Computer Systems Science & Engineering
DOI: 10.32604/csse.2023.030491
Article

Tech Science Press

Optimal Deep Belief Network Based Lung Cancer Detection and Survival Rate Prediction

Sindhuja Manickavasagam[1,*] and Poonkuzhali Sugumaran[2]

[1]Department of Information Technology, Rajalakshmi Engineering College, Chennai, 600125, Tamilnadu, India
[2]Department of Computer Science and Engineering, Rajalakshmi Engineering College, Chennai, 600125, Tamilnadu, India
*Corresponding Author: Sindhuja Manickavasagam. Email: sindhujavignesh06@gmail.com

Received: 27 March 2022; Accepted: 19 May 2022

Abstract: The combination of machine learning (ML) approaches in healthcare is a massive advantage designed at curing illness of millions of persons. Several efforts are used by researchers for detecting and providing primary phase insights as to cancer analysis. Lung cancer remained the essential source of disease connected mortality for both men as well as women and their frequency was increasing around the world. Lung disease is the unrestrained progress of irregular cells which begin off in one or both Lungs. The previous detection of cancer is not simpler procedure however if it can be detected, it can be curable, also finding the survival rate is a major challenging task. This study develops an Ant lion Optimization (ALO) with Deep Belief Network (DBN) for Lung Cancer Detection and Classification with survival rate prediction. The proposed model aims to identify and classify the presence of lung cancer. Initially, the proposed model undergoes min-max data normalization approach to preprocess the input data. Besides, the ALO algorithm gets executed to choose an optimal subset of features. In addition, the DBN model receives the chosen features and performs lung cancer classification. Finally, the optimizer is utilized for hyperparameter optimization of the DBN model. In order to report the enhanced performance of the proposed model, a wide-ranging experimental analysis is performed and the results reported the supremacy of the proposed model.

Keywords: Lung cancer; feature selection; ant lion optimization; classification; disease diagnosis; metaheuristics

1 Introduction

Lung cancer (LC) is the major and primary reason for cancer death in both men and women. Demonstration of LC in the body parts of the patient exposes via earlier symptoms in several persons [1]. Treatments are undergone and projection rest on the diagnosis types of cancers, the stages (extent of spreading), and victim outcome status. Likely medication contains surgery, radiotherapy, and chemotherapy Persistence relies upon the stages, overall health of a person, and other determiners, however entirely only 14% of an individual recognized with LC live 5 years after the recognition [2]. The death rate and prevalence of tobacco consumption are comparatively high. Generally, LC grows inside the wall or epithelium of the respiratory tree [3]. However, it could be started anywhere else in the lung region and affect other parts of the lung system. LC is most probably affected persons between the ages of 55 and 65 and frequently takes several years for development [4]. There are 2 major classifications of LC one is small cell LC (SCLC) or oat cell cancer and another one is Non-small cell LC (NSCLC).

Every type of LC develops and spreads in different means and it is treated accordingly [5]. In case of cancer has characteristics of both verities, it is known as mixed small cell or large cell cancers. Non-small cell LC is more common than SCLC and it has usual growth and also spreads more gradually [6]. SCLC is nearly associated with smoking and growing rapidly and forms huge tumors which may spread extremely all over the body. They frequently started in the bronchi near the centre of the chest. LC death rate is associated with the total number of cigarettes consumed [7]. Pre-diagnosis aids us to recognize or narrow down the probability of screening for LC disease. Signs of LC and risk aspects (obesity, smoking, alcohol drinking, and insulin resistance) had a statistically important effect in early diagnosis phase [8]. The LC diagnostic and prognostic problems are mostly in the scope of the extremely discussed grouping problems [9,10]. These issues have inspired many of the investigators in statistics domain, computer intelligence, and data mining.

In [11], a k-Nearest-Neighbor method, where a genetic algorithm is employed to the effectual feature election to minimize the data dimension and augment classification performance, is applied for identifying the phase of patient disease. To enhance the performance of the presented method, the optimal value for k is defined by the experiment technique. Agrawal et al. [12] examine the lung tumor dataset obtainable from the SEER by proposing precise survival predictive model for lung cancer with data mining technique. Then, the developed pre-processing step led to splitting or removal, or modification of attribute, and 2 of the 11 derived attributes have been found to have important prediction energy. Data mining classification technique is utilized on the pre-processed dataset and validation and data mining optimization. In [13], appropriate integration of Adoptive thresholding segmentation approach was employed to segment input images, a familiar Support Vector Machine (SVM) image classification approach was employed for classifying lung cancer and Content-based image retrieval method is utilized for comparing lung image features namely intensity, contract, shape, and texture. Lim et al. [14] proposed a bioinformatics pipeline and standardized data that is pre-processed by a strong statistical method; allows other to implement largescale meta-analysis, without conducting statistical correction and time-consuming data mining. [15] explored the part of Chinese prescription in non-small cell lung cancer (NSCLC) and offer reference for the prescription and herbs applications. Randomized and quasi-randomized controlled medical trials on Chinese herbal medication in the treatment of NSCLC have been gathered from 7 datasets to determine a dataset of prescriptions on NSCLC.

This study develops a Ant lion Optimization with Optimal Deep Belief Network (ODBN) model named ALO-ODBN for Lung Cancer Detection and Classification. Initially, the ALO-ODBN model undergoes min-max data normalization approach to preprocess the input data. Besides, the ALO algorithm gets executed to choose an optimal subset of features. In addition, the DBN model receives the chosen features and performs lung cancer classification. Finally, the Adam optimizer is utilized for hyperparameter optimization of the DBN model. In order to report the enhanced performance of the ALO-ODBN model, a wide-ranging experimental analysis is performed and the results reported the supremacy of the ALO-ODBN model.

2 The Proposed Model

In this study, a new ALO-ODBN model has been developed for Lung Cancer Detection and Classification. Initially, the ALO-ODBN model undergoes min-max data normalization approach to preprocess the input data. Besides, the ALO algorithm gets executed to choose an optimal subset of features. In addition, the DBN model receives the chosen features and performs lung cancer classification.

