ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A COMPUTER VISION-BASED THYROID DISEASE PREDICTION USING AN EBOLA OPTIMIZATION-ENHANCED DEEP LEARNING MODEL

DR. CH. VIJAYA KUMAR [1], SURESHBABU AMARTHALURI [2], SIRISHA.J [3], RVVN BHEEMA RAO [4], CH.SABITHA [5], V. PURUSHOTHAMA RAJU [6], N.PRASANNA LAKSHMI [7], B SREENIVAS REDDY [8]

[1]Department of Computer Science and Engineering, ACE Engineering College,Hyderabad,Telangana, India
[2]Department of CSE, Vignan's Foundation for Science, Technology & Research, Andhra Pradesh, India
[3]Dept of IT, Prasad V.Potluri Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India
[4]Department of Information Technology, Aditya University, Surampalem, Andhra Pradesh, India
[5]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh, India
[6]Department of CSE, Shri Vishnu Engineering College for Women, Bhimavaram, Andhra Pradesh, India
[7]Department of CSE(AI&ML),RVR&JC COLLEGE OF ENGINEERING, Guntur, Andhra Pradesh, India
[8]Department of EEE, Sri Venkateswara College of Engineering, Tirupati, Andhra Pradesh, India

Email id : [1] vijay.chandarapu@gmail.com, [2] sa_csedip@vignan.ac,[3]siri.jagannadham@gmail.com, [4] rvvnbrao@gmail.com, [5] sabithakiran.ch@kluniversity.in ,[6] praju@svecw.edu.in, [7]lakshmiprasannagali62@gmail.com , [8] sreenivasreddy.b@svcolleges.edu.in

ABSTRACT

The thyroid is an important gland that regulates metabolism through its hormones. Two major thyroid diseases, hyperthyroidism and hypothyroidism, are caused by abnormal thyroid function. Computer vision is becoming increasingly important in the medical diagnosis of various diseases, providing automated analytical models for accurate and rapid diagnosis of lesions. This paper proposed to a computer vision-based thyroid disease prediction framework using an Ebola optimization-based deep learning structure (EODLM). The three basic phases of the suggested structure are pre-processing, feature extraction, and classification. Use open-source databases and a normalisation strategy to substitute missing values during the pre-processing stage. A three-step feature selection approach, involving bidirectional feature removal, backward feature elimination, and forward feature elimination is used during the feature extraction stage. During the classification phase, present the EODLM, an innovative deep learning model that integrates the Elman Recurrent Neural Network (ERNN) with the Ebola Optimisation Algorithm (EOA). The ERNN's weight parameters are optimised using the EOA, the suggested strategy was put into practise in MATLAB, and the following statistical metrics were used to assess its effectiveness: F-measure, sensitivity, kappa, accuracy, recall, specificity, and precision. the suggested approach was contrasted with popular machine learning methods, such as Support Vector Machine (SVM), Efficient Feature Extraction Recurrent Neural Networks (EFERNN), Deep Belief Neural Networks (DBN), and Artificial Neural Networks (ANN). The suggested EODLM performed better than any other method, yielding the following outcomes: F-measure (0.95), specificity (0.99), recall (0.93), accuracy (0.99), kappa (0.94), sensitivity (0.96), and precision (0.94). A promising computer vision-based framework for the prediction of thyroid disease is the suggested EODLM. It works better than other well-known machine learning algorithms and attains excellent accuracy.

Keywords: *Thyroid Prediction, Deep Learning Model, Elman Recurrent Neural Network, Ebola Optimization Algorithm, Bidirectional Feature Elimination, Forward Feature Selection, And Backward Feature Elimination.*

1. INTRODUCTION

The thyroid is a massive gland that looks like a butterfly. It is positioned in the lower neck and aids in the regulation of the body's metabolism [1, 27]. The two active thyroid hormones generated by this gland are levothyroxine and triiodothyronine they are shortened as T4 and T3 [2, 28]. These hormones are essential for managing energy levels generally, protein synthesis, and regulating body temperature.

The thyroid gland can be impacted by a wide range of diseases [3, 29] some of which are more common than others, such as hypothyroidism. Hypothyroidism is caused by insufficient thyroid hormone secretion, whereas hyperthyroidism is caused by excessive production of thyroid hormone [4, 30]. The former situation deals with hypothyroidism, a condition in which there is insufficient or inadequate exhibition of thyroid hormones [5, 31].

In comparison to hyperthyroidism, which is characterised by the thyroid gland producing an excessive amount of thyroid hormone [6, 32], this condition can cause a low pulse rate, swelling in front of the neck, and weight gain. A person may also experience reduced body weight while experiencing elevated blood pressure and a rapid heartbeat. Blood tests, which can measure T4, T3, and TSH levels, are a frequent way to detect thyroid issues. According to experts, detecting, diagnosing, and treating diseases early is essential to halting their progression and even death [7, 33]. For a number of problems, early detection and differential diagnosis increase the likelihood of successful therapy. Clinical diagnosis is often seen as a difficult task, even after multiple tests. Machine learning algorithms are among the best solutions for a wide range of difficult circumstances.

Classification is a data extraction process used to forecast and diagnose a range of ailments, including thyroid disease [8]. We examined and classified thyroid disease here because to machine learning algorithms' significant contributions to the classification of thyroid disease, their high efficacy and efficiency, and their facilitation of classification. Artificial intelligence and computer learning have been used in medicine since the profession's inception. even though there has long been a push to investigate the requirement aimed at solutions in healthcare that are driven by machine learning [9, 34].As a consequence, predictors believe machine learning will soon develop in the healthcare industry. The healthcare sector can use a variety of data mining methods, like classification, regression, association, and clustering for disease identification. Thyroid problems affect a huge number of people each year [10, 35]. As a result, healthcare facilities are finding it harder and harder to make an accurate diagnosis. The following literature analysis will highlight many machine learning algorithms utilized in various research to diagnose thyroid issues. While TieNet [23] and CheXNet [22] laid the foundation for multimodal diagnostics, our approach uniquely integrates EOA-optimized ERNN with CNN-BERT fusion, addressing interpretability and synergy limitations noted in prior works.

Contribution of the research

- In this paper, develop EODLM for thyroid disease prediction. In this proposed technique classification, pre-processing,and feature extraction method is introduced to achieve efficient thyroid detection from the databases.
- To replace missing values using a normalization approach, a pre-processing method is then constructed.
- After that, the feature selection is carried out in three stages: Forward feature removal, backward feature removal, and bidirectional feature removal. Three phases serve as the foundation for the important features.
- Finally, the EODLM is used to identify thyroid from the chosen attributes. The suggested classifier combines the ERNN and EOA algorithms. The EOA is used to choose the weight parameters in the ERNN.
- The proposed method has been evaluated in MATLAB, evaluating statistical parameters such as sensitivity, kappa, accuracy, recall, specificity, precision, and F-measure. Comparing the suggested method to well-known ones like DBN, SVM, ANN, and EFERNN allows for evaluation.

The rest of the study has been organised as follows, with Section 2 providing information on related thyroid detection work. Section 3 describes the proposed methodology in detail. Section 4 provides an analysis of the expected methodology's results. The conclusion is found in Section 5 of the paper.

2. RELATED WORK

Researchers have developed a variety of methods for detecting thyroid conditions. This section only briefly reviews a few research articles.

For a thyroid illness classifier, Vidhushavarshini Suresh Kumar et al. [11] proposed a type-2 fuzzy support vector machine with a feature selection architecture based on a hybrid optimisation technique. The top-n qualities are chosen in this study utilising a hybrid optimisation technique that combines the firefly algorithm (FA) and the butterfly optimisation algorithm (BOA). HFBO-

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

RT2FSVM, which stands for hybrid firefly butterfly optimization-rough type-2 fuzzy support vector machine, was proposed. was assessed using some critical metrics, including specificity, accuracy, and sensitivity. This section compares this method to well-known benchmark techniques like the mixed-kernel support vector machine (MKSVM) techniques and enhanced (GWO Linear SVM).

Rehman *et al.,* [12] have created the K-Nearest Neighbor (KNN) algorithm and its several distance-based standards to recognize thyroid illness. The proposed study is divided into three phases: Three feature selection procedures are used by KNN: chi-square feature selection, L1-based feature selection, and no feature selection. Thyroid datasets from a reputable hospital in Pakistan and the KEEL dataset repository were both utilised in this study. The hot dataset differed from previous datasets due to the addition of three new features: blood pressure (BP), body mass index (BMI), and pulse rate. A range of distance functions were used to assess the KNN model's performance on these two datasets.

