategies. Multilayer neural networks (MLNN) are used to recognize the examination for COPD diagnosis [8]. As a result, two distinct MLNN structures are employed. One structure was a one-hidden-layer MLNN, while the other was a two-hidden-layer MLNN. For neural network training, the Levenberg-Marquardt algorithms and the back-propagation technique with momentum are employed. COPD datasets incorporate respiratory organ illness measurements including two classes. One hundred fifty-five samples are created utilizing the epicrisis reports of patients from the chest disease hospital. The neural network architecture is effectively used to diagnose COPD illness using the MLNN-based decision support system, which effectively assists clinicians in their diagnosis decisions.

The presence of the outwardly evaluated emphysema from CT outcomes and pneumonic capacity tests is a result of the advancement of cellular breakdown in the lungs. COPD may be predicted using CNN on the CT images. In the National Lung Cancer Screening Study, the suggested COPD and emphysema classification based on CNN indicates a cellular breakdown in the lungs. The CNN classifier created this model using COPD gene research data, including spirometric COPD and emphysema [9]. The classifiers are used to calculate the emphysema and COPD scores using 7347 CT images from the National Lung Screening Trial (NLST): element spatial convolution neural network (ESCNN) and cost-sensitive convolution neural network (CSCNN), respectively). The model is built using the cox proportional hazards model to account for the effects of CSCNN, ESCNN, education, gender, age, body mass index (BMI), pack-years of smoking, and time after quitting smoking. The model is used to analyze the cellular breakdown in the lungs. Both CSCNN and ESCNN were established as critical indicators of cellular breakdown in the lungs, which is very dangerous.

Another research investigates the parameters which distinguish asthma and COPD diagnosis. “Machine learning” refers to how computers learn from information. Machines can learn in a variety of ways (known as “algorithms”) [10]. This research aims to manage patients with COPD and asthma in real-time and is included in the clinical decision support systems (CDSSs). The techniques in this research comprise a random forest classifier with 132 patients and 22 features. The performing properties for this method are 97.7% precise for COPD diagnosis and 80.3% precise for asthma diagnosis.

Quantitative analysis of the pulmonary distribution pattern of pulmonary emphysema is an important field of research [14]. This research describes the automatic removal of emphysematous damage from three-dimensional CT images. Computer-aided diagnosis (CAD) based 3-D pulmonary images are used to analyze the emphysematous lesions extracted automatically using the region growing approach. In terms of classification systems, multilayer neural networks (MLNNs) have been effectively employed instead of traditional pattern recognition approaches for illness diagnosis [16], [17].

COPD refers to a collection of lung diseases that includes asthma, emphysema, chronic bronchitis, and some types of bronchiectasis [22]. The research objective is to provide a novel methodology to classify COPD from X-rays. Here, the outlined model was evaluated using 600 x-ray images. The methodology used in this paper included the image processing technique and various classifiers such as LDA, neural network-based classifier, maximum likelihood classifier, and genetic algorithm (GA). The performance of this method classification had 97.9% accuracy.

Syndrome differentiation is an essential component of traditional Chinese medicine and has a therapeutic effect [23]. This paper focuses on the usage of an ANN model to investigate the intelligent syndrome differentiation (ISD) for traditional Chinese medicine. This article’s methodology consists of 18471 genuine clinical records as well as four additional models with ANN and four subgroup datasets. The performance of this method has an accuracy of 86.45% and an F1 score of 82.93%. It is revealed that the four-subgroup model provides better performance when compared to the full group ANN.

Quantitative computed tomography (QCT) measurements show a relationship between pulmonary function tests and symptoms [24]. The work in this paper focuses on the determination of QCT parameters and the prediction of lung function characteristics. The parameters of five tested alternative models of machine learning were compared with QCT values. The methodology discussed in this paper involved the examination of 75 COPD patients. Based on this examination, the delta values were calculated. Four different parameters were estimated to predict five other machine learning models, which included gradient boosting, KNN, median prediction, average prediction, and multilayer perception (MLP). The optimal performance in this method is realized at 16% of the lowest mean relative error for the KNN. This limitation of this research or research gap revealed that the other four models produced a high mean relative error, which resulted in the worst predictive performance.

The examination assists several AI approaches in the detection of COPD utilizing multi-channel lung sounds [25]. DL is a productive AI approach that decreases optimization through unsupervised training and supervised training with the aid of a characteristic-based total distribution of category parameters. It focuses on the investigation of multi-channel lung sounds utilizing statistical aspects of frequency modulations derived using the Hilbert-Huang transform. The suggested model's set of rules was used in the characterization phase of the proposed model to discriminate between patients with COPD and healthy persons. As a result, the DL model based on the Hilbert-Huang transform with statistical features achieved unprecedented classification execution rates of 93.67% accuracy, 91% sensitivity, and 96.33% specificity.

Computerization and digitization of the artery and vein classing of CT scans have been gaining popularity as it aids diagnosticians in diagnosing chronic disorders more precisely [26]. This study describes a unique, automatic method for categorizing vessels such as arteries and veins in chest scans. The methodology is staged into five parts: i) scale-space particle segmentation is employed to separate vessels; ii) a three-dimensional CNN is employed to get the primary vessel classification and; iii) graph cut optimization is used to fine-tune the results; iv) a CNN methodology compared with random forests (RFs) classifier to support the CNN architecture; v) finally, it investigates the modified two-dimensional and three-dimensional CNNs that will incorporate the local data from bronchus and vessel improved pictures to the network using various ways.