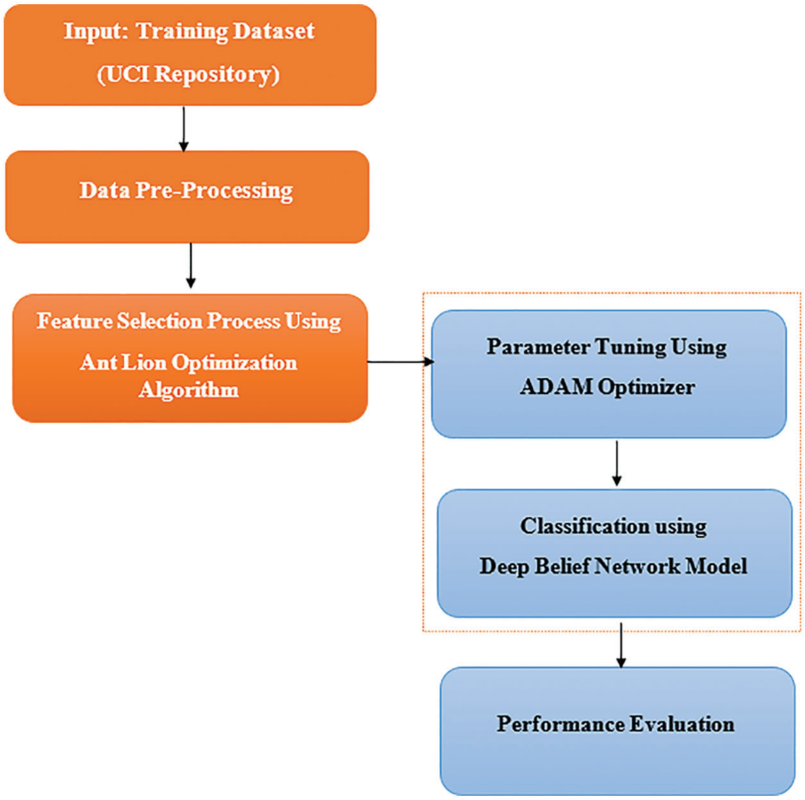

Figure 1: Overall work flow of ALO-ODBN technique

Finally, the Adam optimizer is utilized for hyperparameter optimization of the DBN model. Fig. 1 depicts the overall work flow of ALO-ODBN technique.

2.1 Pre-Processing

Initially, the ALO-ODBN model undergoes min-max data normalization approach to pre-process the input data. In any ML model, data normalization is widely utilized to attain proficient results [16]. The features values can different from small to large values. So, the normalization process is employed for scaling the features as given below.

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} \tag{1}$$

2.2 Process Involved in ALO Technique

The recent meta heuristic method of ant lion optimization algorithm is the interaction of antlions and ants. mall cone shaped traps are made by the antlions, in which they hide and wait for their prey. The major steps involved in the tuning process are illustrated below with the help of steps given below

Ant lion Algorithm

Stage 1: At the first step the location of the ants and ant lions are generated
randomly.

(Continued)

Ant lion Algorithm (continued)

Stage 2: Decide the appropriateness value for each ant and antlion.

Stage 3: Select the elite antlion using roulette wheel.

Stage 4: If end criteria is not met modify the values of inferior and superior using iterations.

Stage 5: Standardize the random walk.

Step 6: Substitute the antlion with ant if it becomes fitter.

Stage 7: If the antlion develops fitter than the elite then update the elite.

Stage 8: Return the elite if the ending criterion is satisfied

The ant lion optimization algorithm simulates the hunting mechanism of the antlions. The following subsections describe the steps of the algorithm.

2.3 Random Walk of Ants

The random walk of ants, where, *csum* calculates the cumulative sum of successive random jumps leading to a random walk, $s(\delta)$ is a stochastic function and is a random number generated with uniform distribution in the interval of [0, 1] and t denotes the max mum number of iterations.in blow Fig. 2 represent that the flow chart for ALO algorithm

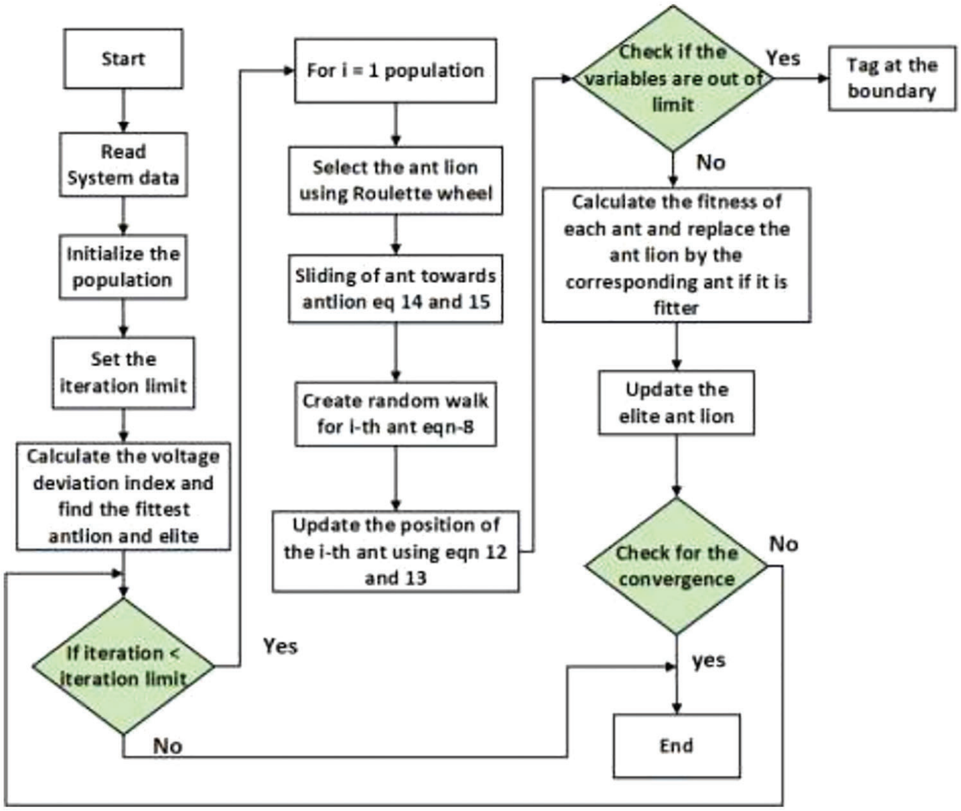
Start
Read System data
Initialize the population
Set the iteration limit
Calculate the voltage deviation index and find the fittest antlion and elite
If iteration < iteration limit
Yes
No
For i = 1 population
Select the ant lion using Roulette wheel
Sliding of ant towards antlion eq 14 and 15
Create random walk for i-th ant eqn-8
Update the position of the i-th ant using eqn 12 and 13
Check if the variables are out of limit
Yes
Tag at the boundary
No
Calculate the fitness of each ant and replace the ant lion by the corresponding ant if it is fitter
Update the elite ant lion
Check for the convergence
No
yes
End