Mehdi Hosseinzadeh *et al.,* [13] have introduced an artificial neural network (ANN) to boost thyroid illness detection precision in IoMT systems using semantic reports and test results. In order to improve generalisation and prevent over-fitting of ANN during the learning phase, a pair of multiple multilayer perceptron (MMLP) neural networks with the ability to back-propagate error were built in this study. Additionally, the back-propagation error technique's difficulties with slow convergence and local minima were solved by employing an adjustable learning rate strategy. The overall classification of thyroid illness was greatly enhanced by the suggested MMLP.

In a recent comprehensive guide on artificial intelligence and neural networks, Saka and Sowjanya [37] provide foundational insights into deep learning architectures, optimization strategies, and their applications in biomedical domains. This work supports the methodological choices in our proposed framework, particularly the integration of CNNs and ERNNs for diagnostic modeling and the use of metaheuristic optimization for parameter tuning.

Table 1: Comparison of related work

S.No	Author	Title	Technique	Advantages	Limitations
1	Alam, Siddique & Adeli (2020) [22]	A dynamic ensemble learning algorithm for neural networks	Innovative, dynamic neural network ensemble learning	It offers extremely few user-designable parameters, automatic ensemble design, and maintenance of accuracy and diversity of the composing neural networks.	Train only with complete sequences.
2	Zhao et al. (2022) [23]	Semantic consistency generative adversarial network for cross-modality domain adaptation in ultrasound thyroid nodule classification	semantic consistency generative adversarial network	Using ultrasound data, it is used to find cancerous thyroid nodules.	high variability
3	Ai et al. (2022) [24]	ResCaps: an improved capsule network and its application in ultrasonic image classification of thyroid papillary	recent type of neural networks, i.e., capsule networks	high accuracy thyroid ultrasound imaging to look for any signs of thyroid cancer	extremely sensitive to the type of data input

		carcinoma			
4	Yadav & Pal (2022) [25]	Thyroid prediction using ensemble data mining techniques	Boosting, Bagging, Stacking, and Voting ensembles	It is used to identify hypothyroidism with the highest degree of accuracy.	much training data needed
5	. Juneja (2022) [26]	Expanded and filtered features-based ELM model for thyroid disease classification	The fuzzy adaptive feature filtration and expansion-based model	It is used to produce a brand-new set of thyroid-related features.	Including excessively many features
6	, Saka and Sowjanya(2025) [37]	The integration of CNNs and ERNNs for diagnostic modeling and the use of metaheuristic optimization for parameter tuning.	CNNs and ERNNs for diagnostic modeling	The use of metaheuristic optimization for parameter tuning.	Optimized Strategies

Tehseen Akhtar et al. [14] developed a powerful homogeneous ensemble of ensembles as well as many feature-selection techniques for improved thyroid illness diagnosis. The database was built using real-time thyroid data from the District Headquarters (DHQ) teaching hospital in Dera Ghazi (DG) Khan, Pakistan. Following that, three other attribute selection techniques were used: Recursive Feature Elimination (RFE), Select K-Best (SKB), and Select from Model (SFM). Among the promising feature estimators were logistic regression (LR), gradient boosting (GB), random forest (RF), and decision tree (DT). The homogeneous ensemble activated the boosting and bagging-based classifiers, which were subsequently classed by the voting ensemble using both soft and hard voting. The following performance evaluation measures were taken into account: sensitivity, accuracy, hamming loss, and mean square error.

Shankar et al. [15] proposed a multi-kernel SVM to categorise thyroid data based on best feature selection and a kernel-based classifier approach. The originality and feature selection objective of the recommended model are used to improve grey wolf optimisation and classification process performance. This optimum feature selection is due to the unique dataset's irrelevant properties and the model's increased computational performance. Comparing the suggested thyroid categorization to the current model, the findings are 97.49, 99.05, and 94.5% accuracy, sensitivity, and specificity.

Research Questions and Hypothesis

Research Questions:

- Can multimodal fusion of image and text data improve diagnostic accuracy over unimodal systems?
- Does EOA optimization enhance ERNN performance in clinical prediction tasks?

Hypothesis:

- The integration of CNN, BERT, and EOA-optimized ERNN will yield superior diagnostic performance and interpretability compared to traditional models.

Suggested additions:

- Open research issues include real-time deployment of multimodal systems in clinical settings, scalability across diverse medical domains, and the need for standardized interpretability benchmarks. Future work should explore federated learning for privacy-preserving multimodal training and domain adaptation for cross-hospital generalization.

3. PROPOSED EBOLA OPTIMIZATION-BASED DEEP LEARNING MODEL

Endocrinology includes thyroid illness, one of the most poorly understood and untreated disorders. According to the World Health Organisation, thyroid gland abnormalities are second only to diabetes in terms of occurrence among endocrine

Journal of Theoretical and Applied Information Technology
30th April 2026. Vol.104. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

disorders. Between 2% and 1% of the population, respectively, have hyperthyroidism and hypothyroidism. Women are nearly ten times more likely to be affected than men.Thyroid gland dysfunction, pituitary gland failure as a cofactor, or hypothalamus dysfunction as a tertiary cause may all contribute to hyper- and hypothyroidism. ,In recent years, early thyroid prediction has become critical to study. As a result, an efficient technique for thyroid prediction using datasets is established. Figure 1 depicts the entire architecture of the proposed approach. Develop EODLM for thyroid illness prediction in this paper. Pre-processing, feature extraction, and classification methods are provided in this proposed strategy to achieve efficient thyroid detection from databases. The initial database is obtained from an open-source system. A pre-processing method is then constructed utilising a normalisation technique to substitute missing values. Following that, three stages of the feature selection procedure are carried out: forward feature selection, backward feature elimination, and bidirectional feature elimination. The essential features are retrieved in three stages. Finally, the EODLM is used to identify thyroid from the chosen attributes. The proposed classifier is an accumulation of ERNN and EOA. The EOA is used to choose the weight parameters in the ERNN. In the parts that follow, the proposed method is explained in more detail

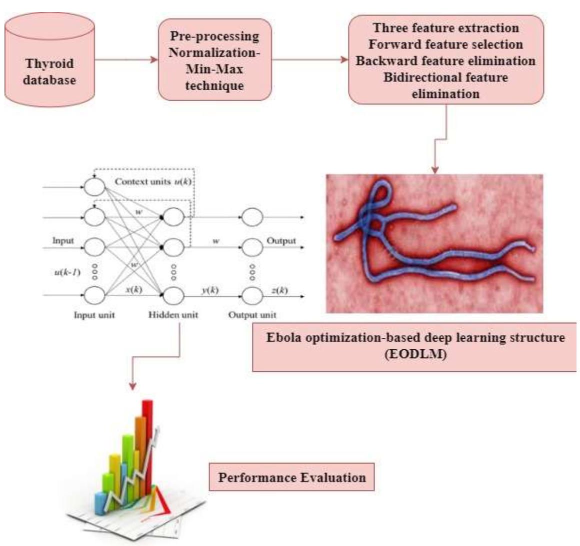


Figure 1: Proposed flow Block Diagram of the Thyroid Disease Prediction

Study Design and Research Protocol

This study follows a structured protocol comprising dataset acquisition, preprocessing, feature selection, model training, and evaluation. Public datasets were normalized and split into training/testing subsets. Feature selection was performed using forward, backward, and bidirectional techniques. The EODLM model was trained using ERNN optimized by EOA, with performance validated through standard metrics.

3.1. Pre-processing model

The methodology's first stage is data processing. which contains cleaning of useless entries or columns and deletion of entries. Cleaning unnecessary data and processing missing values are enhancing the accuracy [16] of the prediction model. Additionally, processing missing parameters can be very efficient because skipping the parameters would negatively impact the outcomes as there is a risk of losing essential data. Based on the procedure, feature scaling related tothe min-max technique is executed to achieve minimum and maximum entry parameters. The initial phase of the implementation with feature selection techniques was used to obtain excellent accuracy and performance of the classifier.

$$X = \left(\frac{X - MIN\ parameter\ of\ X}{MAX\ parameter\ of\ X - MIN\ parameter\ of\ X}\right) * (d - c) + c \qquad (1)$$

input database is characterized with$X$. A predefined border can be defined as d, c. The min-max normalization method is used to normalize the data. The feature selection model receives the normalizeddata. The feature selection model is given in detail in the section that follows.

