International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Efficient Extraction and Automated Thyroid Prediction with an Optimized Gated Recurrent Unit in Recurrent Neural Networks

Nagavali Saka[*1] and S. Murali Krishna[2]
[1*] Research Scholar, Department of Computer Science and Engineering JNTUA, Anantapur, Andhra Pradesh, India. *Corresponding author Email id: vali214@gmail.com
[2]Professor, Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering (SVCE), Tirupati, Andhra Pradesh, India. Email id: muralikrishna.s@svcolleges.edu.in

Abstract: Computer-aided tools are becoming increasingly important in medical diagnostics. This paper introduces the Efficient Feature Extraction Based Recurrent Neural Network (FERNN) for computer-aided thyroid disease prediction. The FERNN model uses a Gated Recurrent Unit Recurrent Neural Network (GRU-RNN) optimized with the COOT Optimization Algorithm.The study begins by gathering data from an open-source system and preprocessing it using min-max normalization to address missing values. The preprocessed data undergoes a two-level feature extraction (TLFE) procedure. In the first level, a ranked filter feature set technique is used to prioritize features based on medical expert recommendations. In the second level, a variety of metrics, including information gain, gain ratio, chi-square, and relief, are used to rank and select features. A composite measure guided by fuzzy logic is then used to select a judicious subset of features. The FERNN model uses the GRU-RNN to classify thyroid diseases in the databases. To optimise, the COOT optimization method is employed. The model's weights. The FERNN model was put into practise in MATLAB and assessed with a variety of statistical metrics, including kappa, accuracy, precision, recall, sensitivity, specificity, and the F-measure. The proposed methodology was benchmarked against traditional techniques, including the deep belief neural network (DBN), artificial neural network (ANN), and support vector machine (SVM).

Keywords: Thyroid disease classification, Feature retrieval, gated recurrent unit (GRU), recurrent neural network (RNN), Two-level feature extraction (TLFE), Fuzzy logic, Information gain, Gain ratio, Chi-square, Relief, Changes.

1. Introduction

The thyroid, resembling a butterfly in shape, is a significant gland situated in the lower neck region and pivotal in regulating the body's metabolism [1]. It produces two vital hormones, triiodothyronine (T3) and levothyroxine (T4), pivotal for protein synthesis, temperature control, and overall energy management [2]. This gland is susceptible to a range of often perplexing diseases, with hypothyroidism and hyperthyroidism being common occurrences [3]. Insufficient thyroid hormone secretion causes hypothyroidism, and excessive thyroid hormone secretion causes hypertension release. The former leads to weight gain, a swollen neck, and a slow heart rate, while the latter entails elevated blood pressure, a heightened heart rate, and weight loss [4] [5].

In developed nations, an estimated 4–5% of individuals are affected by hypothyroidism. Its potential consequences include elevated cholesterol levels, hypertension, cardiovascular complications, stunted growth in adolescents, and depression [7] and [8]. Thus, educating the populace about its symptoms, categories, and origins is of paramount importance. Diagnosing thyroid disorders often involves blood tests to Check the concentrations of TSH, T3, and T4 [9]. The health domain offers an abundance of intricate data, posing significant challenges. To address this, artificial intelligence (AI) techniques are emerging to detect and diagnose evolving diseases [10].

Experts employ diverse classification methods, including Neural Networks, K-nearest neighbor (KNN), Artificial Neural Networks (ANN), Decision Trees (DT), Naive Bayes, Support Vector Machines (SVM), and Bayesian Networks (BNN), to discern thyroid issues. These methods facilitate effective analysis and management of thyroid-related concerns.

Contribution of the Study:

- This research aims to highlight the significance of extracting essential features from unprocessed medical datasets, supporting medical practitioners in achieving precise thyroid disease diagnoses for the wider population.
- The incorporation of a min-max pre-processing method addresses the removal of irrelevant and absent values from the input data, thereby enhancing data integrity.
- The categorization accuracy is expected to experience notable enhancement due to the implementation of a two-level feature extraction approach.
- The development of an optimized RNN-GRU classifier offers the prospect of proficiently classifying thyroid disease categories present within the input dataset.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

This paper's sections are organized as follows: Section 2 discusses pertinent research endeavors focused on employing deep learning techniques for the purpose of detecting thyroid diseases. In Section 3, the complete suggested system architecture model is delineated. The assessment of the technique's outcomes is elucidated in Section 4. Concluding this work, Section 5 presents a concise summary of the conclusions drawn from this study.

2. Literature Review

In recent years, numerous advancements in deep learning methodologies for automatic thyroid disease detection have been introduced by researchers. This section offers a concise overview of notable research endeavors in this field.

Tahir Alyas et al. [11] devised an empirical AI-driven method for predicting thyroid infections. Various AI techniques, including ANN, KNN, random forest algorithms, and decision trees, were evaluated on a dataset to enhance disease prediction within predefined boundaries. They achieved remarkable accuracy of 94.8% and specificity of 91% using the random forest algorithm.

Hafiz Abbad Ur Rehman et al. [12] Developed variants of the K-nearest neighbours (KNN) algorithm for identifying thyroid disorders. KNN using L1-based selection, KNN with chi-square-based choice of features, and KNN without choosing features were among the variants. KEEL Dataset Collection: Thyroid Tumours were used, as well as a dataset from a Pakistani clinic. The later dataset included three additional attributes: pulse rate, BMI, and blood pressure.

Mehdi Hosseinzadeh et al. [13] developed a neural network for thyroid recognition called multiple multi-layer perceptron (MMLP). They improved the network's learning rate computation by resolving issues with local minima and back-propagation error computations. The MMLP model demonstrated enhanced overall classification performance for thyroid disorders, attaining a 0.7% accuracy improvement when compared to a standalone system.

Rajasekhar Chaganti et al. [14] investigated AI and deep learning feature design strategies. Robustness was introduced into their approach via additional resilience, inverted feature removal, forward choice of characteristics, unilateral component choices, and AI-based tree learners' considerations. Their proposed method accurately identified thyroid disease of the immune system, swollen regulation protein, thyroid disease brought on by the condition known as Ha, and non-thyroid disease (NDIS), with the selected additive tree classifier outperforming others and generating an F1 score of 0.94.