Figure 2: Flow chart for Ant lion optimization algorithm

$$\mathrm{Y}(\delta) = [0,\ \mathrm{csum}(2\mathrm{s}(\delta 1) - 1,\ \mathrm{csum}(2\mathrm{s}(\delta) - 1),\ \ldots \mathrm{csum}(2\mathrm{s}(\delta 2) - \tag{2}$$

The fitness values for the ants are stored in the form of a matrix and are a function of the objective function. In the same manner the fitness values of the antlions are also stored in another matrix. During each step f the iteration, the random walk of the ants is to be confined within the boundary of the search space. This is executed by Eq. (3), where *mi* and *ni* are the minimum and maximum step of the random walk of the *i*th ant during the *t*th iteration and *C* and *D* are the inferior and superior restraints of the random walk.

$$\mathrm{yit} = (\mathrm{yit} - \mathrm{mi}) \times (\mathrm{Di} - \mathrm{Cit})\mathrm{nit} - \mathrm{mi} + \mathrm{Ci} \tag{3}$$

2.4 Falling of Ant Towards Antlion

The ants randomly walk without seeing the trap, so, it may fall down into the cone shaped trap. This is realized by adaptively decreasing the radius of the random walk as shown.

$$\mathrm{Ct} + 1 = \mathrm{CtI} \tag{4}$$

$$\mathrm{Dt} + 1 = \mathrm{DtI} \tag{5}$$

where, Ct is the minimum value corresponding to *t*th iteration and *Dt* is the corresponding maximum value, *i* is a ratio given by Eq. (5).

$$\mathrm{I} = 10\mathrm{wtT} \tag{6}$$

Here, *t* is the current iteration and *T* is the maximum iteration and *w* is a constant.

Trapping of ant by antlion

The range of the random walk of the **i**th ant during **t**th iteration is modified as below in order to model the trapping behavior.

$$\mathrm{Ct} + 1 = \mathrm{aliont\ j} + \mathrm{Ct} \tag{7}$$

$$\mathrm{Dt} + 1 = \mathrm{aliont\ j} + \mathrm{Dt} \tag{8}$$

where, antlion is the position of the *j*th antlion during ith iteration

2.5 DBN Classification

At this stage, the DBN model receives the chosen features and performs lung cancer classification. Typically, DBN is constructed by stacking Restricted Boltzmann Machine (RBM) that captures higher-order correlation that is noted in the visible unit. DBN is pretrained in an unsupervised greedy layer-wise manner for learning a stack of RBM through the Contrastive Divergence (CD) approach. The output depiction of RBM is utilized as the input dataset to train the RBM in the stack. Afterward the pretraining, the DBN is finetuned by BP of error derivative and the initial weight and bias of all the layers are corrected. RBM is an approach to represent each training sample in a compact and meaningful manner, by capturing the regularities and inner structure. This is realized by presenting an additional set of neurons named hidden unit to the network that value is indirectly fixed from training dataset [17]. On the other hand, visible unit obtains the value directly from training dataset. Obviously, the network contains three hidden nodes and four visible nodes. In the forward pass, the output y_j is generated through multiplying all the inputs, x_i by respective weight w_{ij} and sum up each product. Then, the summation is included to bias, b, and lastly, the outcome is passed through an activation function f to generate the node output, y_j. For learning the connection weight of the system, we need to emphasize how they learn to recreate data in an unsupervised manner, which makes various forward and backward passes among the visible and hidden layers without including a deep network. In the recreation stage, the output activation of the forward pass, y_j, denotes the input to the backward pass. It is multiplied with the similar weight of the forward pass, $w_1 \ldots w_n$, as well as the summation of the product is added to visible-layer bias that produces the concluding output of operation that is the reconstruction of the original in- put, r_i. Definitely, the recreated value doesn't match the original one. In another word, the recreation is making guess regarding the possibility distribution of the novel input; that is, value of different points simultaneously. Fig. 3 showcases the framework of DBN technique.

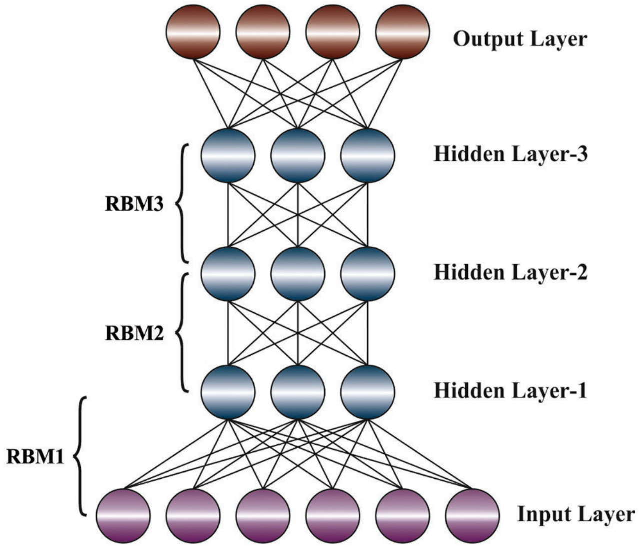

Figure 3: DBN structure

Consider the input dataset and the reconstruction is standard curve of distinct shapes. The aim is to reduce the error or diverging area in the two curves, named Kullback-Leibler $(KL)-$ Divergence, with RBM optimization approach CD [18]. CD learning accurately follows the gradient of the variance of two KL-Divergences. This process assists in iteratively altering the weight and then estimating the actual information that makes the two-possibility distribution converge.

Therefore, the possibility distribution of a group of novel input x, $p(x)$ and the recreated dispersal $q(x)$ whereas the incorporation of the difference.

Consider P and Q represent the distribution of a constant arbitrary parameter; the KL-Divergence equation is determined below:

$$D_{(L}(P\|Q) = \int_{-\infty}^{+\infty} p(x)\log\frac{p(x)}{q(x)}dx, \tag{9}$$

Now, p and q denote the density of P and Q.

For good understanding of the CD approach, mathematical calculation is included.