3.2. Feature selection model

The database contains a variety of attributes, many of which are not necessary for training models to fit best and improve the effectiveness of machine learning designs. In deep learning, feature selection is particularly important for creating the best model and getting the best model performance [17]. To learn the model more quickly, more readily, and with fewer overfitting problems, unwanted and redundant characteristics should be required to be removed from the gathered datasets. In this

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

research, three kinds of feature selection are utilized for achieving efficient thyroid detection such as bidirectional, backward, and forward feature elimination. This method aids in removing the crucial features from the gathered database so that deep learning approaches can be learned.

3.2.1. Backward Feature Elimination

All of the features are included when using this feature selection approach. The component that has a high P-parameter is picked to be dropped from the model and then incorporated into the design. The P-parameter for the deleted feature should be greater than the P-parameter for the specific level. This procedure is continued until all features with high P-parameters have been entirely deleted from the design while ensuring that every feature that has been eliminated has a P-parameter higher than the desired level. The final collection of characteristics that were present could wind up being the most useful and applicable features that were used for effective categorization and detection.

Phase 1: Initiate with complete features to fit the design

$$y_0 = x \quad (2)$$

Phase 2: Find the feature with a high P-parameter from the feature set. The unique level parameter(S)contrasts with the high P-parameter characteristic. The assumption $x > s$ must be compensated to manage the feature elimination.

$$x^- = \arg\max_{x \epsilon y_k} j(y_k - x) \quad (3)$$

Phase 3: The phase 2 management of the next iteration$(K + 1)$ feature elimination is handled by removing the high P-parameter from the list and moving it to that phase. feature elimination. In this case, backward feature elimination is used to define the final feature list when the$K$parameter is set to zero.

$$y_{k+1} = y_k - x^-; k = k + 1 \quad (4)$$

This method is a widely utilized feature selection technique. This technique is achieved with a specific stage of 0.05 with a 95% efficiency.

3.2.2. Forward Feature Selection

This strategy starts with a null design and makes an effort to influence the model using every feature parameter. The feature with a low P-parameter can be picked for the forthcoming process. Then, it begins to include two feature mixtures in the design. The initial low P-parameter feature set must be the candidate of one feature when fitting the designs with two feature combinations. Two features with a low P-parameter are developed to fit the design with three feature mixtures. This process is recurrent as long as every feature in the feature set has a low P-parameter that is below the stated stage.

Phase 1: Specify the level parameter$(S)$and start with the null set

$$y_0 = \{\emptyset\} \quad (5)$$

Phase 2: Select the initial feature in some circumstances. Pick a feature at random from the feature list, for instance. the selection of low P-parameter characteristics from all available features that were used for the choice.

$$x^+ = \arg\max_{x \epsilon y_k} j(y_k + x) \quad (6)$$

Phase 3: The count of all existing complete low P-parameter features is increased for the selected low P-parameter feature. There is a 1 increase in the iteration$K$parameter.This element continued and moved back to phase 3 and the process continued till the complete features P-parameter was lower than the specific stage. The iteration procedure stops when $y_k < s$ and the k parameter is the complete count of features.

$$y_k = y_k + x^+; k = k + 1 \quad (7)$$

This feature selection is utilized for wide adaptation. This is applied in the original collected database. This feature selection is developed with a specified parameter of 0.04 with 96% assurance.

3.2.3. Bi-directional Feature Elimination

This method combines the elimination of backward features with the selection of forward features. A forward feature selection equivalent is this technique. To compare it to previously chosen characteristics, the newly selected feature is put through the backward elimination process in this case. Some selected features whose P-parameter values are above the specified level's "out" parameter are eliminated. Two stated level parameters must be computed using the "out" and "in" of parameter ranges in this method. To manage feature selection, the feature P-parameter must be lower than the provided stage inner parameter and higher than the specified level outer parameter. Otherwise, the feature will be removed from the feature list.

Phase 1: Start with a blank. The next step is to select a feature that is connected to the given condition. The forward feature selection is used in this case to contain the features in the list.

$$y_f = \{\emptyset\}; y_b = x \quad (8)$$

Phase 2: The next optimal feature is chosen utilizing the P-parameter equivalence. The essential features are selected using a traditional forward feature selection technique.

$$x^+ = \arg\max_{\substack{x\epsilon y_k \\ x\epsilon y_{bk}}} j(y_k + x) \quad (9)$$

$$y_{fk+1} = y_{fk} + x^+ \quad (10)$$

Phase 3: The next optimal feature is chosen by utilizing the P-parameter equivalence. The next feature is picked using a standard forward feature selection process, and any unimportant features are then removed using a backward feature removal approach. After that, it moves to phase to continue the process and still, the k parameter achieves the complete count of the feature set.

$$x^- = \arg\max_{\substack{x\epsilon y_{bk} \\ x\epsilon \ fk+1}} j(y_{bk} - x) \quad (11)$$

$$y_{bk+1} = y_{bk} - x^-; k = k + 1 \quad (12)$$

The exact level of 0.04 defined level out and 96% confidence are reached using this strategy. The features are chosen following feature choice formulation. To diagnose thyroid un wellness, the chosen features are given to the suggested classifier. The next section contains a detailed explanation of the suggested deep learning model.

3.3. Ebola optimization-based deep learning model

In the suggested approach, thyroid disease is identified from the datasets using the EODLM. The dataset can be frequently used to learn and test the network following the network design.The EOA method is used in the ERNN to aid in choosing the best weighting parameter. In this section, the ERNN and EOA algorithms are presented in detail.

3.3.1. ERNN -Elman Recurrent Neural Network

The ERNN is designed using the feed forward NN. The focus of this study's development stage is on BP's learning approach in NN and ERNN, as well as the elements and behaviors of EOA that should be considered in NN and ERNN. The research's goal is to improve training methods through the use of EOA. This study supports the evolution of ERNN training from a broad BP technique to a version based on an EOA. The following is an explanation of the feed forward NN and ERNN.

Feedforward Neural Network

Biological neural networks are the foundation of the feedforward ANN data processing technique. Despite technical limits and constraints on computer capacity, the main objective of NN is to process information. The classification and detection of thyroid disorders use data processing. Additionally, the procedure in the advanced, effective systems is really difficult. The NN designs a high count of neurons that are interconnected and communicating together to solve the thyroid detection. This structure is designed of various layers and every layer contains many connected neurons [18]. The NN transmits data between the outer and inner layers while being connected to the outside world. Artificial neurons operate similarly to organic neurons in terms of function. The artificial neuron is a crucial parameter that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

functions as a data source and regulates a certain kind of training process. The artificial neurons were operated and then brought to the neuron's incentive frequency using input from various information sources. Inputs and outputs used by the NN can be connected using two different weights. The neuron in this phase uses an activation function technique to adjust the overall weight of inputs in a nonlinear manner to reach Y quantity.

$$Y = \sum_{I=1}^{D} W_I X_I + B \qquad (13)$$

The output of the neuron Y determined about the commencement purpose technique and the neuron data. B can be included in the design in this case to define a biased parameter. It is regarded as 1 when it is steady and seen as a weight. The most common and commonly used activation functions are the hyperbolic tangent and sigmoid function, the sigmoid function, and the logistic function of the sigmoid volume. The classifier segments training instances into compact feature clusters to extract the features. The cluster features that are defined as tokens are the bone and air conduction audiology references. A specific decibel and frequency degree that the patient hears at a pure tone might reflect the reference to audiology hearing. Every learning example saved as a token list could have two parameters, like 1 or 2. It is treated as 1, in the training sample, and as 0, otherwise. The ANN training process makes use of a combination of training instances, probabilistic and full features, and their conditional probabilities.