Some researchers established COPD as a respiratory disease characterized by variations in lung architecture. The research objective involves the formulation of a novel classification approach established on deep erudition and parametric response mapping [27]. The approach described in this research used the 3D CNN and the 3D parametric response map (PRM) dataset. This approach established an accuracy of 89.3 percent and a sensitivity of 88.3% using a 5-fold cross-validation strategy. Due to this, the parenchymal abnormalities in COPD patients can be effectively detected. Other researchers analyzed the discovery of different levels of COPD by forecasting acute respiratory disease. The methodology explained in this paper implemented 2D CNN using original CT slices from the COPDGene testing cohort. The research utilized 7,983 COPDGene participants, 1,000 non-overlapping COPDGene participants and 1,672 ECLIPSE participants [28]. The COPD was assessed, and the acute respiratory disease (ARD) was predicted using logistic regression. The cox regression method was used to calculate mortality among smokers. This technique achieved a 77.3% accuracy level.

In a point-of-care (PoC) environment [29], AI may play a critical role in decomposing information and providing complete results for COPD detection. The research goal was to utilize equipment pedagogy devices on data collected from patients with COPD and healthy saliva tests, and their segment data was used to detect infection at the point of care. The dielectric properties of the saliva samples were depicted using a permittivity biosensor, and the resulting data were categorized using ML algorithms. The XGBoost gradient boosting set of rules achieved a classification accuracy of 91.25% and a sensitivity of 100% making it an effective variant for COPD assessment.

Explains that COPD is related to the morphological anomalies of airways with different examples and severities [30]. This paper focuses on utilizing the deep learning models CNN to determine COPD from multi-view images of three-dimensional pulmonary airway trees. The image inputs used are: i) air-way tree withdrawal and conception; ii) multi-view CNN-based classification utilizing colorful snapshots; iii) multi-view CNN-based classification utilizing grey snapshots; iv) multi-view CNN-based classification utilizing binary snapshots in which the Bayesian optimization algorithms are utilized to distinguish the COPD for all four inputs. The outcome accuracy is (ACC) is 86.8%, 87.5%, and 86.7% for a colorful snapshot in the wake of casting a ballot accomplishes the ACC of 88.2%. Individually, grey, and binary snapshots achieve 88.6% and 86.4%. First, limitations or research gaps in the research manuscript focused on the smallness of the sample size since all patients are from one focus. Second, additional patients are selected from different communities for the upgradation of the datasets. Third, the CNN model uses 2D and 2.5D modalities. Now, 3D is used extensively, which renders it unfeasible. In this assessment, they focused on researching COPD and health conditions. Other than that, we have many airway illnesses like asthma, and pneumonia, resulting in a transition of gold stage 2 to 4. Fourth, the lung airway is locked here. In any case, lung parenchyma, instances of left atrial appendage (LAA), air catching, pneumonic veins, and clinical records (e.g., segment data, biomarkers, and some known clinical disorders) are rejected from the significant CNN model. The division of an airway is in a like manner and not agreeable.

Gupta *et al.* [31] expects to distinguish and classify lung CT pictures as healthy or unwell, such as COPD and fibrosis. Three processes are necessary to attain these objectives as it collects significant parameters from bronchial pictures, attribute assortment technique, as well as AI classifiers to predict lung infections. First, this study employs a methodology that segregates the Haralick consistency characteristics from a pool of features utilizing gray-level co-occurrence matrix (GLCM), Gabor features Zernike's stages, and Tamura disposition profiles. Next, three developmental subroutines are used as a characteristic determination method. This selects the most appropriate feature subset from a vast pool of clinical pictures to enhance the category accuracy and decrease the computational fees. Finally, 4 artificial intelligence ML classifiers: KNN, decision tree, SVM, and RF classifiers are implemented to each feature subset to determine the feature selection techniques. The outcomes show that the improvised crow search algorithm dispensed with the most significant and unimportant features by approximately 71%, and the improvised grey wolf algorithm eliminated just 52.3% of the all-out removed features. The improvised cuttlefish algorithm shifted through a minimal measure of elements. However, the improvised grey wolf algorithm provided a high-quality accuracy of 99.4% for classifying the respiratory illness observed through the improvised crow search algorithm with an accuracy of 99.0% separately.

The spatial distribution of COPD in the left and right lungs reflects the severity of airflow limitation [32]. Conventional strategies for finding COPD often require identifying the entire lung region, including the left and right lungs. If both the left lung and the right lung are utilized to recognize the spatial distribution of chronic obstructive pulmonary disease, some regions will not be regions of interest. Therefore, our article suggests a method to extract a single pulmonary lobe with Hounsfield units from the CT pulmonary images for diagnosis or teaching. After abstracting five pulmonary lobes, a specific pulmonary lobe is selected, observed, and analyzed for concentrating on the region of interest, such as locating the pulmonary disease [32].