Now v and h denote visible and hidden layers of RBM, correspondingly. The energy of joint configuration, (v, h) of the visible and hidden layers are determined below:

$$E(v,\ h) = -\sum_{i \in visib1e} a_i v_i - \sum_{j \in hidden} b_j h_j - \sum_{i,j} v_i h_j w_{ij}, \tag{10}$$

Now v_i, h_j indicates the binary state of visible node i and hidden node j, a_i and b_j represent the bias value, and w_{ij} indicates the weight among v_i and h_j.

Next, the joint possibility through v and h is calculated by:

$$p(v,\ h) = \frac{1}{\sum_{v,h} e^{-E(v,h)}} e^{-E(v,h)}. \tag{11}$$

Assume $p(v)$ indicates the possibility that RBM takes visible vector v over summing by each probably hidden vector:

$$p(v) = \frac{1}{\sum_{v,h} e^{-E(v,h)}} \sum_h e^{-E(v,h)}. \tag{12}$$

The equation to update weight is estimated by taking derivative of $\log p(v)$ regarding weight w_{ij}:

$$\triangle w_{i,j} = \varepsilon(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{model}), \tag{13}$$

Here ε indicates the learning rate and $\langle\rangle$ represent the expectancy in the model or data distribution.

The conditional probability of $h_j = 1$ provided v and $v_i = 1$ shown h are given below:

$$p(h_j = 1|v) = \sigma\left(b_j + \sum_i v_i w_{ij}\right), \tag{14}$$

Now, $\sigma(x)$ indicates the sigmoid function.

$$p(v_i = 1|h) = \sigma\left(a_i + \sum_j h_j w_{ij}\right). \tag{15}$$

Having Eqs. (12), (14), and (15), it is easier to attain an un-biased instance of $\langle v_i h_j \rangle_{data}$ while there are no straightforward interconnections among visible and hidden nodes in RBM. In contrast, an unbiased instance of $\langle v_i h_j \}_{model}$ is very complex. Gibb's sampling is utilized by upgrading each hidden and visible node separately resulting in a slower convergency. To resolve this problem, the CD approach generates a recreation of a visible vector by setting visible node to one with possibility $p(v_i = 1|h)$ when the binary state has been calculated for hidden node. The weight upgrade is shown below

$$\triangle w_{i,j} = \epsilon(\langle v_i h_j \rangle_{data} - \langle v_i h_j \rangle_{recon}), \tag{16}$$

Now, recon represents the recreation stage.

Finally, the Adam optimizer is utilized for hyperparameter optimization of the DBN model. Simultaneously, the hyperparameter optimization of the DBN was performed by Adam optimizer. It is applied for estimating an adaptive learning value where the variable is employed for training the variable of the DNN [19]. It is effective and elegant technique for the first-order gradient with controlled storage for stochastic optimization. Here, the lately presented technique has been applied for resolving the ML problem with higher dimension variable space, and massive data evaluates the learning rate for various characteristics from calculation with initial and another order moments. In addition, the Adam optimizer is commonly employed according to the momentum and gradient descent (GD) methods, and a variation of an interval. Consequently, initial momentum has been gained b7:

$$m_i = \beta_1 m_{i-1} + (1 - \beta_1)\frac{\partial C}{\partial w}. \quad (17)$$

The next momentum is expressed by,

$$v_i = \beta_2 v_{i-1} + (1 - \beta_2)\left(\frac{\partial C}{\partial w}\right)^2. \quad (18)$$

$$w_{i+1} = w_i - \eta \frac{\hat{m}_i}{\sqrt{\hat{v}_i + \epsilon}}, \quad (19)$$

Here $\hat{m}_i = m_i/(1 - \beta_1)$ and $\hat{v}_i = v_i/(1 - \beta_2)$.

3 Performance Evaluation

This section inspects the performance validation of the ALO-ODBN model using benchmark lung cancer dataset [20]. It comprises 32 samples with 56 features and 3 class labels.

Fig. 4 demonstrates a collection of confusion matrices produced by the ALO-ODBN model on distinct runs. On run-1, the ALO-ODBN model has identified 9, 13, and 9 class labels respectively. Moreover, on run-2, the ALO-ODBN methodology has identified 9, 12, and 10 class labels correspondingly. Furthermore, on run-3, the ALO-ODBN approach has identified 9, 12, and 9 class labels respectively. Along with that, on run-4, the ALO-ODBN technique has identified 9, 11, and 9 class labels respectively. In line with, on run-5, the ALO-ODBN algorithm has identified 9, 13, and 8 class labels correspondingly.

Tab. 1 and Fig. 5 demonstrates brief classification results of the ALO-ODBN model using distinct runs. The results indicated that the ALO-ODBN model has accomplished maximum classification performance. For instance, with run-1, the ALO-ODBN model has provided an average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 97.92%, 97.62%, 96.67%, and 97.01% respectively. In addition, with run-2, the ALO-ODBN technique has offered an average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 97.92%, 96.67%, 97.44%, and 96.91% respectively. Meanwhile, with run-4, the ALO-ODBN method has provided an average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 93.75%, 91.67%, 91.54%, and 90.71% correspondingly. Eventually, with run-5, the ALO-ODBN system has provided an average $accu_y$, $prec_n$, $reca_l$, and F_{score} of 95.83%, 93.94%, 93.33%, and 92.96% correspondingly.

The training accuracy (TA) and validation accuracy (VA) attained by the ALO-ODBN approach on lung cancer classification is demonstrated in Fig. 6. The experimental outcomes implied that the ALO-ODBN model has gained maximum values of TA and VA. In specific, the VA is seemed to be higher than TA.

The training loss (TL) and validation loss (VL) achieved by the ALO-ODBN model on lung cancer classification are established in Fig. 7. The experimental outcomes inferred that the ALO-ODBN technique has accomplished least values of TL and VL. In specific, the VL is seemed to be lower than TL.

A brief precision-recall examination of the ALO-ODBN model on test dataset is represented in Fig. 8. By observing the figure, it can be stated that the ALO-ODBN approach has accomplished maximal precision-recall performance under test dataset.

A brief ROC investigation of the ALO-ODBN method on test dataset is depicted in Fig. 9. The results exposed that theALO-ODBN model has exhibited its ability in categorizing three different classes such as class 1, class 2, and class 3 on the test datasets.