Elman recurrent neural network

The general ERNN contains three layers managed as X, Y and Z horizontally combined with the context units U in this hidden layer. The structure of the ERNN is given in the figure 2. The general version of the ERNN uses BP functions to save the activations of hidden units and input units throughout the learning phases [19]. The BP outputs an original count between 0 and 1 by executing the sigmoid function which is presented as follows,

$$Y = \frac{1}{(1 + E - sum)} \qquad (14)$$

Here, different parameters are to be considered in the back propagation, like minimum error, transfer function or activation function, momentum rate, and learning rate. This parameter and the convergence of the BP learning algorithm are directly related. Here, the major problem with the NN method and the ERNN is that it might therefore become formed in local minimum when it learns and propagates across networking. Convergence problems are the primary concern, which arises since the NN design itself has several layers such as including hidden, output, and input layers. Therefore, the complexity of the design for databases increases as the parameter increases

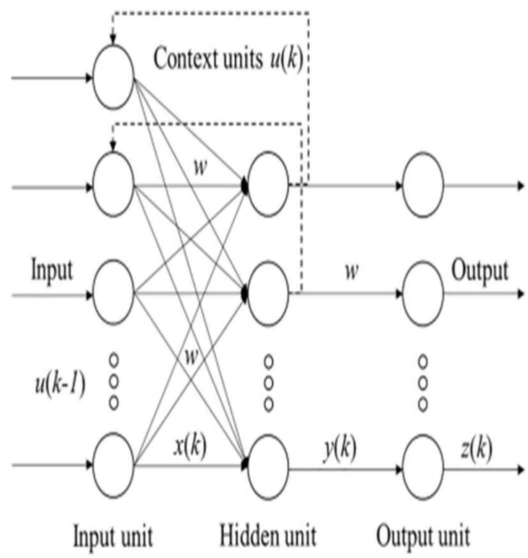


Figure 2: ERNN architecture

.3.3.2. Ebola optimization algorithm

The EOA is a metaheuristic algorithm utilized to select the optimal weighting parameter of ERNN. To formalize the EOA, the SEIR model is utilized. The SEIR model is designed by SEIR-HDFVQ such as Infected (I), Susceptible (S), Recovered(R), Vaccinated (V), Hospitalized (H), Death or dead (D), Exposed (E), Quarantine (Q)and Funeral (F).Additionally, it was discovered when creating the model that a small count of recovered stages could still have the virus present in their bodily fluids and be capable of posing a threat to healthy people. The most effective weighting parameter is selected using the model [20]. In this part, the EOA's mathematical model is presented. To upgrade the positions of every exposed individual, the formulation is presented as follows,

$$Mi_I^{T+1} = Mi_I^T + \rho m(i) \qquad (15)$$

Here, $M(i)$ is defined as the movement rate developed by individuals, Mi_I^T is defined as original positions at time T, Mi_I^{T+1} is defined as updated positions at time $T+1$. ρ is characterized as the scale factor of displacement individual.

$$m(i) = SRATE * RAND(0,1) + m(Ind_{BEST}) \quad (16)$$

$$m(s) = LRATE * RAND(0,1) + m(Ind_{BEST}) \quad (17)$$

The investigation stage can be based on the fact that the sick person travels farther than the typical neighborhood range $LRATE$. The premise that the contaminated person either appears inside a time interval of null or is replaced within a time frame not to exceed $SRATE$will determines how the exploitation stage is produced. Moving over a short distance is what is meant by$SRATE$here. The study takes into account the possibility of infection among a large number of S persons the more remote the replacements. According to the aforementioned equations, the and $SRATE$ and $LRATE$ can be controlled by a neighborhood variable in such a way that when the neighborhood is >=0.5, a particular has to travel outside of the neighborhood, empowering to the high infection; otherwise, it stays inside the neighborhood, which reduces infection [21].

Stage 1: Initialization of super-sensitive population

An initial population with zero-based initial positions can be produced by using a random count distribution. The person could be created and presented as follows,

$$Individual_I = l_I + RAND(0,1) * (u_i + l_i) \quad (18)$$

Here, l_i is defined as lower bounds, u_i is defined as upper bounds for the ith individual (i.e., i=1,2, 3...N) in the population size. The actual best chosen is calculated based on the pair of infected individuals in time T as follows,

$$BESTs = \begin{cases} GBEST, Fitness(CBEST) < Fitness(GBEST) \\ CBEST, Fitness(CBEST) \geq Fitness(GBEST) \end{cases} \quad (19)$$

Here, $CBEST$ is described as the actual best solution, $GBEST$ is described as the global best solution, $CBEST$ is described as the actual best solution at period T. The difference between$CBEST$ and $GBEST$ as infected individuals who are spreaders and super broadcasters of virusebola. The fitness function of the EOA is reducing learning rate error in the ERNN network architecture.

The fitness function for training the network

This article selects the loss function as the sum function of learning the network which is presented as follows,

$$C = \frac{1}{2}\sum_{p}^{n}\sum_{k}^{0}(D_{PK} - Y_{PK})^2 \quad (20)$$

Create a neural network with more than two layers to compute historical data that is longer and is capable of being transmitted accompanied by an error parameter. Consequently, the error function is defined as follows:

$$\delta = -\frac{\partial C}{\partial (Net)} \quad (21)$$

$$\delta_{PJ}(T-1) = \sum_{H}^{M} \delta_{PH}(T)\, U_{HI} F^{`}\left(\delta_{PJ}(T-1)\right) \quad (22)$$

Here, J can be defined as the hidden layer node index with the period $T-1$, H can be defined as the hidden layer node index with the period T.

Algorithm 1: Pseudocode of EOA
Output: Optimal weighting parameter
Input: Objective function, lb, ub, epoch, population size, evdincub

$S, E, I, H, R, V, Q, Sols \leftarrow \emptyset$
$S \leftarrow Createsusceptibleindud(psize, S)$
$icase \leftarrow generatedIndexCase(S);$
$GBEST, CBEST \leftarrow icase$
$While\ e \geq epoch\ \wedge Len(I) > 0\ do$
$Q \leftarrow RAND(0, Eq\ 38 \times i);$
$fraci = i - q;$
For $i \leftarrow 1$ to $Len(fraci) do$
$POS_i \leftarrow movrate()$
$d_i \leftarrow RAND();$

$Newi \leftarrow \emptyset;$
If $d_i > evdincub$ then
$neighborhood \leftarrow prob(POS_i);$
If $neighborhood < 0.5$ then
$tmp \leftarrow RAND(0, eq\ 33 \times i \times SRATE)$
End
Else
$tmp \leftarrow RAND(0, eq\ 33 \times i \times IRATE)$
End
$Newi + \leftarrow tmp;$
End
$i + \leftarrow newi$
End
$H \leftarrow RAND(0, eq\ 25 \times i), h + \leftarrow H;$
$R \leftarrow RAND(0, eq\ 26 \times i), r + \leftarrow R;$
$V \leftarrow RAND(0, eq\ 27 \times i), v + \leftarrow V;$
$D \leftarrow RAND(0, eq\ 2837 \times i), d + \leftarrow D;$
$I + \leftarrow i - Add(R, D);$
$S + \leftarrow R;$
$S + \leftarrow D;$
$CBest = Fitness(objfunction, i);$
$if\ GBest > GBest$ then
$GBest = CBest;$
$Sols \leftarrow GBest;$
End
End
Save the optimal weighting parameter

Updating process

Update of quarantine (Q), funeral (F), recovery (R), immunization (V), exposure (E), hospitalization (H), infection (I), susceptibility (S), and an arrangement of common computation equations deregulate dead (D).Calculus is a branch of mathematics, and one of its branches is differential calculus. In the former, the range of modification of one conception about another is dealt with, whereas in the latter, it is dealt with computing various properties of derivatives and integrals. In this instance, differential calculus was used to calculate the rate of change of the variables Q, D, V, R, H, I, and S as they relate to the time T. In light of this, the equations are stated as follows:

$$\frac{\partial S(T)}{\partial T} = \pi - \left(\beta_1 i + \beta_3 d + \beta_4 r + \beta_2 (PE)\right)S - (\tau S + \Gamma i)) \quad (23)$$

$$\frac{\partial i(T)}{\partial T} = (\beta_1 i + \beta_3 d + \beta_4 r + \beta_2 (PE)\lambda)S - (\Gamma + \gamma)i - (\tau)S) \quad (24)$$

$$\frac{\partial h(T)}{\partial t} = \alpha i(\gamma + \varpi)h \quad (25)$$

$$\frac{\partial r(T)}{\partial t} = \gamma i - \Gamma r \quad (26)$$

$$\frac{\partial v(T)}{\partial t} = \gamma i - (\mu + v)v) \quad (27)$$

$$\frac{\partial d(T)}{\partial t} = (\tau s + \Gamma i) - \delta d \quad (28)$$

$$\frac{\partial d(T)}{\partial t} = (\pi i - (\gamma r + \Gamma d)) - \xi q \quad (29)$$

The aforementioned equations are scalar functions, which means that each of them has a single parameter that is a float-defined number.This is not far off from a handful of general scalar computation equations and their associated f functions, such as population increase or exponential money evolution.

$$u^{`} = \alpha u \quad (30)$$

Here, u is defined as the growth rate. Here, calculates the population's rate of change and applies it to the vector's current size to determine the population's susceptible person count at periodT.A similar process is proceeding to estimate the set of individuals in vectors q, v, r, h, i and q exploitation rates. This research assumes the initial conditions. The classifier's ideal weighting parameter is calculated based on the EOA. After that, the efficient training process is enabled which empowers the thyroid classification.