K. Shankar et al. [15] suggested a feature-based multi-kernel SVM method for thyroid identification diseases. Their approach made use of fine component selection as well as a region-based classification process. Using "Multi Kernel Support Vector Machine," they organized data sequences and implemented a comprehensive feature selection method. Through advanced dark wolf development, the model attempted to improve characterization performance by raising the significance of non-essential variables and refining models from varied datasets.

3. The suggested system and methods

A recent report by the Indian daily "Seasons in India" suggests that one out of every ten Indians has hypothyroidism. This estimate is founded on studies carried out by the Indian Thyroid Organization. Hypothyroidism is the tenth most common chronic condition in India, ranking behind asthma, obesity, diabetes, heart disease, and sleep disorders.

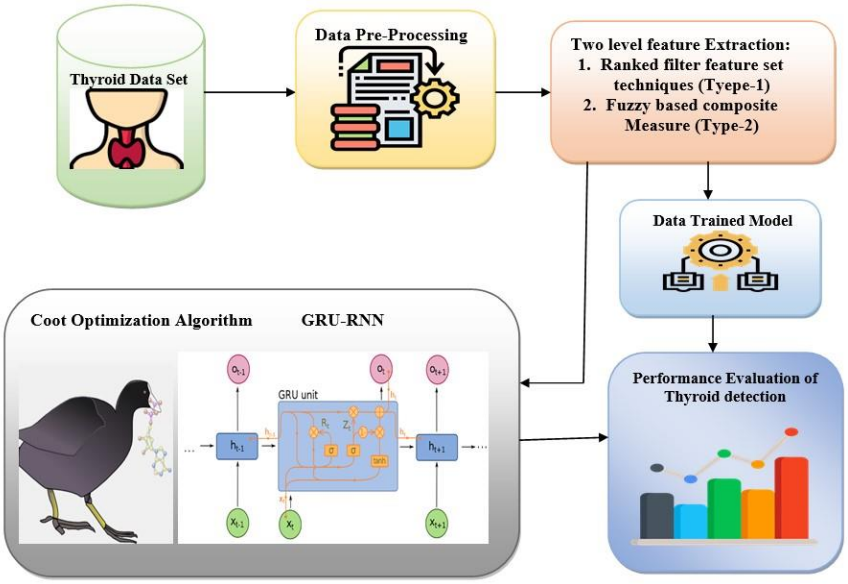

Figure 1. The suggested methodology's block diagram

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Despite the fact that thyroid symptoms can be similar to those of other disorders, only half of the surveyed population is aware of thyroid problems and the availability of symptom tests for diagnosis. This paper proposes a semi-automatic model for the detection and classification of thyroid diseases using database resources, called the EFERNN model. The potential of artificial intelligence (AI) to provide intelligent and rational solutions to a variety of problems makes it a valuable tool in the medical field. As society relies on productive individuals, maintaining good health is essential. People who are ill expend a significant amount of energy on their health, which detracts from their ability to fulfill their obligations and excel. 1. The open-source system starts by acquiring the necessary databases. A preprocessing technique is then introduced to address missing values through normalization methods. Following preprocessing, the data undergoes feature extraction using the Two-Level Feature Extraction (TLFE) approach. In the initial level (Level 1), to prioritize characteristics using suggestions, the ranked filtering collection method is applied. From medical experts. This process employs ranking measures and label-driven validation. Progressing to Level 2, a variety of metrics, including information gain, gain ratio, chi-square, and Relief-F, are employed to rank and select features. Within this stage, a composite measure based on fuzziness identifies effective features within the feature set. Finally, the GRU-RNN model is applied for the classification of thyroid diseases using the databases. The COOT Optimization Algorithm is utilized to determine optimal weights for the GRU-RNN.

3.1. Data Pre-processing

Normalization plays a pivotal role in the initial data preprocessing phase, purging inessential information and mitigating contextual noise present in the acquired databases. Employing the Min-Max normalization technique [16], the data is subjected to a linear transformation that aligns it with a predefined range. This method effectively preserves the inherent relationships within the data. Characterized by its simplicity, the technique seamlessly fits the data within predetermined boundaries. The min-max normalization technique can be succinctly represented as follows:

Min-Max normalization technique =

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

Here, X the term "Input Data" signifies the dataset to be normalized, while d, c "Minimum Value" and "Maximum Value" represent predefined range boundaries. Following the min-max normalization process, the normalized information is subsequently channeled into the feature extraction phase. This normalization technique serves as an essential preparatory step, ensuring that the subsequent analytical processes are conducted on data that has been refined and standardized.

3.2. Level 1: Feature Extraction:

This phase of analysis is geared towards identifying thyroid illness type and presence within a patient. The database encompasses an array of 29 diverse variables, encompassing patient query data, age, and gender, an array of test measurements, physiological features, symptoms, and patient information. The query data component records pertinent information regarding the patient's history of hyperthyroid and hypothyroid instances, antithyroid medication, and thyroxine intake. Accurate data is meticulously collated to differentiate between high-level and low-level symptoms. The latter includes factors such as lithium levels, 1131 infection, and the presence of goiter. High-level symptoms encompass physiological shifts such as pregnancy and nausea. Overall, the patient data indicates a prevalent inclination towards relying on medical consultation and comprehensive thyroid assessment for issue diagnosis [17]. Table 1 effectively summarizes low-level and high-level inquiries, relevant clinical aspects discussed with medical professionals, and crucial recommendations. The table illuminates the significant impact that lithium levels and the presence of 1131 infection can exert on thyroid hormones. These medical endorsements assume the role of weighted factors, which, when factored into feature matrices, facilitate the generation of ranked and filtered feature pairs. This approach underscores the meticulousness of the analysis, drawing insights from a comprehensive dataset to extract and rank features that hold the potential to be critical indicators of thyroid disease.