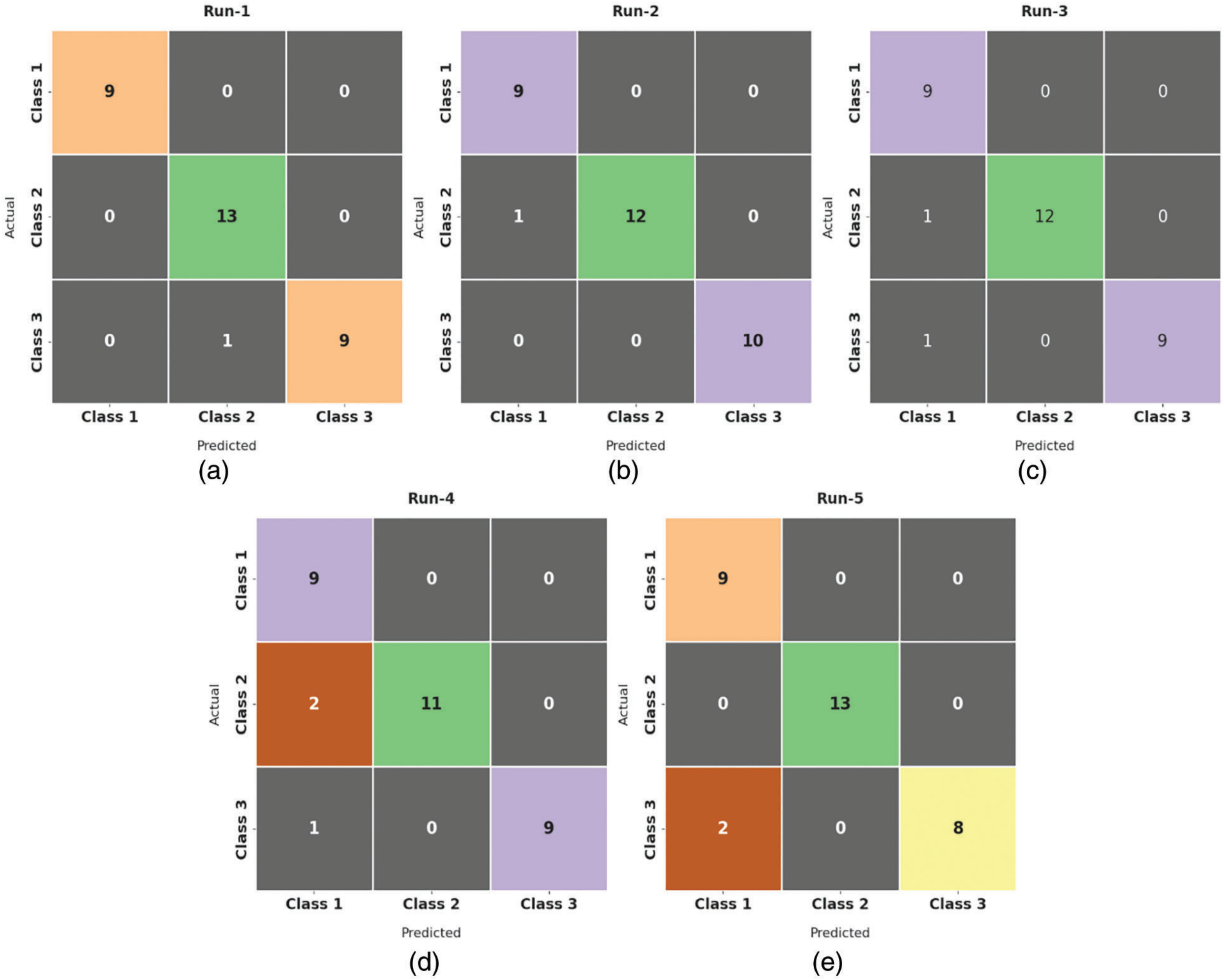

Figure 4: Confusion matrix of ALO-ODBN technique with distinct runs

Table 1: Result analysis of ALO-ODBN technique with distinct measures and runs

Class labels	Accuracy	Precision	Recall	F-score
Run-1				
Class 1	100.00	100.00	100.00	100.00
Class 2	96.88	92.86	100.00	96.30
Class 3	96.88	100.00	90.00	94.74
Average	**97.92**	**97.62**	**96.67**	**97.01**
Run-2				
Class 1	96.88	90.00	100.00	94.74
Class 2	96.88	100.00	92.31	96.00
Class 3	100.00	100.00	100.00	100.00
Average	**97.92**	**96.67**	**97.44**	**96.91**

(Continued)

Table 1 (continued)

Class labels	Accuracy	Precision	Recall	F-score
Run-3				
Class 1	93.75	81.82	100.00	90.00
Class 2	96.88	100.00	92.31	96.00
Class 3	96.88	100.00	90.00	94.74
Average	**95.83**	**93.94**	**94.10**	**93.58**
Run-4				
Class 1	90.62	75.00	100.00	85.71
Class 2	93.75	100.00	84.62	91.67
Class 3	96.88	100.00	90.00	94.74
Average	**93.75**	**91.67**	**91.54**	**90.71**
Run-5				
Class 1	93.75	81.82	100.00	90.00
Class 2	100.00	100.00	100.00	100.00
Class 3	93.75	100.00	80.00	88.89
Average	**95.83**	**93.94**	**93.33**	**92.96**