4. RESULTS AND DISCUSSION

The expected automatic thyroid detection is analyzed and verified in this part. The normalized database is compiled from online sources to verify the indicated thyroid detection [36]. The assembled database contains both the 7201 databases and the three classes. The network is trained using 80% of the database's thyroid data. Additionally, the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

network is tested using the remaining 20% of the thyroid data. The projected strategy is executed in MATLAB, and the presentation is validated by considering statistical indicators such as kappa, F-Measuree, AUC, precision, sensitivity, recall, ROC, accuracy, and specificity. The suggested technique is compared to the current methods, including SVM, EFERNN, ANN, and DBNN. Table 2 provides the suggested method's implementation variables. The projected technique, which comprises the confusion matrix, is determined to be valid based on the statistical measures. Table 3 displays the example attribute data.

Table 2: Simulation variables

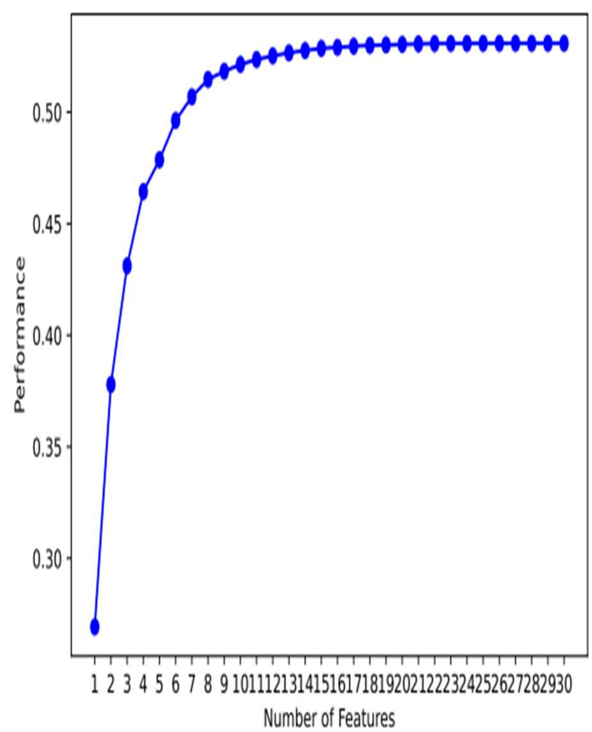
S. No	Name of the Variable	Parameters
1	Stop requirements with minimal error	0.005
2	Maximum iteration/Limit	500
3	The minimum value of weight	-0.50
4	The maximum value of weight	0.50
5	Total count of bias and weight	N
6	Search agents	20
7	Dimension	10
8	Upper bound	500
9	Lower bound	200
10	Hidden nodes	2
11	Learning rate	0.3
12	Training time	600s

Table 3:Parameters of Confusion matrix

N=7201		Expected		
		Class 3	Class 2	Class1
Actual	Class 1	1	1	2998
	Class 2	1	2800	0
	Class 3	1404	1	2

Table 4: Sample attributes of data

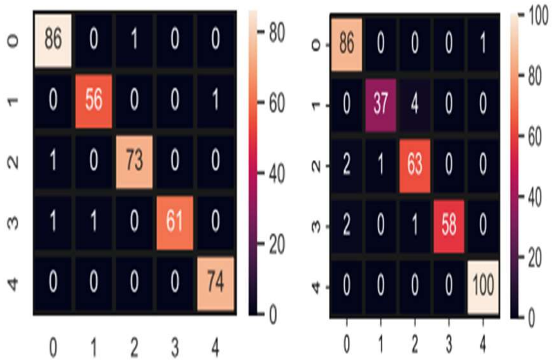
Description	Attribute	Data type
The singular id of the patient	patient id	str
Hyperthyroidism medical diagnosis	target	str
TBG level blood from lab work	TBG	float
FTI level in blood from lab work	FIT	float
If FTI was premeditated in the blood	FTT measured	bool
T4U level in blood from lab work	TT4	float
If TT4 was premeditated in the blood TT4 measured	bool	
T3 level in blood from lab work	T3	float
If T3 was premeditated in the blood	T3 measured	bool
TSH level in blood from lab work	TSH	float
H was premeditated in the blood	TSH measured	bool
Persevering psych	psych	bool
patient hyper pituitary gland	hypopituitary	float
patient has tumor	tumor	bool
patient has goitre	goitre	bool
patient lithium	lithium	bool
The patient expect they have hyperthyroid	query hyperthyroid	bool
The patient expect they have hypothyroid	query hypothyroid	bool
patient is undergoing 1131 treatment	1131treatment	bool
patient has suffered thyroid surgery	thyroid surgery	bool
patient is pregnant	pregnant	bool
patient is sick	sick	bool
whether the patient is on antithyroid meds	on antithyroid meds	bool
whether patient is on thyroxine	on thyroxine	bool
sex patient identifies	sex	str
age of the patient	age	int



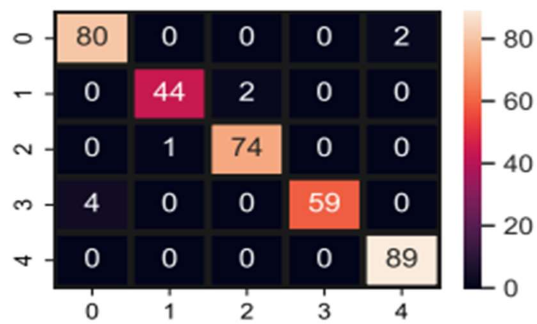
Journal of Theoretical and Applied Information Technology
30th April 2026. Vol.104. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 3: Features



(a)



(c)

Figure 4: Feature selection technique (a)Backward Feature Elimination, (b) Forward Feature Selection and (c) Bidirectional Feature Elimination

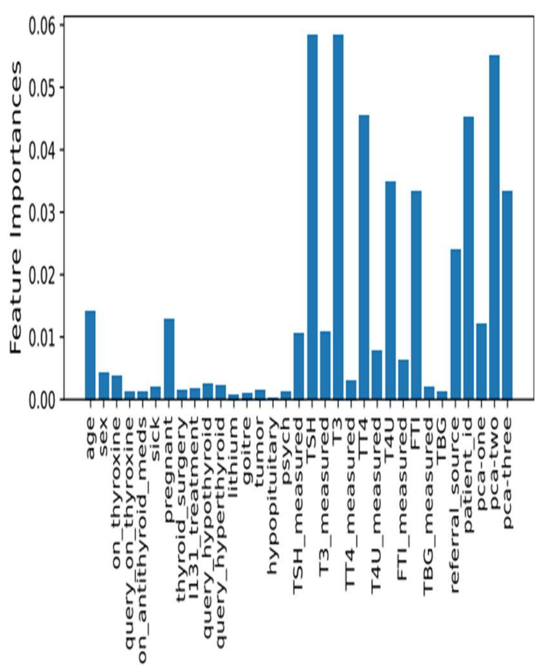


Figure 5: Feature importance

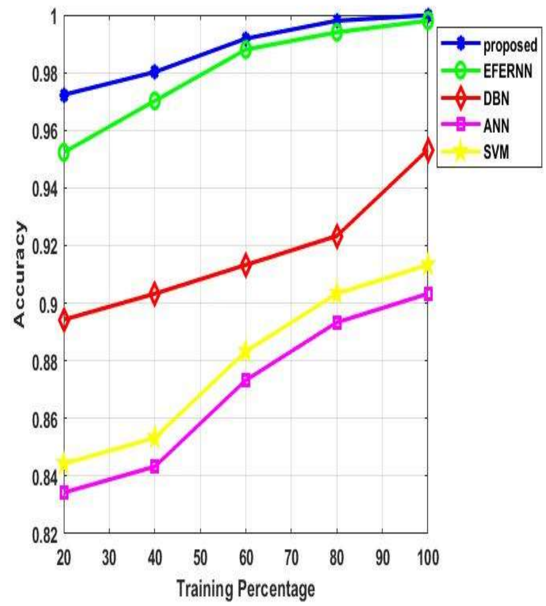


Figure 6: A comparison of accuracy performance

The collected database consists of 30 features utilized for thyroid prediction. The features count with performance is presented in the figure 3. The feature selection technique with the confusion matrix is presented in Figure 4. It consists of three feature selection techniques bidirectional, forward, and backward feature selection. The collected database contains 30 features, and the importance of the features is organized by using the feature selection technique which is presented in Figure 5. Using performance metrics including kappa, F_measure, AUC, precision, accuracy, specificity, ROC, recall, specificity, and sensitivity, the demonstration of the proposed method is evaluated. The accuracy metric is employed to assess the proposed methodology, as illustrated in Figure 6. The technique proposed is contrasted with longstanding methods like ANN, EFERNN, SVM, and DBN. The proposed method has attained an accuracy of 0.999. After that, the ANN, EFERNN,SVM, andDBN attained 0.909, 0.91 and 0.95. With the consideration of accuracy validation, the proposed method achieved efficient results.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