Table 1: Medical Expert-Recommended Level 1 Feature Extraction (Symptoms

Doctor Advice	Description/Effect	Features
Low	Emotional, physical, and mental impact. Varied detection due to thyroid.	Psych
Low	Thyroid-related sickness with multiple potential causes.	Sick
High	Affects thyroid function. Associated with hypoglycaemia and hyperthyroidism, and goitre.	Lithium
Medium	Influences growth and blood pressure, rare disorder. Relates to hypopituitarism.	Hypopituitary
Low	Subtle symptoms, often present in specific cancer at a later age.	Tumor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Medium	Enlargement of the thyroid gland.	Goiter
Low	Hormonal and size changes during pregnancy.	Pregnancy
Low	Risk of complications during thyroid surgery is below 2 percent.	Thyroid surgery
High	Destruction of thyroid cells.	1131

Certain physiological symptoms, like high blood pressure, hair loss, or weakness, don't necessarily point to thyroid illness. To accurately assess and detect thyroid disorders, medical professionals administer a range of blood tests, including the Free Thyroxine Index (FTI), Triiodothyronine (T3), Thyroid-Stimulating Hormone (TSH), Thyroxine (T4), and Thyroxine-Binding Globulin (TBG). These tests form the cornerstone of low-level evaluations, enabling the identification of thyroid abnormalities and specific disorder types. One notable Diagnostic tool is the blood test, which provides insights into the nature and anomalies of thyroid disease. Table 2 offers a comprehensive summary of the relationship between abnormality impact, associated symptoms, disease types, and the influence of anomalies. This approach underscores the significance of employing comprehensive blood tests in combination with a thorough analysis of related symptoms.

Table 2: Medical Expert-Recommended Level 1 Feature Extraction Tests

Doctor Advice	Symptoms	Features
Low	Hypersensitivity, hair loss, tiredness, dryness	TBG (Thyroxine Binding Globulin)
High	-	FT1 (Free T4)
High	Test for blood	TSH (Thyroid Stimulating Hormone)
Medium	Hair loss, dryness, sensitivity, weight loss, weariness, weakness	T3 Test
High	High blood pressure, hair loss, eye problems, dryness	TT4 Test

The preceding table serves to elucidate both the functional attributes and the comprehensive scope associated with low and high abnormality scenarios. These diagnostic tests serve as potent tools for detecting a spectrum of thyroid diseases. By synthesizing insights from medical professionals' test recommendations and symptom-based evaluations, Level 1 features are meticulously extracted to facilitate thyroid detection. These extracted features subsequently undergo scrutiny in the context of Level 2 feature ranking and selection measures. This second-tier evaluation rigorously identifies and selects the indispensable features that hold the utmost potential for accurate thyroid disease detection. The cohesive interplay between the two levels underscores the methodology's precision in delineating crucial features, ultimately contributing to the model's ability to discern thyroid illnesses effectively.

3.3. Fuzzy-Based Composite Measure for Level 2 Feature Ranking and Selection

This section explores the detailed process of feature selection and ranking at Level 2, using a sophisticated fuzzy-based composite measure.

Information Gain (Info Gain): By adopting an information theory-based approach, Info Gain evaluates the relevance of features in relation to their association with the corresponding class [18]. This assessment metric is computed using training set data and can be formulated as follows:

$$iv\emptyset o(\Sigma) = -\sum_{j=1}^{k} P(A_J)\, Log_2 P(A_J) \quad (2)$$

Here, denote pairings of classes, shows the number of variables inside a certain sample set allocated to a class, and relates to the number of classes associated with thyroid disease disorder. Making use of the gathered data, specific independent attributes are ranked, favoring higher values as more significant. This technique computes Information Gain parameters for various features within thyroid datasets, enhancing the proposed model's efficacy through the incorporation of the fuzzy-based composite measure.

Relief: Primarily applicable to binary classification problems, Relief functions as an effective feature selector. This technique assesses feature significance by computing variations among closest neighbors. Hit and miss computations, pertaining to class specification, are incorporated to generate weighted scores validating the relevance of features.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Chi-Square: This statistical measure quantifies relationships between parameters, functioning as a coefficient to gauge similarities.

$$\varkappa^2 = \sum_{I=1}^{K} \sum_{J=1}^{C} \frac{(N_{IJ} - \mu_{IJ})^2}{\mu_{IJ}} \quad (3)$$

Here, N_{IJ} signifies the count of samples, corresponds to K various parameters, and C denotes different classes.

Gain Ratio: An improved version of Information Gain, Gain Ratio operates within normalized information to mitigate bias by considering higher attributes.

$$SInfo = -\sum_{I=1}^{V} \frac{|d_I|}{d} \times Log_2\left(\frac{|d_I|}{d}\right) \quad (4)$$

Here, d can be defined as a data parameter and V can be defined as an attribute parameter. The gain ratio is presented as follows,

$$GR = \frac{Gain}{SInfo} \quad (5)$$

The relevance of the traits is determined using the measures mentioned above. Fuzzy rules are developed to identify critical traits across multiple metrics in order to do this, as exemplified in the following.

$$INFOgain = \begin{matrix} INFOgain \geq a & Significant \\ INFOgain < a & Nn\ significant \end{matrix} \quad a = 0.005 \quad (6)$$

$$GAIN\ ratio = \begin{matrix} GAIN\ ratio \geq a & Significant \\ GAIN\ ratio < a & Nn\ significant \end{matrix} \quad a = 0.2 \quad (7)$$

$$Chi\ Square = \begin{matrix} Chi\ Square \geq a & Significant \\ Chi\ Square < a & Nn\ significant \end{matrix} \quad a = 50 \quad (8)$$

Here, a represents an optimized fuzzy decision parameter. These K composite fuzzy measures guide the selection of significant features, as depicted in the following equation:

$$fMeasure = \begin{matrix} (sg \cap sg \cap sg \cap sg) \cup (nsg \cap sg \cap sg \cap sg) & High \\ (nsg \cap sg \cap sg \cap sg) & Medium \\ otherwise & Low \end{matrix} \quad (9)$$

This approach ranks and selects features for accurate thyroid detection and categorization. These selected attributes are sent into the ultimate classifier to identify thyroid disorders from aggregated databases. The next section describes this classifier's structure.

This level of granularity underscores the complexity and thoroughness of the feature selection process, utilizing sophisticated techniques to pinpoint and incorporate the most relevant attributes for robust thyroid detection and classification.

3.4. Gated REU Recurrent Neural Network (GRU-RNN) with COOT Optimization

The GRU-RNN is a key component of the proposed system to identify hypothyroidism disorders in assembled datasets. Using the preset network structure, the dataset is used to train and evaluate the network. The COOT technique is used to determine the appropriate weighting parameter within the GRU-RNN architecture. This section digs into the inner workings of both the GRU-RNN and the COOT algorithms, providing an in-depth understanding of their capabilities.

3.4.1. GRU-RNN

Recurrent Neural Networks (RNNs) enhance conventional neural networks by enabling neurons within the same layer to transmit data to one another. This architecture lends itself to superior functionality, particularly in handling time sequences, making RNNs well-suited for time series operations [19]. Figure 2 visually represents the design architecture of a recurrent neural network.