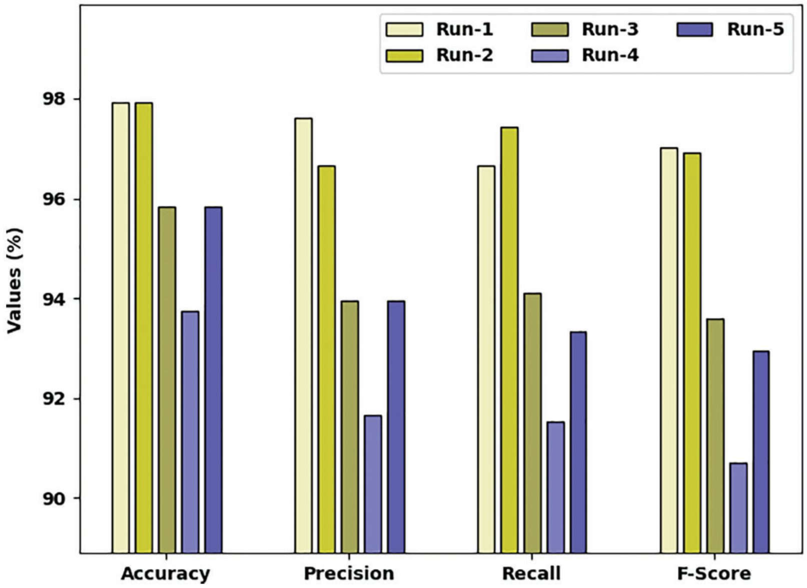

Figure 5: Result analysis of ALO-ODBN technique with distinct runs

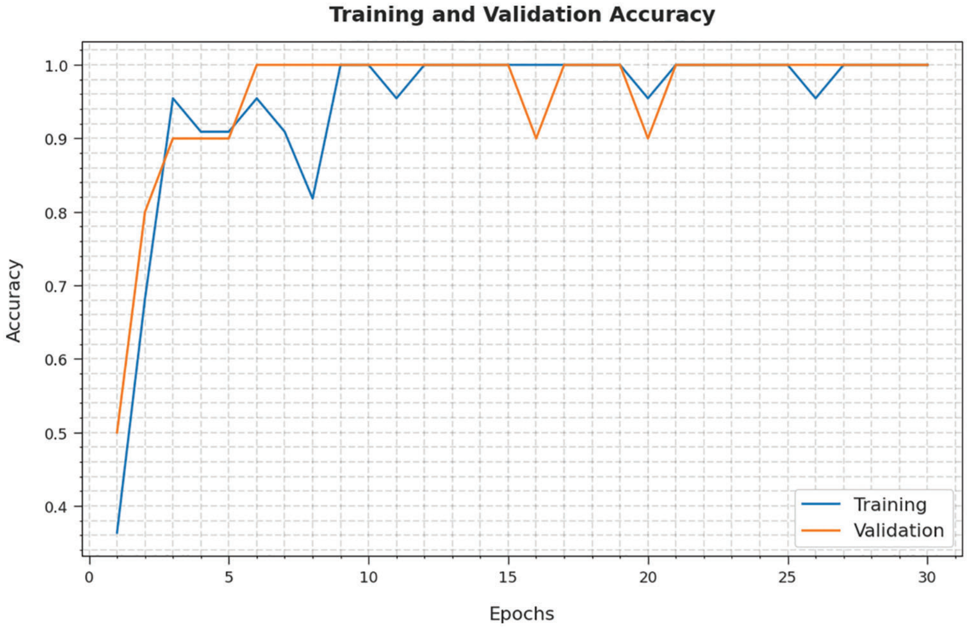

Figure 6: TA and VA analysis of ALO-ODBN technique

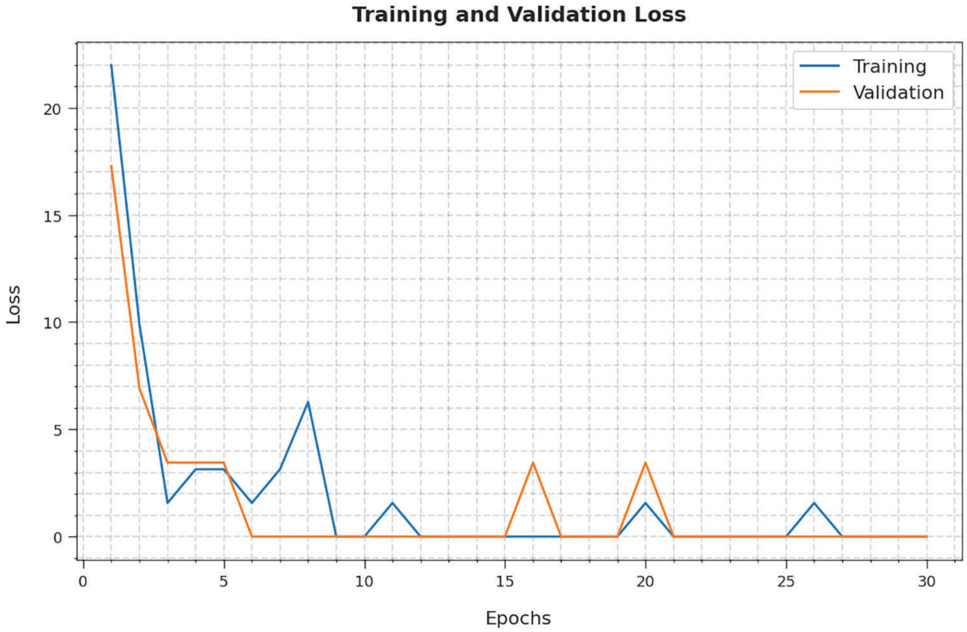

Figure 7: TL and VL analysis of ALO-ODBN technique

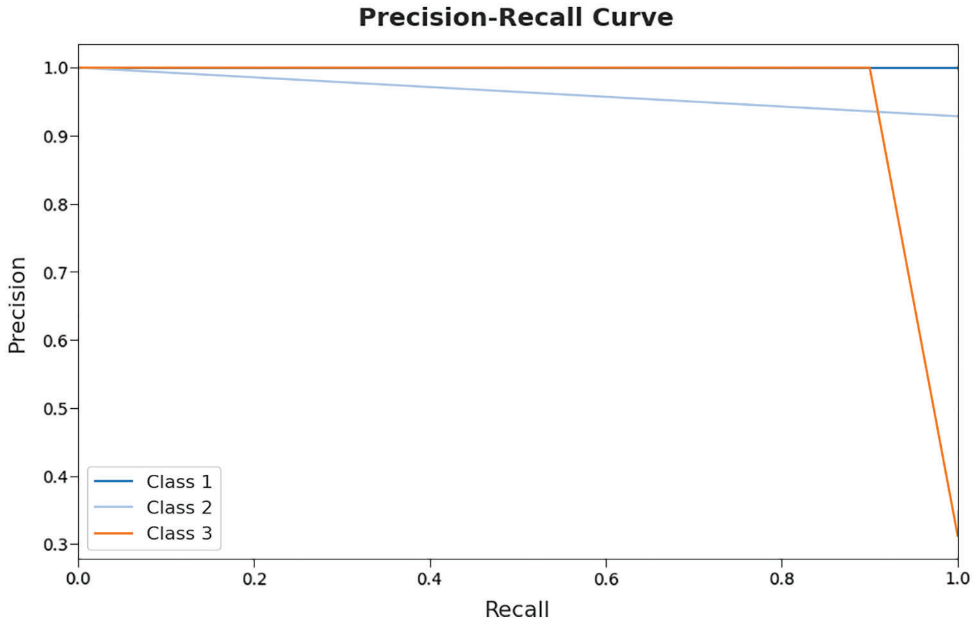

Figure 8: Precision-recall curve analysis of ALO-ODBN technique

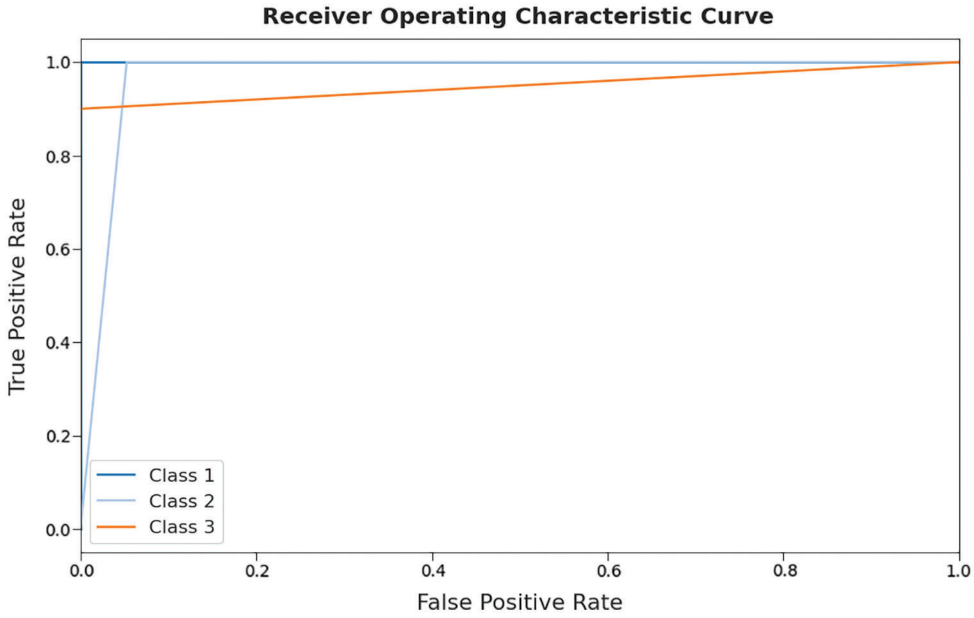

Figure 9: ROC curve analysis of ALO-ODBN technique

For ensuring the betterment of the ALO-ODBN model, a comparison study is made with existing models in Tab. 2 [21]. A detailed $accu_y$ and $prec_n$ investigation of the ALO-ODBN model with existing models is performed in Fig. 9. The result indicated that the SVM model has shown least performance with minimal values of $accu_y$ and $prec_n$. In line with, the Radial Basis Function (RBF) Neural Network has accomplished slightly improved outcome with values of $accu_y$ and $prec_n$. In addition, the DNN and DT (C4.5) model has shown reasonable values of and $prec_n$. However, the ALO-DBN model has exhibited superior $accu_y$ and $prec_n$ of 97.92% and 97.62% respectively.

Table 2: Comparative analysis of ALO-ODBN technique with existing approaches