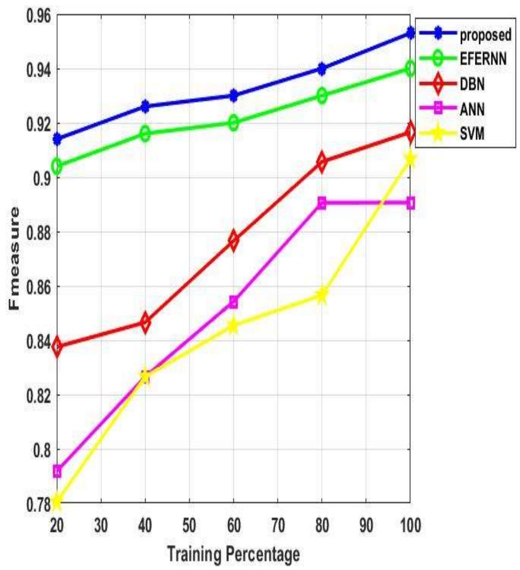


Figure 7: A comparison of F_Measure performance

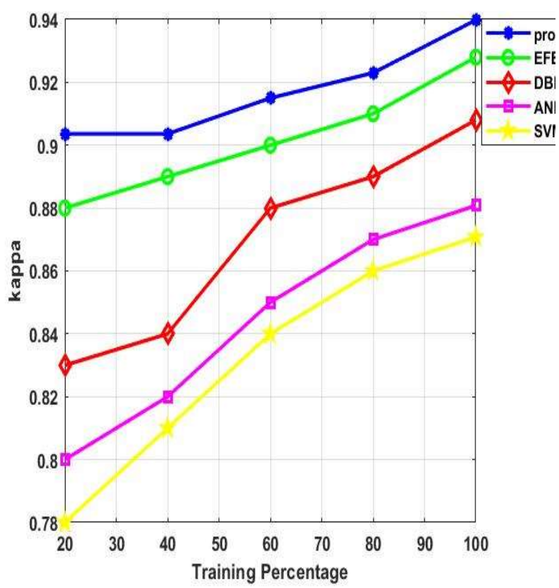


Figure 8: A comparison of Kappa performance

Figure 7 presents the planned technique that is evaluated using the F_measure. The proposed method is contrasted with traditional techniques such as SVM, ANN, DBN, and EFERNN. The F_measure for the suggested approach is 0.95. Following that, the results for the SVM, ANN, DBN, and EFERNN were 0.89, 0.91, 0.915, and 0.94. When F_measure validation was taken into account, the suggested approach produced effective results. The planned technique shown in figure 8 is evaluated using the kappa. The suggested method is compared to more established approaches such as SVM, ANN, DBN, and EFERNN. The kappa in 0.94 has been achieved with the suggested approach. Subsequently, the ANN, DBN, EFERNN, and SVM achieved 0.889, 0.918, 0.91, and 0.93. When kappa validation was taken into account, the suggested approach produced effective outcomes. Figure 9 shows the predicted technique that is evaluated using precision. The suggested technique is compared to traditional methods such as SVM, ANN, DBN, and EFERNN. The proposed method has attained a precision of 0.94. After that, the SVM, ANN, DBN, and EFERNN attained 0.905, 0.918,0.928, and 0.93. With the consideration of precision validation, the proposed method achieved efficient results.

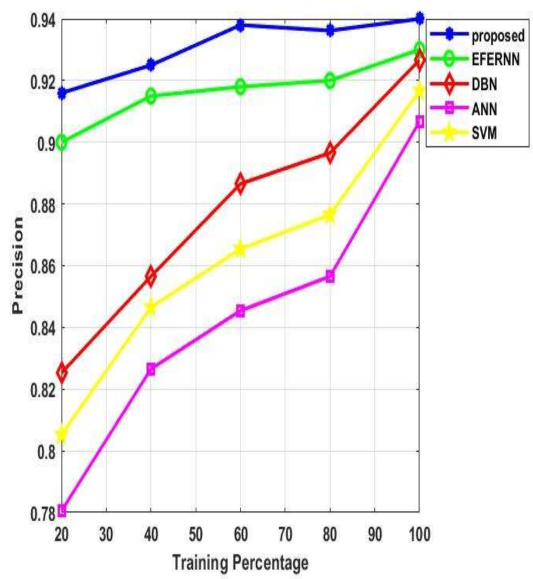


Figure 9: A comparison of Precision performance

Journal of Theoretical and Applied Information Technology
30th April 2026. Vol.104. No.8