In this figure 2, W represents hidden-to-hidden weight matrices, V signifies hidden-to-output weight matrices, corresponds to U input-to-hidden weight matrices, Y denotes the predicted outcome, S represents the hidden state, and X represents the input. The RNN time series model validates theS_{t-1} network's characteristics and X_i state over time. The combination of the prior time state and current input computes the S neuron state at the specified time as follows:

$$S_t = M(UX_{t-1} + WS_{t-1} + B_H) \quad (10)$$

Here, *z* can be defined as a bias term, and *f* can be defined as a mechanism for actuation. As a neuron, it is used as outputs at time *t* and the network state input at the next time *t*+1 at the same time. Hence, *z* isn't directly linked to the final result. It needs: to divide by the factor Z, then add using the offset function. The equation that follows can be used to define this method mathematically:

$$Y_t = ACT(ZS_t + B_Y) \quad (11)$$

Different from LSTM, GRU does not have an output gate. Instead, one gate combines the input gate and forget gate, further merging the cell state of the hidden state into a single phase. As a result, GRU is simpler and offers faster training, making it a preferred choice. The architecture of the GRU cell is visualized in Figure 3.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

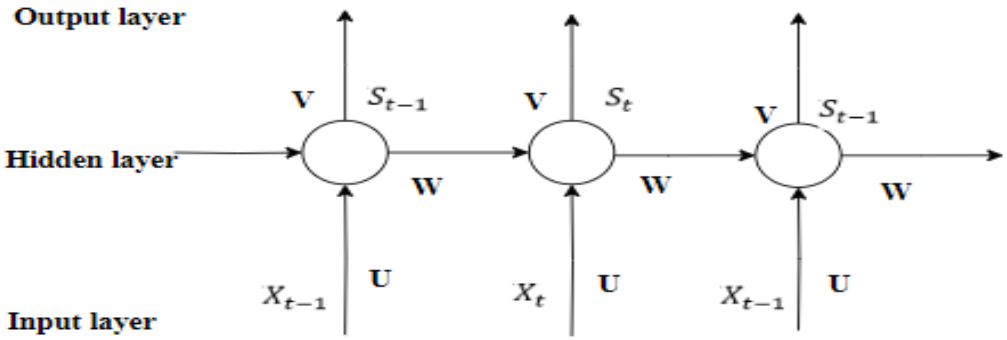

Figure 2: Visualization of Recurrent Neural Network Architecture

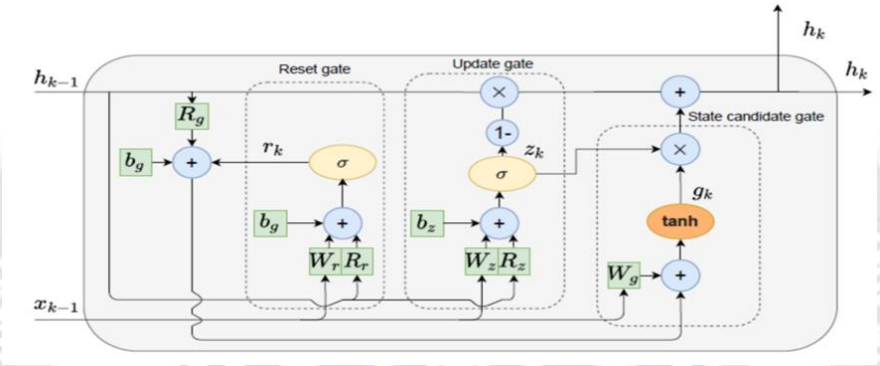

Figure 3: The architecture of the GRU

Figure 3 portrays the GRU architecture. The hidden state can be calculated using the expressions in this section.

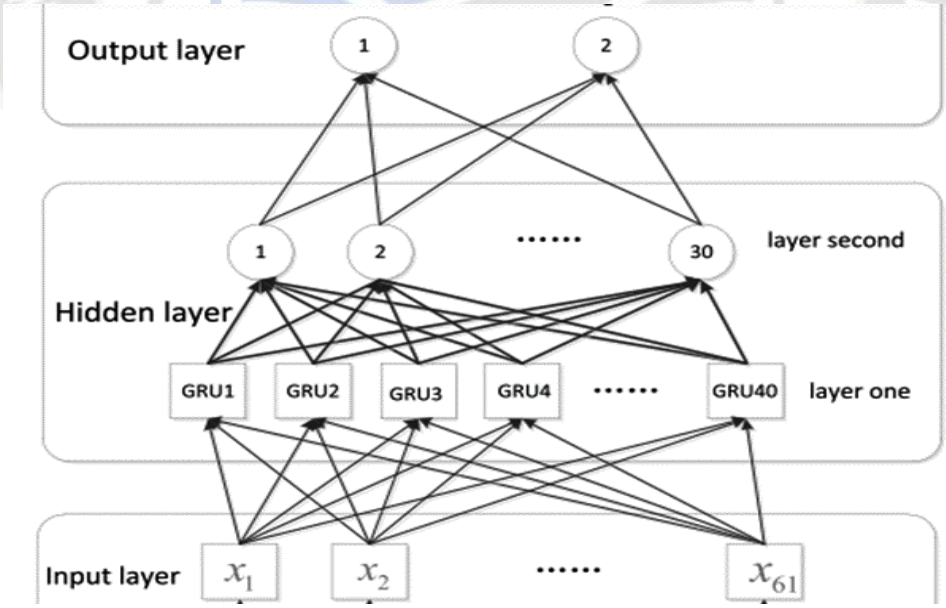

Figure 4 presents the overarching architecture of GRU-RNN.