COPD is a diverse infection with several classifications describing the major phenotypes and having an evaluation outline that provides ready-to-customize medication for COPD patients [33]. This COPD diagnosis necessitates the spirometric demonstration of the post-bronchodilator FEV1/FVC continuous airflow restriction. The classification enables the sorting of the patients in a definitive manner that is appropriate for anticipating symptoms, practical results, prognosis, and therapy response. Airflow restriction is the most common symptom of COPD. A basic strategy for COPD therapy was entirely focused on reducing FEV1 using the global initiative for chronic obstructive lung disease. The later GOLD rules "grade" COPD seriousness depends on the rate anticipated FEV1 that will keep away the disarray for a class dependent on either FEV1 decrease or exacerbation history. Independently, the severity of the dyspnea and the history of exacerbations are consolidated into a 2×2 grid, producing four "groups" [33].

The COPD results using ANN are presented. The dataset utilized in the examination includes 15 variables, four COPD disease levels (mild, moderate, severe, very severe), and information on 507 patients [34]. In this research, the investigation (MATLAB code with two hidden layers) of the result findings can be assessed using a five-layered cross-validation approach and a mean of 5-layer mistakes. The mean squared error (MSE) and mean absolute error (MAE) esteem were discovered to be 0,00996, and 0,02478 separately. As a result, the accuracy rate is determined to be 99%. Finally, the results establish that just 15 variables have very high rates of accuracy. As a result, it appears that ANN is effectively implemented in the categorization of lung diseases. The technology thus enables the clinician to determine the COPD degree rapidly and effectively.

QCT is successfully utilized to assess the respiratory organ parenchymatic patients with COPD illness [35]. The pathologic appearance of the respiratory organ pulmonary emphysema is briefly examined using image quantization, which needs a CT scan, Phantom CT, and image acquisition variables in CT. Pulmonary emphysema is depicted as a state of respiratory organs that is an extension of air gaps in the bronchioles. Centri-lobular emphysema is an initial kind of emphysema that is related to inflammation in the lungs and affects the respiratory bronchioles and the nearby air spaces. The second type of pulmonary emphysema is the pan-acinar pulmonary emphysema, which affects the entire respiratory organ acinus evenly. The third kind of pulmonary emphysema affects the alveolar system [35]. A multidetector computed tomography (MDCT) scanner with sixteen detectors is used in QCT investigations due to the duration of CT scanners with fewer channels. The QCT investigation does not exclusively aggregate the pulmonary emphysema load. However, alternative phenotypic parameters like provincial spread, airway illness, evaluation, and air trappings are also analyzed using expiratory CT images.

A hybrid approach combining probabilistic neural networks (PNN) and the principal component analysis (PCA) is utilized in the processed examination of the lung sounds to assist in the identification of concomitant pneumonia in COPD patients [36]. Fifty-eight COPD patients took part in the research (25 people were admitted to the hospital with community-acquired pneumonia and 33 for severe COPD exacerbation). The actual patients did auscultations on their suprasternal notch. The short-time Fourier transform (STFT) is utilized to distinguish the characteristics beginning with the recorded breathing sounds, and the PCA is used for spatial property reduction. In addition, the classifier resulted in the preparation of a PNN. The ten-fold cross-validation approach and a recipient working performance plot analysis are implemented to procure the system's functionality. This work achieves a specificity and sensitivity of 81.8 and 72%, respectively, using the proposed method. The findings show that the usage of electronic self-auscultation (suprasternal notch) at one site helps in the identification of pneumonia in individuals suffering from COPD.

The fastened flow of air restriction and ventilation non-uniformity is typical of chronic pulmonary infection. The conventional non-contrast CT offers airway and parenchyma inferences which cannot be utilized straightforwardly to decide lung function. The creator trains and tests the model using the CT surface examination and AI calculations to anticipate the lung ventilation non-uniformity in COPD. During this close examination, the members are randomly allocated for the optimization (m=1), testing (m=27), and training (m=67) datasets [21]. To establish ground truth labels, hyperpolarized (HP) helium-3 (HP3He) magnetic resonance imaging ventilation maps are combined with CT thorax scans [21]. Eighty-seven quantitative imaging highlights are extracted and standardized according to the respiratory organ midpoints, which came up with 174 highlights. To limit the number of features, the forward feature selection was used. The fivefold cross-validation was used to train quadratic SVM classifiers, linear SVM, and logistic regression. The test data set was subjected to the application of the best-performing classification model. The correlation between the model, magnetic resonance imaging, and lung function data was determined using the Pearson coefficients. The quantum enhanced SVM (QSVM) outperformed the others in training and is now being utilized on the test dataset. The model's predictable ventilation maps showed 88% accuracy and an area under the curve (AUC) of 0.82 using the HP3He MRI ventilation map as the reference standard.

As a result of the neural network's training, the method for categorizing COPD and asthma used fuzzy criteria [37]. This technique utilizes a neuro-fuzzy system to classify asthma and COPD diseases with neural network (NN) parameters outlined in line with the guidelines of the genetic information nondiscrimination act (GINA) and global initiative for chronic obstructive lung disease (GOLD) rules. Asthma is an inflammatory airway impairment that leads to hyper-reactive and increased mucous, mucosal edema, and airway muscle spasm, leading to airway obstruction. Spirometry and impulse oscillometry system are the two most frequent pulmonary function tests used for detecting asthma and COPD [37]. For NN training, doctors from the Sarajevo University Clinical Centre utilized the care-fusion database. The system was tested on 455 patients. 99.41% of 170 asthma patients were accurately identified. Also, 99.19% of 248 patients with COPD were precisely categorized. The system model was shown to be successful after being implemented on 37 people with healthy lungs. In classifying COPD and asthma patients, the impulse oscillometry system (IOS) check data provide a sensitivity of 99.28% and a specificity of 100%.