Methods	Accuracy	Precision	Recall	F1-score
ALO-ODBN	97.92	97.62	96.67	97.01
DNN Model	97.16	90.87	89.23	81.30
SVM algorithm	70.56	88.34	87.03	85.60
RBF neural network	95.83	90.15	88.48	92.15
Decision tree (C4.5)	96.96	93.85	93.12	94.64

A brief $reca_l$ and $F1_{score}$ investigation of the ALO-ODBN method with existing techniques is implemented in Fig. 10. The outcome referred that the SVM approach has shown least performance with minimal values of $reca_l$ and $F1_{score}$. Likewise, the RBF Neural Network has accomplished somewhat enhanced outcome with values of $reca_l$ and $F1_{score}$. Besides, the DNN and DT (C4.5) approach has revealed reasonable values of $reca_l$ and $F1_{score}$ shown in Fig. 11. At last, the ALO-DBN approach has exhibited superior $reca_l$ and $F1_{score}$ of 96.67% and 97.01% correspondingly.

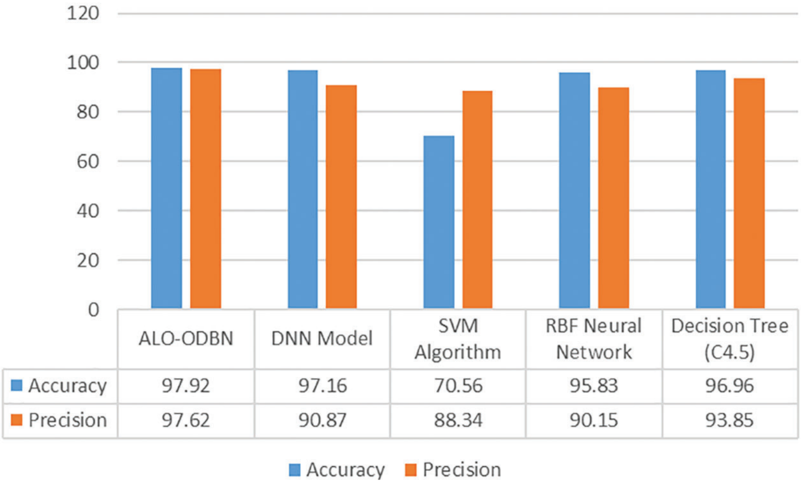

Figure 10: Acc_y and $Prec_n$ analysis of ALO-ODBN technique with existing methodologies

From the above mentioned tables and figures, it is apparent that the ALO-ODBN model has outperformed the other methods interms of different measures. The overall survival rate detection of patient is improved when compared with other existing methods.