If discarding the information from the state of concealment prior to inputs is necessary, the reset gate $r(K)$is employed. The update gate $z(k)$ manages the amount of data to retain and pass on to subsequent phases. By multiplying the output of the reset gate from the previous phase, unnecessary data from the prior phase can be eliminated. To put it differently, if the output of the Zupdate gate tends towards zero, the current state incorporates newer data; conversely, if the output approaches one, the current data is retained from the last iteration. The equations below formalize the aforementioned details within a specified time interval$k$:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

$$h(k) = (1_n N \times 1 - z(k)) \times g(k) + z(k) \times h(k-1)) \quad (12)$$

$$g(k) - tanh\left(w_g x(k) + z(k) \times R_g h(k-1) + b_g\right) \quad (13)$$

$$z(k) = \sigma(w_z x(k) + R_z h(k-1) + b_z) \quad (14)$$

$$r(k) = \sigma(w_r x(k) + R_r h(k-1) + b_r) \quad (15)$$

Here, R and W signify learned weight matrices, ×represents element-wise multiplication, denotes the σ logistic sigmoid function, h corresponds to candidate activation, g represents the activation function, signifies the update gate, and $r$denotes the reset gate. The GRU-RNN was tailored for thyroid disease prediction. Within the GRU-RNN network model, the GRU module that captures information from the thyroid data emerges as a critical design. To optimize the weighting parameter within the GRU-RNN, the COOT optimization algorithm is harnessed. The forthcoming section offers a comprehensive delineation of the COOT algorithm's mechanics. This comprehensive exposition clarifies the intricate dynamics of the GRU-RNN, shedding light on its integral role in thyroid disease detection. The subsequent integration of the COOT optimization algorithm showcases a well-rounded approach to enhancing the model's performance.

3.4.2. COOT Algorithm for Optimization

In the architecture of the GRU-RNN, the weight parameter's selection is a crucial step accomplished through the application of the COOT optimization algorithm. To mitigate the potential compromise in system accuracy due to random weight updates within the GRU-RNN, the COOT optimization algorithm is adopted. This ensures the identification of the optimal weight parameter, thereby enhancing the accuracy of thyroid disease identification. This section presents a comprehensive elucidation of the COOT optimization algorithm [21].

The term "coots" refers to a type of water bird found in the rail family, specifically the genus "Rallis." The term "Felica," denoting these birds, originates from the Latin term for "coot." Coots display diverse collective behaviors, primarily focused on mimicking flocking movements. A small cluster of coots positioned at the forefront of the flock serves as cluster leaders, guiding the entire group toward a common destination.

Four distinct coot movements on water surfaces are taken into account:

- Random deviations to both sides.
- Sequential deviations.
- Positional adjustments relative to cluster leaders.
- Enhancement of the cluster's positioning through the leaders toward an optimal location.

This algorithm's mathematical model is detailed in the following section.

The fundamental structure of the complete optimization algorithm closely resembles that of other meta-heuristic algorithms. The coots' positions correspond to the weights of the GRU-RNN networks. The algorithm commences with an initial random population, expressed as:

$$\left(\vec{X}\right) = \left\{\overrightarrow{X_1}, \overrightarrow{X_2}, \ldots., \overrightarrow{X_N}\right\} \quad (16)$$

The goal function is continuously applied to this randomized population, resulting in the following final value:

$$\left(\vec{O}\right) = \{O_1, O_2, \ldots, O_N\} \quad (17)$$

The fundamental principles underpinning the optimization approach are harnessed to further refine it. Population-based optimization methods are employed to search for optimal solutions across a spectrum of optimization problems, without the assurance of reaching a solution within a single iteration. The probability of attaining global optimality improves with an ample number of random solutions and optimization cycles. The population is generated randomly within a predefined range, following this mathematical formulation:

$$CootPOS(I) = RAND(1, D).* (UB - LB) + LB \quad (18)$$

Here, UB and LB denote the upper and lower bounds of the search space, D represents the number of variables or problem variables, and $CootPOS(I)$ corresponds to the coot's position. Each parameter can encompass various upper and lower bounds:

$$UB = [UB_1, UB_2, \ldots, UB_D], LB = [LB_1, LB_2, \ldots, LB_D] \quad (19)$$

Upon constructing the initial population, the position of each search agent must be determined, and the fitness of each solution must be evaluated using the objective function:

$$O_I = f\left(\vec{X}\right) \quad (20)$$

This equation represents the fitness function. In this scenario, the NL number of coots serves as the cluster leader, with leaders being randomly selected. The algorithm is revised in light of the four motions of the coots.

Fitness Evaluation

Fitness function parameters are computed using the variables of the function of fitness for every coot derived from the difference between detection and its corresponding observations:

$$MSE_J = \frac{1}{N}\sum_{T=1}^{N}\left(x_T - \widehat{x_T}\right)^2, J = 1,2,.., pn \quad (21)$$

Random deviations to both sides.

To effect random changes within the search space, connectedness is established through the formula below. The coot is then moved in that direction:

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

$$q = RAND(1, D).* (UB - LB) + LB \quad (22)$$

This variation allows the coot to explore different regions of the search space. If these modifications impact the local optimum, they can guide the technique away from it. The new position of the coot can be calculated using the following formula:

$$CootPOS(I) = CootPOS(I) + a \times r_2 \times (q - CootPOS(I)) \quad (23)$$

Here, r_2 represents a random variable, and signifies a random number within the range [0, 1].Equation (23) ensures that the movement is constrained within the search space limits. It guarantees a balance between exploration and exploitation by utilizing a random term that decreases over time.

Chain Variation

Chain variation involves moving one coot to the average location of two adjacent coots. The process starts by calculating the distance between two coots and moving one of them halfway toward the other [22]. The formula for computing chain variation is:

$$a = 1 - l \times \left(\frac{1}{Iteration}\right) \quad (24)$$

Here, $Iteration$ refers to the second coot. This l variation promotes convergence by averaging the coots' positions.

Position Management Relative to Cluster Leaders

In a cluster, a few coots typically serve as leaders at the forefront. The other coots need to adjust their positions in relation to these cluster leaders and move towards them. Coots that are not leaders should alter their positions based on the dominant leader. Additionally, coots improve their positions based on the leaders' average position. However, utilizing average positions can result in convergence issues. To calculate these variations, the following formula is employed:

$$CootPOS(I) = 0.5 \times (CootPOS(I - 1) + CootPOS(I) \quad (25)$$

Here, $CootPOS(I - 1)$ represents the leader index number, denotes the count of leaders, and signifies the index number of the current coot. The coot adjusts its position based on the leader's index. The coot's future position is calculated based on the selected leader.

$$k = 1 + (I \, mod \, NL) \quad (26)$$

Here, represents an arbitrary number within the interval [-1, 1], signifies the value of π (approximately 3.14), represents a random number within the interval [0, 1], represents the leader's chosen position, and signifies the current coot's position.

$$CootPOS(I) = LeaderPOS(K) + 2 \times r_1 \times \cos(2r\pi) \times (LeaderPOS(K) - CootPOS(I) \quad (27)$$

The comprehensive explanation provided above delves into the intricate mechanics of the COOT optimization algorithm. By strategically integrating this algorithm into the framework of GRU-RNN, the proposed approach enhances the selection of weight parameters, leading to improved accuracy and robustness in the diagnosis of thyroid conditions.