COPD is analyzed using patches of CT pictures which contribute to the overall diagnosis [38]. The image labels are used for labeling healthy patches in COPD patients. At this point, the measurement of COPD from lung pictures as a multiple instance learning (MIL) strategy is a better fit for such sparsely labeled data. The different MIL assumptions of the pulmonary disease show that the idea area is contrasted. COPD-related infection designs contribute to the consideration of the entire distribution of lung tissue patches for improved performance.

The use of paired CT lung images of the metric response mapping (MRM) has improved the phenotyping of COPD [39]. The huge changeability in the typical approach for evaluating PRM functional small airways disease (fSAD) uses the three-dimensional data from the PRM fSAD scalar variable (PRM fSAD relative ability), distorting the initial three-dimensional information to distinguish the movement of the COPD as differing subtypes. In this investigation, another methodology was created to examine the PRM which relies on topological procedures. 3D contours were produced using the regional statistical parameters. The surface area was the most essential and significant independent predictor with clinically substantial COPD proportions. These discoveries establish the suitability of our method in the utilization of different COPD qualities. The established method gave considerable freedom to arrive at improvised conclusions in COPD patients.

The same discovery of COPD subgroups using imaging biomarkers is critical for personalized therapy [40]. A PRM was used to analyze the COPD phenotype using a voxel-based image processing method. This examines the entire respiratory organ estimated tomography scans from 194 patients with COPD using COPDGene research. The PRM was aware of the magnitude of emphysema. The PRM offered CT evidence that preceded the respiratory disorder with the growing severity of COPD. PRM is an adaptable bio-indicator for imagery, detecting ailments and configurations. It gives dimensional information regarding conditions and spaces. The capacity of the PRM to distinguish among particular COPD configurations enables the correct identification of individual patients by complementing quality clinical techniques.

According to experimental studies, ADO (alveolar gas-arterial blood oxygenation) was demonstrated to be better in predicting COPD-related mortality in primary care [41]. COPD patients were readmitted to the hospital within 30 days more than 89% of the time when SVM was used [42]. It is highly vital to diagnose COPD early in a curable stage and even save a patient's life. The work in this paper focuses on the knowledge graph [43]. The method in this paper includes building a COPD knowledge graph, using the CMFS-η algorithm, which chooses multidimensional features and creates a hybrid decision model differential search algorithm (DSA)-SVM. The performance of this method is obtained at 95.01% accuracy using the DSA-SVM algorithm. Hence, the DSA-SVM algorithm is better than the standard SVM algorithm.

Serum metabolic biomarkers can be used to diagnose COPD effectively [44]. The research outlined in this paper is focused on serum metabolic biomarkers, which state that polynomial and linear least-squares support the vector machine models. The method in this paper includes least-squares SVM (LS-SVM) and serum metabolic biomarkers. The main focus of the current study is to build an integrated AI approach for the supplemental diagnosis of COPD and bio fluid-based biomarkers. The performance of this method is obtained at 80.77% and 84.62% of accuracy values, and the AUC values at 0.87 and 0.90. The research established that the LS-SVM classifier is better than the radial basis function (RBF) classifier.

To determine the best model for another research, the classification of acute chronic obstructive pulmonary disease (AECOPDs) may be broadly classified according to patient symptoms [45], [46]. AECOPD detection models were developed to analyze the effectiveness of different modeling archetypes [47]. The research mentioned in this paper focused on identifying a robust method for the acute diagnosis of COPD. The paper includes five equipment erudition subroutines which were KNN, logistic regression, RF, and naive Bayes. For discovering the best hyper-parameters for each model, ten-fold cross-validation was utilized. As a result of this SVM method, the region under the recipient working distinctive curvature was 0.90, the specificity at 0.83, sensitivity at 0.80, the perquisite prognostic number at 0.81, and the contrary prognostic number at 0.85. This result was derived from a total of 303 Electronic medical records (135 AECOPDs patients and 168 non-AECOPDs patients).

Early identification of COPD can lead to timely treatment and a lower mortality rate. After analyzing 75 patient records, researchers used quantitative computed tomography and five partial ML models to predict COPD, with the KNN model, polynomial regression, and gradient boosting achieving the lowest mean relative error: 16% [48], [49]. The research outlined focuses on a multistage ensemble model to help clinicians deliver the appropriate medication at the right time. The methodology explained in this paper includes the genetic algorithm, KNN, and light gradient boosting machine (GBM)–recursive feature elimination (RFE) algorithm. Here, the multistage ensemble model (MSEM) utilized the Exasens datasets. The proposed technique performed well at precision values of 0.9800, F1-measure values of 0.9667, AUC values of 0.9912, recall values of 0.9600, and accuracy values of 0.9820. Hence, the proposed model was better than the other machine learning models due to its excellent performance. However, many neural networks can be combined to provide better classification. COPD patients who were infected with COVID-19 had greater rates of hospitalization and death, which were mostly due to pneumonia [50].

3. REVIEW OF METHODS

Based on the above-related works, it is evident that several techniques and algorithms were available for COPD analysis. Some techniques are more suitable for COPD identification. The analyzed literature mentioned the conventional machine learning techniques and deep learning models explored by various researchers. The major papers involving the detection of COPD using different machine learning and deep learning methods are listed below. As shown in Table 1, there are numerous approaches for classifying COPD with different datasets.