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

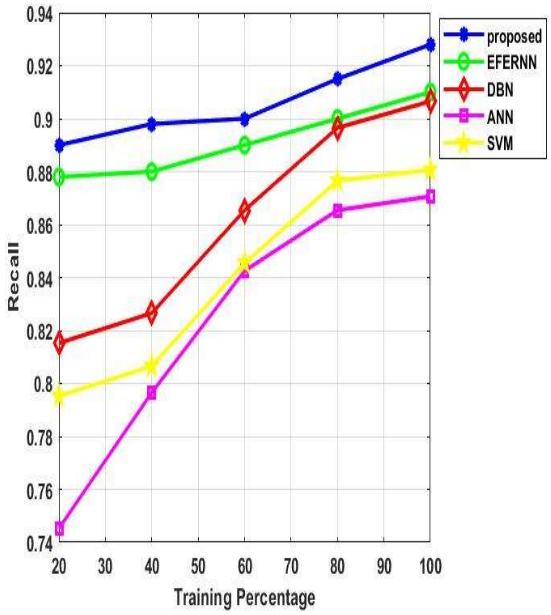


Figure 10: A comparison of Recall performance

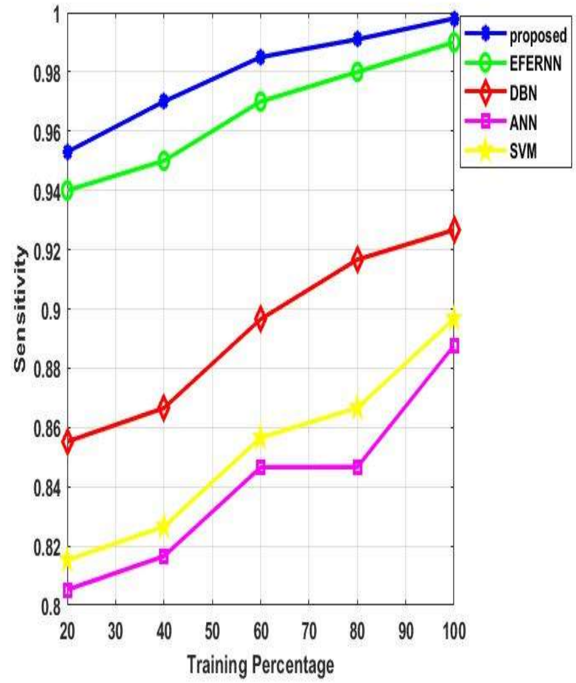


Figure 12: A comparison of Sensitivity performance

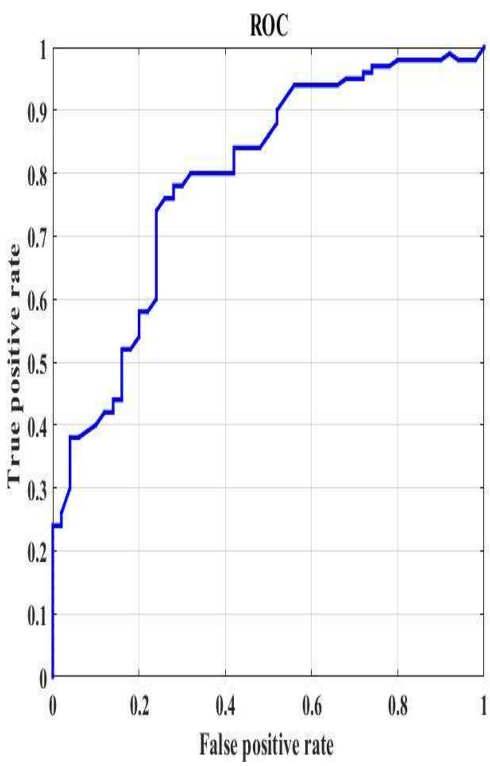


Figure 11: A comparison of ROC performance

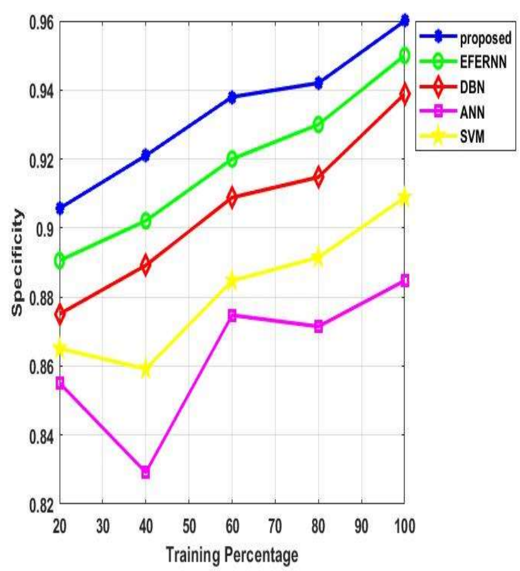


Figure 13:A comparison of Specificity performance

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4:Validation analysis

S. No	Technique	References	Precision	Accuracy	Specificity	Sensitivity	F_measure	Recall
1	hybrid optimization algorithm	Vidhushavarshini Suresh Kumar*et al.,* [11]	0.98	0.92	0.99	0.86	0.85	0.89
2	K-Nearest neighbor (KNN)	Hafiz Abbad Ur Rehman *et al.,* [12]	0.97	0.95	0.89	0.91	0.87	0.85
3	artificial neural network (ANN)	Mehdi et al., [13]	0.91	0.93	0.942	0.92	0.89	0.87
4	homogeneous ensemble approach	Tehseen Akhtar *et al.,* [14]	0.92	0.94	0.920	0.891	0.912	0.926
5	multi-kernel support vector machine	K. Shankar *et al.,* [15]	0.94	0.96	0.91	0.91	0.920	0.950
6	B-LDA (Bayes Linear Discriminant Analysis) and CSWDT (Cuckoo search-based Weighted Decision tree)	-	0.964	0.995	0.93	0.88	0.87	0.89
7	Efficient Feature Extraction Based Recurrent Neural Network (EFERNN)	-	0.93	0.998	0.95	0.93	0.92	0.91
8	Proposed method	-	0.94	0.999	0.99	0.96	0.95	0.93

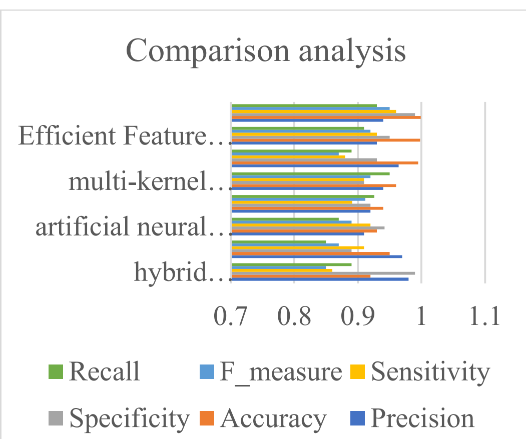


Figure 12: Comparison Analysis

The proposed technique, shown in Figure 10, is evaluated using the recall. The proposed method is contrasted with conventional approaches like SVM, ANN, DBN, and EFERNN. The proposed method has been attained recall in 0.93. After that, the SVM, ANN, DBN, and EFERNN attained 0.87, 0.88,0.908, and 0.91. With the consideration of recall validation, the proposed method achieved efficient results. The ROC characteristics of the proposed technique are given in Figure 11. The suggested method, shown in Figure 12, is assessed using the sensitivity. The suggested technique is compared to traditional methods such as SVM, ANN, DBN, and EFERNN. The recommended method has a sensitivity of 0.999. Then, at 0.88, 0.89, 0.918, and 0.99, respectively, are the SVM, ANN, DBN, and EFERNN. When sensitivity validation was taken into account, the suggested approach produced effective results. The suggested method, shown in Figure 13, is assessed using the specificity. The suggested technique is compared to traditional methods such as SVM, ANN, DBN, and EFERNN. The specificity of the suggested approach is 0.999. After that, the SVM, ANN, DBN, and EFERNN attained 0.885, 0.92,0.94, and 0.95. With the consideration of specificity validation, the proposed method achieved efficient results.

Research Problems and Open Issues

Despite significant progress in multimodal AI for medical diagnostics, several research challenges remain:

- Limited cross-modal reasoning due to simplistic fusion strategies.
- Suboptimal interpretability from tools like Grad-CAM and LIME.
- Poor generalization across clinical settings and institutions.
- Scalability and real-time deployment constraints.
- Lack of standardized evaluation protocols for multimodal systems.

These open issues motivate the development of the proposed EODLM framework, which aims to address interpretability, fusion synergy, and optimization for clinical reliability**.**

5. CONCLUSION

Early identification of thyroid disorders is critical for timely treatment and can be life-saving. Leveraging advanced computational techniques enables automated diagnosis from diverse medical datasets, minimizing the need for manual intervention. This study introduces an Ebola Optimization Deep Learning Model (EODLM) tailored for thyroid disease classification using computer vision and machine learning strategies. The proposed framework encompasses three key stages: data preprocessing, feature selection, and classification. Initially, thyroid-related datasets were sourced from publicly available repositories. To address missing values and ensure consistency, a normalization-based preprocessing method was applied. Feature selection was conducted in three phases—forward selection, backward elimination, and bidirectional refinement—to isolate the most relevant attributes. The classification model integrates an Elman Recurrent Neural Network (ERNN) with the Ebola Optimization Algorithm (EOA), which dynamically tunes the network's weight parameters for optimal performance. The system was evaluated using standard metrics including accuracy, precision, recall, specificity, sensitivity, F1-score, kappa coefficient, and ROC curve analysis. Comparative results indicate that the proposed EODLM outperforms existing models such as B-LDA-CSWDT (96% accuracy) and EFERNN (99.8% accuracy), achieving a superior accuracy of 99.9%.While the model demonstrates high diagnostic precision, it currently lacks real-time deployment capabilities. Future work will focus on integrating the framework into real-time clinical environments for dynamic thyroid detection and classification.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the Work reported in this paper.

Acknowledgments

The authors thank the anonymous reviewers for their helpful comments and suggestions.

Funding

Not applicable

Ethical approval

Not applicable.

Consent to participate

Not applicable.

Consent to publish

Not applicable.

Conflict of interest

The authors declare no competing interests.

Data Availability

No data availability

REFERENCES

[1].Luo, Xianzhu, Rui Wang, Chuanzhu Lv, Guang Chen, Jinmao You, and Fabiao Yu. "Detection of seleno cysteine with a radiometric near-infrared fluorescent probe in cells and in mice thyroid diseases model." *Analytical chemistry* 92, no. 1 (2019): 1589-1597.

[2].Mishra, Sushruta, Yeshihareg Tadesse, Anuttam Dash, Lambodar Jena, and Piyush Ranjan. "Thyroid disorder analysis using random forest classifier." *In Intelligent and cloud computing*, pp. 385-390. Springer, Singapore, 2021.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[3].Pal, Rekha, Tanvi Anand, and Sanjay Kumar Dubey. "Evaluation and performance analysis of classification techniques for thyroid detection." *International Journal of Business Information Systems* 28, no. 2 (2018): 163-177.

[4].Banu, G. Rasitha. "A Role of decision Tree classification data Mining Technique in Diagnosing Thyroid disease." *International Journal of Computer Sciences and Engineering* 4, no. 11 (2016): 64-70.

[5].Sankar, S., Anupama Potti, G. Naga Chandrika, and Somula Ramasubbareddy. "Thyroid Disease Prediction Using XGBoost Algorithms." *Journal of Mobile Multimedia* 18, no. 3 (2022): 1-18.

[6].Juneja, Kapil. "Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification." *Wireless Personal Communications* 126, no. 2 (2022): 1805-1842.