Here is the suggested method's pseudocode:

Algorithm 1: pseudocode of the suggested method

```
initialized COOT's weights and GRU-RNN biases
Initialize P=0.5 and NL (coots).
Randomly choose coot leaders.
        Calculate coot and leader fitness.
        Find the finest leader or coot globally
        While condition is satisfied
          Calculate variables a and b
          If NL > 10 then
            coot can be random vectors
          Else
            coot are random variables
          End
          For each coot from 1 to NL
            Compute variable k
            If k < P then
              Update the position
            Else
              Update the position
            End
          End
          Determine the coots' fitness.
            If coot's fitness is higher than leader's,
             Then assign leader (k) = coot.
              Assign leader (k) to coot.
      End
            For each leader
              If RAND < 0.5 then
                Update position of leader
              Else
          Update the leader's position
          End

            If a leader's fitness is higher
                Keep the optimal settings
            End
          End
        End
```

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Improving the cluster by strategically positioning the leaders

An essential aspect of the COOT optimization algorithm involves steering the cluster toward an optimal location. To achieve this, cluster leaders must adapt their positions in alignment with the desired destination. This section focuses on the mechanism of enhancing the cluster's positioning through the leaders to reach an optimal location [21]. For this purpose, leaders are prompted to update their positions toward the desired goal. This involves revising the positions of leaders to align with the optimal outcome. The formulation aims to identify optimal positions surrounding the current optimal point. Leaders strive to locate these optimal positions by departing from the present ideal position. The formula facilitates the identification of ideal positions while simultaneously avoiding convergence to the current optimum. The process of improving the cluster's positioning from the leaders to the ideal spot mathematically represented as:

$$LeaderPOS(I) = \begin{cases} b \times r3 \times cos(2\pi r) \times (gbest - LeaderPOS(I)) + gbest & r3 < 0.5 \\ b \times r3 \times cos(2\pi r) \times (gbest - LeaderPOS(I)) - gbest & r4 \geq 0.5 \end{cases} \quad (28)$$

In this (28) equation, r represents an arbitrary number within the interval [-1, 1], signifies the value of π (approximately 3.14), $r3$, and $r4$ represents a random variable within the interval [0, 1], and $gbest$ corresponds to the best position ever discovered.

The iterative process of updating the leader's position is further guided by the iteration count. The formula to determine the leader's new position is given b

$$b = 2 - l \times \left(\frac{1}{Iteration}\right) \quad (29)$$

Here, $Iteration$ represents the maximum number of iterations, and l signifies the current iteration. By incorporating the COOT algorithm, the proposed classifier effectively identifies the optimal weighting parameter. Ultimately, this optimized classifier is employed for the accurate identification of thyroid cases from the collected databases.

4. Results and Analysis

This section presents a comprehensive evaluation of the proposed automated thyroid detection system. The validation process was conducted using a standardized database obtained from an open-source system, which is essential for establishing the credibility of the proposed thyroid detection methodology. The database contains three distinct classes and a total of 7201 records. To facilitate network training, 80% of the database was used, while the remaining 20% was reserved for network testing. The performance of the proposed technique was meticulously evaluated using MATLAB using a variety of statistical metrics, including accuracy, precision, recall, specificity, kappa, F1 score, AUC, ROC, and sensitivity. To ensure a fair evaluation of the proposed method, a comparative analysis was conducted against well-established benchmarks such as ANN, SVM, and DBNN. The simulation parameters that form the foundation of the proposed approach are meticulously detailed in Table 3, offering transparency into the method's configuration. Additionally, the validation of the proposed technique was reinforced by applying statistical measures, including the use of a confusion matrix, to enhance the evaluation process. The intricacies of the confusion matrix and its associated metrics are encapsulated in Table 4. This comprehensive performance evaluation sheds light on the pragmatic applicability and potential superiority of the proposed thyroid detection methodology, reinforcing its significance in the domain of medical diagnostics.

Table 3: The simulation tools for parameters

S. No	Variable term	Parameters
1	Find agent quantity	2
2	Lower limit	200
3	The upper limit, measurement	500
4	Maximum iterations	10
5	A dropout rate	100
6	Speed of learning	0.1
7	Parametric regularity	0.001
8	The start of the function	0.01
9	Losing capability	ReLU
10	Size of batch Epoch Neuron count	MSE
11	Variable term	50
12	Epoch	50
13	Number of neurons	32 for the GRU layer

Table 4: Confusion matrix

N=7201		**Predicted**		
		Class-1	**Class -2**	**Class -3**
Actual	**Class- 1**	2995	2	3
	Class -2	3	2798	2
	Class- 3	2	2	1401

Table 5: Feature evaluation utilizing the proposed technique

S. No	Features	Info gain	Gain ratio	Chisq
1.	Taking the hormone thy	0.0128	0.0011	0.638
2.	Query with thyroxine	0.0212	0.0017	6.5128
3.	Query with thyroxine	0.0019	0.0088	7.6544
4.	Taking ant thyroid drugs	0.0914	0.0458	7554.4197

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: 25 July 2023 Revised: 12 September 2023 Accepted: 30 September 2023

5.	Tired	0.0056	0.0843	922.4810
6.	Being pregnant	0.0489	0.0253	2316.9053
7.	Thoracic surgeries	0.0065	0.0634	3248.1219
8.	I131_treatment	0.1293	0.0441	249.9336
9.	Search hyperthyroidism	0.0743	0.1332	1485.5598
10.	testosterone	0.0994	0.0106	385.4197
11.	the third stage	0.0975	0.0855	125.255
12.	TT4	0.3549	0.0488	69.549
13.	T4U	0.1549	0.0688	125.655
14.	FTI	0.0486	0.0155	55.757

The effectiveness of the proposed methodology hinges on its innovative two-level feature extraction technique, which takes into consideration expert recommendations to pinpoint pivotal features. Information gain, gain ratio, and chi-square measures are also used to pick features. These measurements extracted characteristics and selected traits, which are precisely described in Tables 5 and 6. Several robust performance measures are used to

evaluate the suggested strategy. Accuracy, sensitivity, specificity, F measure, kappa, ROC, AUC, precision, and recall are these measurements. By systematically evaluating these measures, the suggested method's capabilities are understood. Advanced feature extraction, expert insights, and multidimensional performance evaluation measures make the suggested thyroid illness categorization strategy promising.