Table 1. The techniques used to detect COPD using different datasets

Paper	Technique (Classification)	Datasets
[1]	Naive Bayes Classifier	56 thoracic CT images
[26]	3D CNN + Garbage Collection Optimization Algorithm and RFs Classifier	18 CT images
[30]	Bayesian Optimization Algorithm	280 CT images
[25]	Deep Belief Network Algorithm; Hilbert–Huang Transform	Multichannel lung sound
[31]	Improvised Cuttlefish Algorithm Improvised Crow Search Algorithm Improvised Grey Wolf Algorithm	36- CT image
[29]	XGBoost Gradient Boosting Algorithm (SVM, Gaussian Naive Bayes, Logistic Regression, and ANN)	Saliva samples
[3]	SVM KNN	Spirometric data
[21]	QSVM	CT images
[16]	RFs Classifier	132 samples
[44]	Linear SVM Polynomial SVM	Serum metabolic biomarkers based on 26 subjects
[49]	Multistage Ensemble Learning Approach Light GBM Algorithm	Exasens datasets-239 saliva samples
[39]	3D CNN	596- CT images
[28]	2D CNN	COPD gene datasets 1000 CT images
[22]	Maximum-likelihood, Latent Dirichlet allocation, neural network, GA	600- Chest X-ray

4. RESULTS AND DISCUSSION

The tables and figures reveal the summary and findings from the related works of COPD analysis and diagnosis shown below. Various datasets and parameters are used for assessment. Metrics of evaluation are based on the rate of accuracy or classification. Table 2 shows the relative accuracy of several methods. Figure 2 shows a comparison of accuracy achieved by using various techniques. Table 3 shows the sensitivity and specificity. Table 4 shows the AUC comparison.

Table 2. Comparative analysis of existing COPD classification techniques

Paper	Technique	Accuracy (%)
[1]	Naive Bayes Classifier	84.00
[26]	3D CNN + Garbage Collection Optimization Algorithm and RFs Classifier	94.00
[30]	Bayesian Optimization Algorithm	88.60
[25]	Deep Belief Network Algorithm; Hilbert–Huang Transform	93.67
[31]	Improvised Cuttlefish Algorithm	97.30
	Improvised Crow Search Algorithm	99.00
	Improvised Grey Wolf Algorithm	99.40
[29]	XGBoost Gradient Boosting Algorithm (SVM, Gaussian Naive Bayes, Logistic Regression, and ANN)	91.25
[3]	SVM	96.97
	KNN	92.30
[21]	QSVM	88.00
[16]	RFs Classifier	97.70
[44]	Linear SVM	80.77
	Polynomial SVM	84.62
[49]	Multistage Ensemble Learning Approach	98.20
	Light GBM Algorithm	
[39]	3D CNN	89.30
[28]	2D CNN	77.30
[22]	GA	97.9

Table 3. Sensitivity and specificity

Paper	Sensitivity (%)	Specificity (%)
[25]	91.00	96.33
[29]	100.00	88.89
[36]	72.00	81.80
[37]	99.28	100.00
[47]	83.00	80.00

Table 4. AUC

Paper	AUC
[21]	0.82
[38]	0.776
[44]	0.87 (Linear SVM)
	0.90 (Polynomial SVM)
[47]	0.90
[49]	0.9912

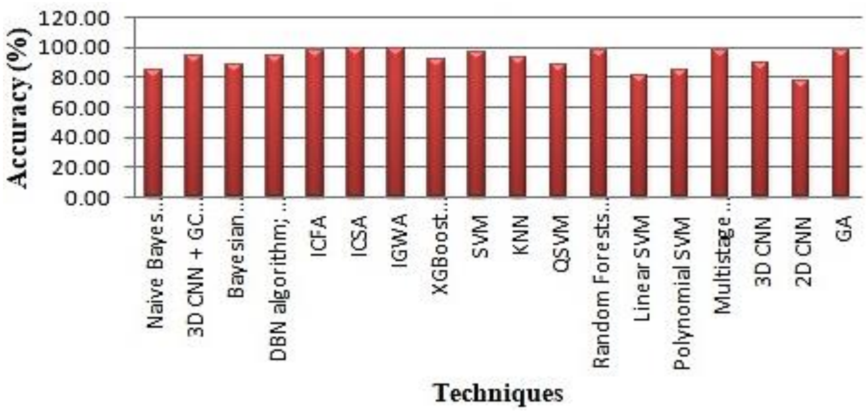

Figure 2. Accuracy comparison

5. CONCLUSION

The researchers mentioned establishing COPD as a long-term lung disease that is principally caused due to smoking, inhaling toxic substances, and pollutants. An expert resolution is needed after a COPD diagnosis. COPD is classified into two types: pulmonary emphysema and chronic bronchitis. The inflammatory substances damage the alveolar walls due to pulmonary emphysema, which destroys the air sacs. Shortness of breath and wheezing are the common symptoms of pulmonary emphysema. Chronic bronchitis is caused due to air-duct tapering as a result of inflammation. People affected with chronic bronchitis are likely to experience chest tightness, frequent coughing, and mucus secretion excretion. The positive note is that COPD can be prevented and cured using long-term regulated treatments and drastic lifestyle changes. Within the deep learning technique, CNN is established with successful outcomes in medical imaging classification. This review paper explores the relevant analysis done related to respiratory organ diseases, asthma, emphysema, COPD, phenotype, biomarkers, pulmonary lobe, ANN, and machine learning ideas.