[7].Namdeo, Rajole Bhausaheb, and Gond Vitthal Janardan. "Thyroid disorder diagnosis by optimal convolutional neuron-based CNN architecture." *Journal of Experimental & Theoretical Artificial Intelligence* 34, no. 5 (2022): 871-890.

[8].Vanitha, R., and K. Perumal. "A Novel Classification Methodology for Thyroid Cancer using C4. 5 with Firefly Optimization Algorithm (CFOA)." *Indian Journal of Science and Technology* 15, no. 26 (2022): 1324-1335.

[9].Santin, M., C. Brama, H. Théro, E. Ketheeswaran, I. El-Karoui, F. Bidault, R. Gillet, P. Gondim Teixeira, and A. Blum. "Detecting abnormal thyroid cartilages on CT using deep learning." *Diagnostic and interventional imaging* 100, no. 4 (2019): 251-257.

[10]. Abdolali, Fatemeh, Jeevesh Kapur, Jacob L. Jaremko, Michelle Noga, Abhilash R. Hareendranathan, and Kumaradevan Punithakumar. "Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks." *Computers in Biology and Medicine* 122 (2020): 103871.

[11]. .Sureshkumar, Vidhushavarshini, Sathiyabhama Balasubramaniam, Vijayakumar Ravi, and Ajay Arunachalam. "A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine." *Expert Systems* 39, no. 1 (2022): e12811.

[12]. Abbad Ur Rehman, Hafiz, Chyi-Yeu Lin, and Zohaib Mushtaq. "Effective K-Nearest Neighbor Algorithms Performance Analysis of Thyroid Disease." *Journal of the Chinese Institute of Engineers* 44, no. 1 (2021): 77-87.

[13]. Hosseinzadeh, Mehdi, Omed Hassan Ahmed, Marwan Yassin Ghafour, Fatemeh Safara, Saqib Ali, Bay Vo, and Hsiu-Sen Chiang. "A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things." *The Journal of Supercomputing* 77, no. 4 (2021): 3616-3637.

[14]. Akhtar, Tehseen, Syed Omer Gilani, Zohaib Mushtaq, Saad Arif, Mohsin Jamil, Yasar Ayaz, Shahid Ikramullah Butt, and Asim Waris. "Effective Voting Ensemble of Homogenous Ensembling with Multiple Attribute-Selection Approaches for Improved Identification of Thyroid Disorder." *Electronics* 10, no. 23 (2021): 3026.

[15]. Shankar, K., S. K. Lakshmanaprabu, Deepak Gupta, Andino Maseleno, and Victor Hugo C. De Albuquerque. "Optimal feature-based multi-kernel SVM approach for thyroid disease classification." *The journal of supercomputing* 76, no. 2 (2020): 1128-1143.

[16]. Abbad Ur Rehman, Hafiz, Chyi-Yeu Lin, Zohaib Mushtaq, and Shun-Feng Su. "Performance analysis of machine learning algorithms for thyroid disease." *Arabian Journal for Science and Engineering* 46, no. 10 (2021): 9437-9449.

[17]. Chaganti, Rajasekhar, Furqan Rustam, Isabel De La Torre Díez, Juan Luis Vidal Mazón, Carmen Lili Rodríguez, and Imran Ashraf. "Thyroid Disease Prediction Using Selective Features and Machine Learning Techniques." *Cancers* 14, no. 16 (2022): 3914.

[18]. Ab Aziz, Mohamad Firdaus, Salama A. Mostafa, Cik Feresa Mohd Foozy, Mazin Abed Mohammed, Mohamed Elhoseny, and Abedallah Zaid Abualkishik. "Integrating Elman recurrent neural network with particle swarm optimization algorithms for an improved hybrid training of multidisciplinary datasets." *Expert Systems with Applications* 183 (2021): 115441.

[19]. Sadeghi-Niaraki, Abolghasem, Parima Mirshafiei, Maryam Shakeri, and Soo-Mi Choi. "Short-term traffic flow prediction using the modified elman recurrent neural network optimized through a genetic algorithm." *IEEE Access* 8 (2020): 217526-217540.

ISSN: **1992-8645** www.jatit.org E-ISSN: **1817-3195**

[20]. Oyelade, Olaide Nathaniel, Absalom El-Shamir Ezugwu, Tehnan IA Mohamed, and Laith Abualigah. "Ebola optimization search algorithm: A new nature-inspired metaheuristic optimization algorithm." *IEEE Access* 10 (2022): 16150-16177.

[21]. Oyelade, Olaide N., and Absalom E. Ezugwu. "Ebola Optimization Search Algorithm (EOSA): A new metaheuristic algorithm based on the propagation model of Ebola virus disease." *arXiv preprint arXiv*:2106.01416 (2021).

[22]. Alam K, Siddique N, Adeli H. 2020. A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications* 32:8675–8690

[23]. Zhao J, Zhou X, Shi G, Xiao N, Song K, Zhao J, Hao R, Li K. 2022. Semantic consistency generative adversarial network for cross-modality domain adaptation in ultrasound thyroid nodule classification. *Applied Intelligence* 52(9):10369–10383

[24]. Ai X, Zhuang J, Wang Y, Wan P, Fu Y. 2022. ResCaps: an improved capsule network and its application in ultrasonic image classification of thyroid papillary carcinoma. *Complex & Intelligent Systems* 8(3):1865–1873

[25]. Yadav DC, Pal S. 2022. Thyroid prediction using ensemble data mining techniques. *International Journal of Information Technology* 14(3):1273–1283

[26]. Jandoubi, B., & Akhloufi, M. A., "Multimodal Artificial Intelligence in Medical Diagnostics," *Information*, vol. 16, no. 7, p. 591, 2025. [Online]. Available: https://doi.org/10.3390/info16070591

[27]. Kaczmarczyk, R., Wilhelm, T. I., Martin, R., & Roos, J., "Evaluating multimodal AI in medical diagnostics," *npj Digit. Med.*, 2024. [Online]. Available: https://doi.org/10.1038/s41746-024-01208-3

[28]. Li, J., et al., "CFFormer: Cross CNN-Transformer Channel Attention and Spatial Feature Fusion for Improved Segmentation of Heterogeneous Medical Images," *arXiv preprint*, arXiv:2501.03629, 2025. [Online]. Available: https://arxiv.org/abs/2501.03629

[29]. Yadav, Dhyan Chandra, and Saurabh Pal. "To generate an ensemble model for women thyroid prediction using data mining techniques." *Asian Pacific journal of cancer prevention*: APJCP 20, no. 4 (2019): 1275.

[30]. Jha, Ritesh, Vandana Bhattacharjee, and Abhijit Mustafi. "Increasing the prediction accuracy for thyroid disease: a step towards better health for society." *Wireless Personal Communications* 122, no. 2 (2022): 1921-1938.

[31]. Liu, Xiaoyun, Marco Medici, Norra Kwong, Trevor E. Angell, Ellen Marqusee, Matthew I. Kim, P. Reed Larsen et al. "Bethesda categorization of thyroid nodule cytology and prediction of thyroid cancer type and prognosis." *Thyroid* 26, no. 2 (2016): 256-261.

[32]. Islam, Saima Sharleen, MdSamiul Haque, M. SaefUllah Miah, Talan Bin Sarwar, and RamadhanNugraha. "Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study." *Peer Computer Science* 8 (2022): e898.

[33]. Liu, Wenfei, Shoufei Wang, Ziheng Ye, Peipei Xu, Xiaotian Xia, and Minggao Guo. "Prediction of lung metastases in thyroid cancer using machine learning based on SEER database." *Cancer Medicine* 11, no. 12 (2022): 2503-2515.

[34]. Lai, Shengwei, Yun-long Fan, Yu-hua Zhu, Fei Zhang, Zheng Guo, Bing Wang, Zheng Wan, Peilin Liu, Ning Yu, and Handai Qin. "Machine learning-based dynamic prediction of lateral lymph node metastasis in patients with papillary thyroid cancer." *Frontiers in Endocrinology* 13 (2022): 1019037.

[35]. Duc, Nguyen Thanh, Yong-Moon Lee, Jae Hyun Park, and Boreom Lee. "An ensemble deep learning for automatic prediction of papillary thyroid carcinoma using fine needle aspiration cytology." *Expert Systems with Applications* 188 (2022): 115927.

[36]. https://www.kaggle.com/datasets/yasserhessein/thyroid-disease-data-set

[37]. Nagavali Saka Munjeti Sowjanya, Deep Learning: A Comprehensive Guide to Artificial Intelligence and Neural Networks, Book Bytes International Publications, Vol 1,P-153(2025)