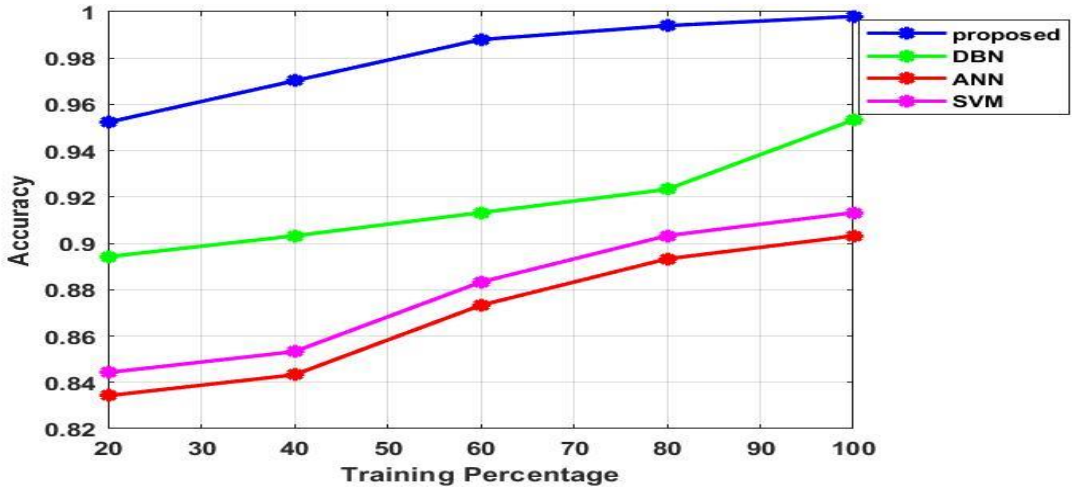

Figure 5: Comparative evaluation of Accuracy

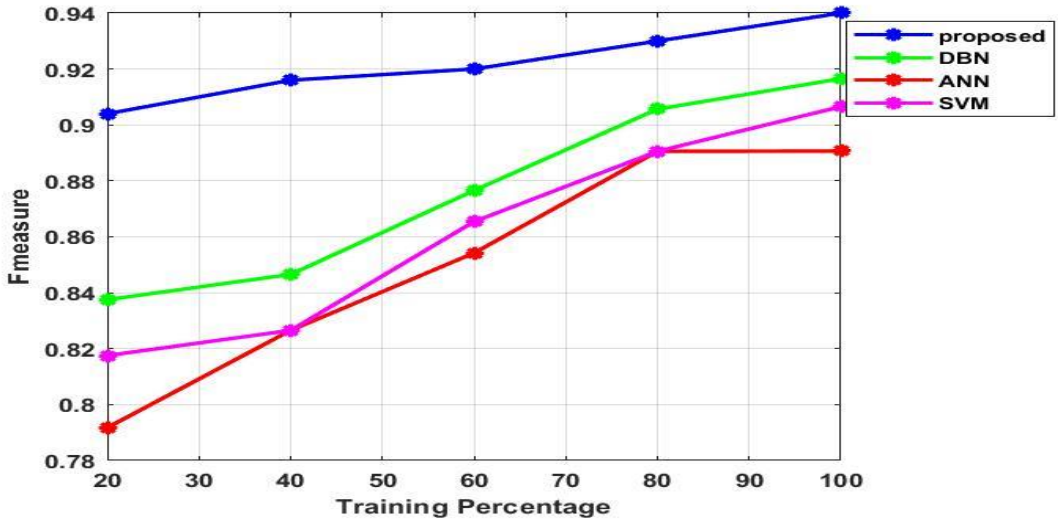

Figure 6: Comparative evaluation of F_ measure

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

To substantiate the efficacy of the suggested model for hypothyroidism identification, the metric of accuracy is employed as a benchmark, as depicted in Figure 5. The performance of the proposed method is juxtaposed against conventional methodologies, specifically DBN, ANN, and SVM. Impressively, the proposed technique attains an accuracy of 0.998. In comparison, DBN, ANN, and SVM yield accuracy scores of 0.95, 0.91, and 0.909 respectively. Evidently, the proposed technique yields the most favorable results, as discerned from Figure 5. Similarly, to validate t the F_measure of the suggested hypothyroidism-detecting method he merits of the proposed thyroid detection model, the F_measure is utilized as an evaluation metric, showcased in Figure 6. The proposed approach is assessed against established methods, namely DBN, ANN, and SVM. Remarkably, the proposed technique garners an F_measure of 0.94. In contrast, DBN, ANN, and SVM achieve F_measure scores of 0.915, 0.91, and 0.89 respectively. This reiterates the superior performance of the proposed technique, as underscored by Figure 6. Furthermore, precision is employed to further affirm the efficacy of the proposed thyroid detection model, presented in Figure 7. A comparative analysis is conducted against customary techniques including DBN, ANN, and SVM. Impressively, the proposed approach yields a precision of 0.93. Contrastingly, DBN, ANN, and SVM exhibit precision scores of 0.928, 0.918, and 0.905 respectively. The discerning eye can deduct from Figure 7 that the proposed technique consistently outperforms its counterparts. This detailed performance evaluation underscores the proposed thyroid detection model's superiority across multiple critical dimensions, establishing its potential for highly accurate and reliable thyroid disease classification.