6. FUTURE WORK

During the COVID-19 pandemic, the important points for managing stable COPD were outlined. First, precautionary measures like fundamental infection control procedures, putting on a face mask, and considering a safe-haven or shelter-in-place alert need to be established. Second, the necessary spirometry should be used for inquiry. Third, enough pharmaceutical supplies should be maintained during pharmacological treatment. Fourth, physical activity should be done, and a yearly flu injection should be taken while on non-pharmacological therapy.

COVID-19 is an excellent realm for further research, which may provide insight into COPD due to the increased patient life risk. More prescriptions and prompt treatment for recovery are still required due to COVID-19. As a result of SARS-CoV-2 infection, patients with COPD are more vulnerable. Because of the restricted airways, COPD may make it more challenging to fight the coronavirus. The three main symptoms to look for while diagnosing COVID-19 are fever, coughing, and shortness of breath. COPD symptoms include coughing and shortness of breath but do not generally result in a fever. Physiologists use the real time polymerase chain reaction (PCR) test to confirm the diagnosis and therapy if the fever is 100.4 F or above, in addition to the other COVID-19 symptoms. Without expert help, around 80% of patients recover completely from COVID-19. On the other hand, COPD puts the patient at a higher chance of becoming unwell.

REFERENCES

[1] H. Moghadas-Dastjerdi, M. Ahmadzadeh, E. Karami, M. Karami, and A. Samani, "Lung CT image based automatic technique for COPD GOLD stage assessment," *Expert Systems with Applications*, vol. 85, pp. 194–203, Nov. 2017, doi: 10.1016/j.eswa.2017.05.036.

[2] D. M. G. Halpin *et al.*, "Global initiative for the diagnosis, management, and prevention of chronic obstructive lung disease. The 2020 GOLD science committee report on COVID-19 and chronic obstructive pulmonary disease," *American Journal of Respiratory and Critical Care Medicine*, vol. 203, no. 1, pp. 24–36, Jan. 2021, doi: 10.1164/rccm.202009-3533SO.

[3] S. Vora and S. Chintan, "COPD classification using machine learning algorithms," *International Research Journal of Engineering and Technology*, vol. 6, pp. 608–611, 2008.

[4] D. P. Rao, A. Talatam, B. Chakradhar, K. Bhargavi, and A. Bhagyaraj, "High resolution computed tomography in chronic obstructive pulmonary disease," *International Journal of Advances in Medicine*, vol. 5, no. 5, Sep. 2018, doi: 10.18203/2349-3933.ijam20183898.

[5] J. I. Stewart and G. J. Criner, "The small airways in chronic obstructive pulmonary disease," *Current Opinion in Pulmonary Medicine*, vol. 19, no. 2, pp. 109–115, Mar. 2013, doi: 10.1097/MCP.0b013e32835ceefc.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[7] N. Raju, A. H. B., and P. Augustine, "Identification of interstitial lung diseases using deep learning," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 6, pp. 6283–6291, Dec. 2020, doi: 10.11591/ijece.v10i6.pp6283-6291.

[8] Y. H. Cho *et al.*, "Quantitative assessment of pulmonary vascular alterations in chronic obstructive lung disease: Associations with pulmonary function test and survival in the KOLD cohort," *European Journal of Radiology*, vol. 108, pp. 276–282, Nov. 2018, doi: 10.1016/j.ejrad.2018.09.013.

[9] C. Hatt, C. Galban, W. Labaki, E. Kazerooni, D. Lynch, and M. Han, "Convolutional neural network based COPD and emphysema classifications are predictive of lung cancer diagnosis," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2018, pp. 302–309. doi: 10.1007/978-3-030-00946-5_30.

[10] L. Peng *et al.*, "Classification and quantification of emphysema using a multi-scale residual network," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 6, pp. 2526–2536, Nov. 2019, doi: 10.1109/JBHI.2018.2890045.

[11] W. L. Bi *et al.*, "Artificial intelligence in cancer imaging: Clinical challenges and applications," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 2, Feb. 2019, doi: 10.3322/caac.21552.

[12] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine*, vol. 380, no. 14, pp. 1347–1358, Apr. 2019, doi: 10.1056/NEJMra1814259.

[13] S. Soffer, A. Ben-Cohen, O. Shimon, M. M. Amitai, H. Greenspan, and E. Klang, "Convolutional neural networks for radiologic images: a radiologist's guide," *Radiology*, vol. 290, no. 3, pp. 590–606, Mar. 2019, doi: 10.1148/radiol.2018180547.

[14] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, no. 1995, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[15] O. Er and F. Temurtas, "A study on chronic obstructive pulmonary disease diagnosis using multilayer neural networks," *Journal of Medical Systems*, vol. 32, no. 5, pp. 429–432, Oct. 2008, doi: 10.1007/s10916-008-9148-6.

[16] D. Spathis and P. Vlamos, "Diagnosing asthma and chronic obstructive pulmonary disease with machine learning," *Health Informatics Journal*, vol. 25, no. 3, pp. 811–827, Sep. 2019, doi: 10.1177/1460458217723169.