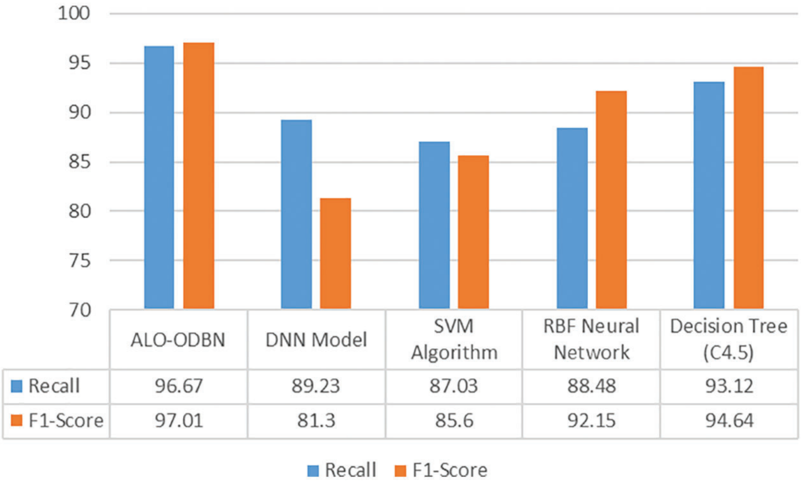

Figure 11: $Reca_l$ and $F1_{score}$ analysis of ALO-ODBN technique with existing methodologies

4 Conclusion

In this study, a new ALO-ODBN model has been developed for Lung Cancer Detection, Classification and survival rate prediction. The proposed ALO-ODBN model aims to identify and classify the presence of lung cancer. Initially, the ALO-ODBN model undergoes min-max data normalization approach to preprocess the input data. Besides, the ALO algorithm gets executed to choose an optimal subset of features. In addition, the DBN model receives the chosen features and performs lung cancer classification. Finally, the Adam optimizer is utilized for hyperparameter optimization of the DBN model. In order to report the enhanced performance of the ALO-ODBN model, a wide-ranging experimental analysis is performed and the results reported the supremacy of the ALO-ODBN model.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study

References

[1] K. Ahmed, A. A. Kawsar, E. Kawsar, A. A. Emran, T. Jesmin *et al.,* “Early detection of lung cancer risk using data mining,” *Asian Pacific Journal of Cancer Prevention*, vol. 14, no. 1, pp. 595–598, 2013.

[2] P. Ramachandran, N. Girija and T. Bhuvaneswari, “Early detection and prevention of cancer using data mining techniques,” *International Journal of Computer Applications*, vol. 97, no. 13, pp. 48–53, 2014.

[3] V. Krishnaiah, G. Narsimha and N. S. Chandra, “Diagnosis of lung cancer prediction system using data mining classification techniques,” *International Journal of Computer Science and Information Technologies*, vol. 4, no. 1, pp. 39–45, 2013.

[4] Z. S. Zubi and R. A. Saad, “Using some data mining techniques for early diagnosis of lung cancer,” in *Proc. of the 10th WSEAS International Conf. on Artificial Intelligence Knowledge Engineering and Data Bases*, Cambridge, UK, pp. 32–37, 2011.

[5] R. G. Ramani and S. G. Jacob, “Improved classification of lung cancer tumors based on structural and physicochemical properties of proteins using data mining models,” *PloS One*, vol. 8, no. 3, pp. e58772, 2013.

[6] S. Shah and A. Kusiak, “Cancer gene search with data-mining and genetic algorithms,” *Computers in Biology and Medicine*, vol. 37, no. 2, pp. 251–261, 2007.

[7] K. Ahmed, T. Jesmin and M. Z. Rahman, “Early prevention and detection of skin cancer risk using data mining,” *International Journal of Computer Applications*, vol. 62, no. 4, pp. 1–6, 2013.

[8] D. Chauhan and V. Jaiswal, “An efficient data mining classification approach for detecting lung cancer disease,” in *Proc. Int. Conf. on Communication and Electronics Systems (ICCES), IEEE*, Coimbatore, India, pp. 1–8, 2016.

[9] F. Heydari and M. K. Rafsanjani, “A review on lung cancer diagnosis using data mining algorithms,” *Current Medical Imaging*, vol. 17, no. 1, pp. 16–26, 2021.

[10] K. Juma, M. He and Y. Zhao, “Lung cancer detection and analysis using data mining techniques, principal component analysis and artificial neural network,” *American Academic Scientific Research Journal for Engineering, Technology, and Sciences*, vol. 26, no. 3, pp. 254–265, 2016.

[11] N. Maleki, Y. Zeinali and S. T. A. Niaki, “A K-NN method for lung cancer prognosis with the use of a genetic algorithm for feature selection,” *Expert Systems with Applications*, vol. 164, pp. 113981, 2021.

[12] A. Agrawal, S. Misra, R. Narayanan, L. Polepeddi and A. Choudhary, “A lung cancer outcome calculator using ensemble data mining on SEER data,” in *Proc. of the Tenth Int. Workshop on Data Mining in Bioinformatics*, San Diego, CA, USA, pp. 1–9, 2011.

[13] B. Muthazhagan, T. Ravi and D. Rajinigirinath, “An enhanced computer-assisted lung cancer detection method using content based image retrieval and data mining techniques,” *Journal of Ambient Intelligence and Humanized Computing* vol. early access, pp. 1–9, 2020.

[14] S. B. Lim, S. J. Tan, W. T. Lim and C. T. Lim, “A merged lung cancer transcriptome dataset for clinical predictive modeling,” *Scientific Data*, vol. 5, no. 1, pp. 1–8, 2018.

[15] X. Qi, Z. Guo, Q. Chen, W. Lan, Z. Chen *et al.,* “A data mining-based analysis of core herbs on different patterns (zheng) of non-small cell lung cancer,” *Evidence-Based Complementary and Alternative Medicine*, vol. 2021, pp. 1–13, 2021.

[16] Y. Wu, B. Yan and X. Qu, “Improved chicken swarm optimization method for reentry trajectory optimization,” *Mathematical Problems in Engineering*, vol. 2018, pp. 1–13, 2018.

[17] A. Fischer and C. Igel, “An introduction to restricted boltzmann machines,” in *Iberoamerican Congress on Pattern Recognitio*, Berlin, Heidelberg, Germany, Springer, pp. 14–36, 2012.

[18] W. Almanaseer, M. Alshraideh and O. Alkadi, “A deep belief network classification approach for automatic diacritization of arabic text,” *Applied Sciences*, vol. 11, no. 11, pp. 5228, 2021.

[19] Z. Zhang, “Improved adam optimizer for deep neural networks,” in *Proc. IEEE/ACM 26th Int. Symp. on Quality of Service (IWQoS), IEEE*, Banff, AB, Canada, pp. 1–2, 2018.

[20] Z. Q. Hong and J. Y. Yang, “Optimal discriminant plane for a small number of samples and design method of classifier on the plane,” *Pattern Recognition*, vol. 24, no. 4, pp. 317–324, 1991.

[21] N. R. Murty and M. P. Babu, “A critical study of classification algorithms for lung cancer disease detection and diagnosis,” *International Journal of Computational Intelligence Research*, vol. 13, no. 5, pp. 1041–1048, 2017.