[17] A. A. Abdulmunem, Z. A. Abutiheen, and H. J. Aleqabie, "Recognition of corona virus disease (COVID-19) using deep learning network," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 1, pp. 365–374, Feb. 2021,

doi: 10.11591/ijece.v11i1.pp365-374.
[18] J. Nagao *et al.*, “A CAD system for quantifying COPD based on 3-D CT images,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 2878, pp. 730–737, 2003, doi: 10.1007/978-3-540-39899-8_89.
[19] S. T. H. Kieu, A. Bade, M. H. A. Hijazi, and H. Kolivand, “A survey of deep learning for lung disease detection on medical images: state-of-the-art, taxonomy, issues and future directions,” *Journal of Imaging*, vol. 6, no. 12, Dec. 2020, doi: 10.3390/jimaging6120131.
[20] N. Alrefai and O. Ibrahim, “Deep learning for COVID-19 diagnosis based on chest X-ray images,” *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 11, no. 5, pp. 4531–4541, Oct. 2021, doi: 10.11591/ijece.v11i5.pp4531-4541.
[21] A. Westcott, D. P. I. Capaldi, D. G. McCormack, A. D. Ward, A. Fenster, and G. Parraga, “Chronic obstructive pulmonary disease: thoracic CT texture analysis and machine learning to predict pulmonary ventilation,” *Radiology*, vol. 293, no. 3, pp. 676–684, Dec. 2019, doi: 10.1148/radiol.2019190450.
[22] P. Bhuvaneswari and A. B. Therese, “Feature extraction and classification of COPD chest X-ray images,” *International Journal of Computer Aided Engineering and Technology*, vol. 12, no. 3, 2020, doi: 10.1504/IJCAET.2020.106212.
[23] Q. Xu *et al.*, “Intelligent syndrome differentiation of traditional chinese medicine by ANN: a case study of chronic obstructive pulmonary disease,” *IEEE Access*, vol. 7, pp. 76167–76175, 2019, doi: 10.1109/ACCESS.2019.2921318.
[24] J. D. Newell, “Quantitative computed tomography of lung parenchyma in chronic obstructive pulmonary disease: an overview,” *Proceedings of the American Thoracic Society*, vol. 5, no. 9, pp. 915–918, Dec. 2008, doi: 10.1513/pats.200804-034QC.
[25] G. Altan, Y. Kutlu, and N. Allahverdi, “Deep learning on computerized analysis of chronic obstructive pulmonary disease,” *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 5, pp. 1344–1350, May 2020, doi: 10.1109/JBHI.2019.2931395.
[26] P. Nardelli *et al.*, “Pulmonary artery-vein classification in CT images using deep learning,” *IEEE Transactions on Medical Imaging*, vol. 37, no. 11, pp. 2428–2440, Nov. 2018, doi: 10.1109/TMI.2018.2833385.
[27] B. A. Hoff *et al.*, “CT-based local distribution metric improves characterization of COPD,” *Scientific Reports*, vol. 7, no. 1, Dec. 2017, doi: 10.1038/s41598-017-02871-1.
[28] G. González *et al.*, “Disease staging and prognosis in smokers using deep learning in chest computed tomography,” *American Journal of Respiratory and Critical Care Medicine*, vol. 197, no. 2, pp. 193–203, Jan. 2018, doi: 10.1164/rccm.201705-0860OC.
[29] P. S. Zarrin, N. Roeckendorf, and C. Wenger, “In-vitro classification of saliva samples of COPD patients and healthy controls using machine learning tools,” *IEEE Access*, vol. 8, pp. 168053–168060, 2020, doi: 10.1109/ACCESS.2020.3023971.
[30] R. Du *et al.*, “Identification of COPD from multi-view snapshots of 3D lung airway tree via deep CNN,” *IEEE Access*, vol. 8, pp. 38907–38919, 2020, doi: 10.1109/ACCESS.2020.2974617.
[31] N. Gupta, D. Gupta, A. Khanna, P. P. Rebouças Filho, and V. H. C. de Albuquerque, “Evolutionary algorithms for automatic lung disease detection,” *Measurement*, vol. 140, pp. 590–608, Jul. 2019, doi: 10.1016/j.measurement.2019.02.042.
[32] Y. Yang, Y. Guo, J. Guo, Y. Gao, and Y. Kang, “A method of abstracting single pulmonary lobe from computed tomography pulmonary images for locating COPD,” in *Proceedings of the Fourth International Conference on Biological Information and Biomedical Engineering*, Jul. 2020, pp. 1–6. doi: 10.1145/3403782.3403805.
[33] P. Manian, “Chronic obstructive pulmonary disease classification, phenotypes and risk assessment,” *Journal of Thoracic Disease*, vol. 11, no. 14, pp. 1761–1766, Sep. 2019, doi: 10.21037/jtd.2019.05.10.
[34] H. Isik, U., Guven, A., Buyukoglan, “2015 Medical Technologies National Conference, TIPTEKNO 2015,” in *2015 Medical Technologies National Conference, TIPTEKNO 2015*, 2015, pp. 1–4.
[35] J. Gawlitza *et al.*, “Predicting pulmonary function testing from quantified computed tomography using machine learning algorithms in patients with COPD,” *Diagnostics*, vol. 9, no. 1, Mar. 2019, doi: 10.3390/diagnostics9010033.
[36] D. Sanchez Morillo, A. Leon Jimenez, and S. A. Moreno, “Computer-aided diagnosis of pneumonia in patients with chronic obstructive pulmonary disease,” *Journal of the American Medical Informatics Association*, vol. 20, no. 1, pp. 111–117, Jun. 2013, doi: 10.1136/amiajnl-2012-001171.
[37] A. Badnjevic, M. Cifrek, D. Koruga, and D. Osmankovic, “Neuro-fuzzy classification of asthma and chronic obstructive pulmonary disease,” *BMC Medical Informatics and Decision Making*, vol. 15, Dec. 2015, doi: 10.1186/1472-6947-15-S3-S1.
[38] V. Cheplygina, L. Sorensen, D. M. J. Tax, J. H. Pedersen, M. Loog, and M. de Bruijne, “Classification of COPD with multiple instance learning,” in *2014 22nd International Conference on Pattern Recognition*, Aug. 2014, pp. 1508–1513. doi: 10.1109/ICPR.2014.268.
[39] T. T. Ho *et al.*, “A 3D-CNN model with CT-based parametric response mapping for classifying COPD subjects,” *Scientific Reports*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-020-79336-5.
[40] C. J. Galbán *et al.*, “Computed tomography-based biomarker provides unique signature for diagnosis of COPD phenotypes and disease progression,” *Nature Medicine*, vol. 18, no. 11, pp. 1711–1715, Nov. 2012, doi: 10.1038/nm.2971.
[41] J. Li and Y. Hong, “Crackles detection method based on time-frequency features analysis and SVM,” in *2016 IEEE 13th International Conference on Signal Processing (ICSP)*, Nov. 2016, vol. 0, no. 1, pp. 1412–1416. doi: 10.1109/ICSP.2016.7878059.
[42] M. A. Hakim, F. Garden, M. Jennings, and C. Dobler, “Performance of the LACE index to predict 30-day hospital readmissions in patients with chronic obstructive pulmonary disease,” *Clinical Epidemiology*, vol. 10, pp. 51–59, Dec. 2017, doi: 10.2147/CLEP.S149574.
[43] Y. Fang, H. Wang, L. Wang, R. Di, and Y. Song, “Diagnosis of COPD based on a knowledge graph and integrated model,” *IEEE Access*, vol. 7, pp. 46004–46013, 2019, doi: 10.1109/ACCESS.2019.2909069.
[44] H. Zheng *et al.*, “Predictive diagnosis of chronic obstructive pulmonary disease using serum metabolic biomarkers and least-squares support vector machine,” *Journal of Clinical Laboratory Analysis*, vol. 35, no. 2, pp. 1–8, Feb. 2021, doi: 10.1002/jcla.23641.
[45] S. A. Shah, C. Velardo, A. Farmer, and L. Tarassenko, “Exacerbations in chronic obstructive pulmonary disease: identification and prediction using a digital health system,” *Journal of Medical Internet Research*, vol. 19, no. 3, Mar. 2017, doi: 10.2196/jmir.7207.
[46] R. Pauwels *et al.*, “COPD exacerbations: the importance of a standard definition,” *Respiratory Medicine*, vol. 98, no. 2, pp. 99–107, Feb. 2004, doi: 10.1016/j.rmed.2003.09.001.
[47] C. Wang, X. Chen, L. Du, Q. Zhan, T. Yang, and Z. Fang, “Comparison of machine learning algorithms for the identification of acute exacerbations in chronic obstructive pulmonary disease,” *Computer Methods and Programs in Biomedicine*, vol. 188, May 2020, doi: 10.1016/j.cmpb.2019.105267.
[48] J. Gawlitza *et al.*, “Predicting Pulmonary function testing from quantified computed tomography using machine learning algorithms in patients with COPD,” *Diagnostics*, vol. 9, no. 1, Mar. 2019, doi: 10.3390/diagnostics9010033.

[49] J. Dhar, "Multistage ensemble learning model with weighted voting and genetic algorithm optimization strategy for detecting chronic obstructive pulmonary disease," *IEEE Access*, vol. 9, pp. 48640–48657, 2021, doi: 10.1109/ACCESS.2021.3067949.

[50] D. Graziani *et al.*, "Characteristics and prognosis of COVID-19 in patients with COPD," *Journal of Clinical Medicine*, vol. 9, no. 10, Oct. 2020, doi: 10.3390/jcm9103259.

BIOGRAPHIES OF AUTHORS

Ramadoss Ramalingam graduated in Electronics and Communication Engineering from Maharaja Engineering College, Avinashi affiliated to Anna University Chennai, India in the year 2009. He obtained his M.E degree in Applied Electronics from Velammal Engineering College, affiliated with Anna University Chennai, India in the year 2012. He is currently pursuing a Ph.D. degree in Electronics and Communication Engineering from SRM Institute of Science and Technology, Kattankulathur, India. His current research area is image processing and deep learning. He can be contacted at rr5606@srmist.edu.in.

Vimala Chinnaiyan received a B.E (Electronics and Communication Engineering), M.E (Applied Electronics), and degrees from Bharathidasan University-Trichy, Anna University Chennai, India in 2003 and 2008 respectively. She received her Ph.D. (Electronics and Communication Engineering) at SRM Institute of Science and Technology, 603203, Kattankulathur, India. Currently, she is an assistant professor in the Department of Electronics and Communication Engineering at SRM Institute of Science and Technology, India. Her research interests are image processing and signal processing. She can be contacted at vimalac@srmist.edu.in.