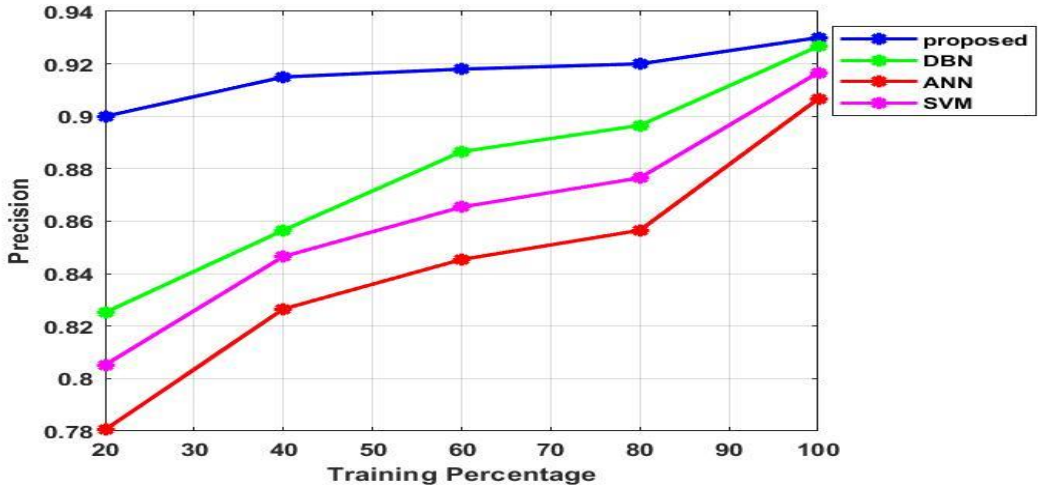

Figure 7: Comparative evaluation of Precision

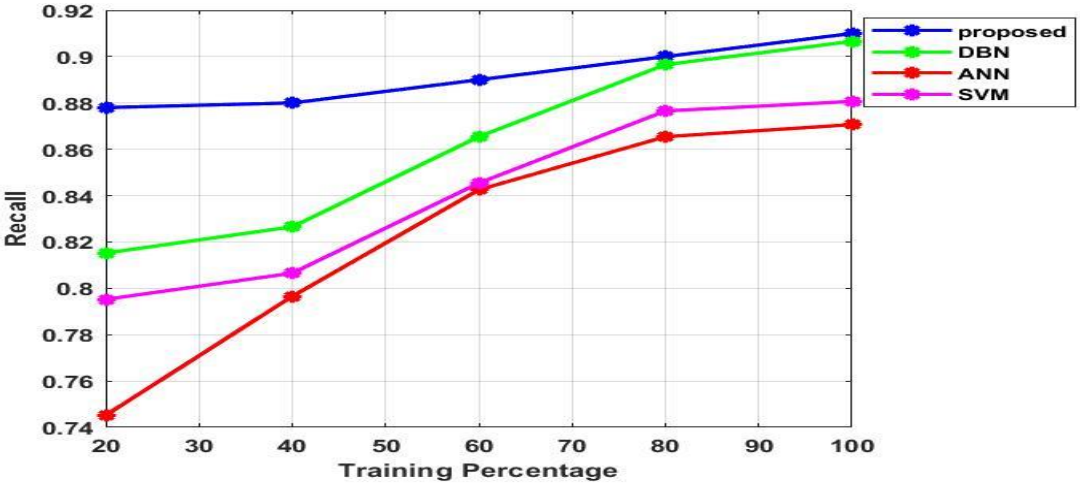

Figure 8: Comparative evaluation of Recall

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

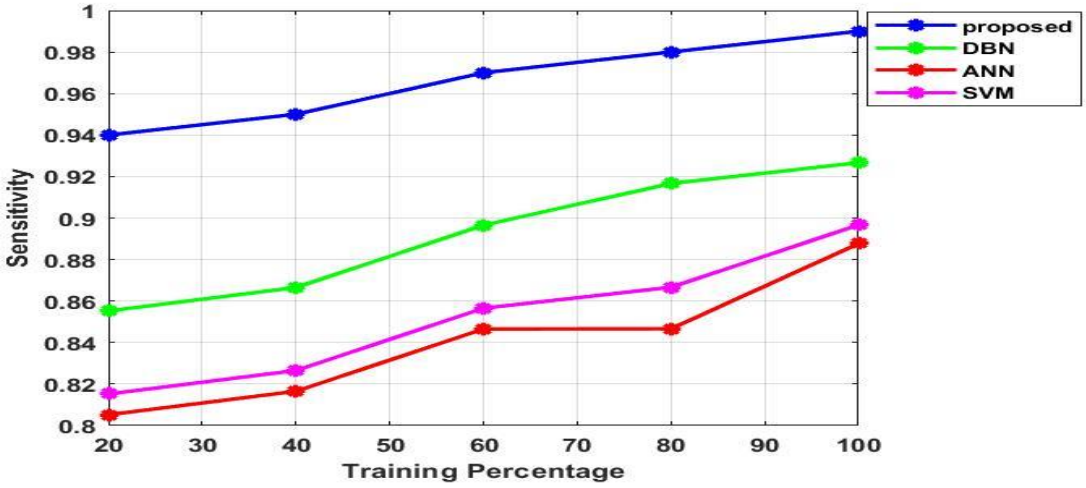

Figure 9: Comparative evaluation of Sensitivity

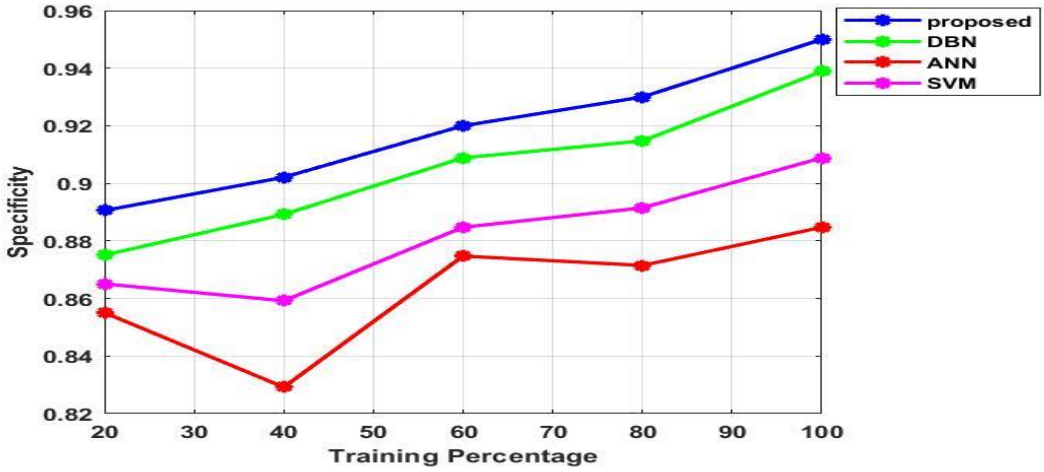

Figure 10: Comparative evaluation of Specificity

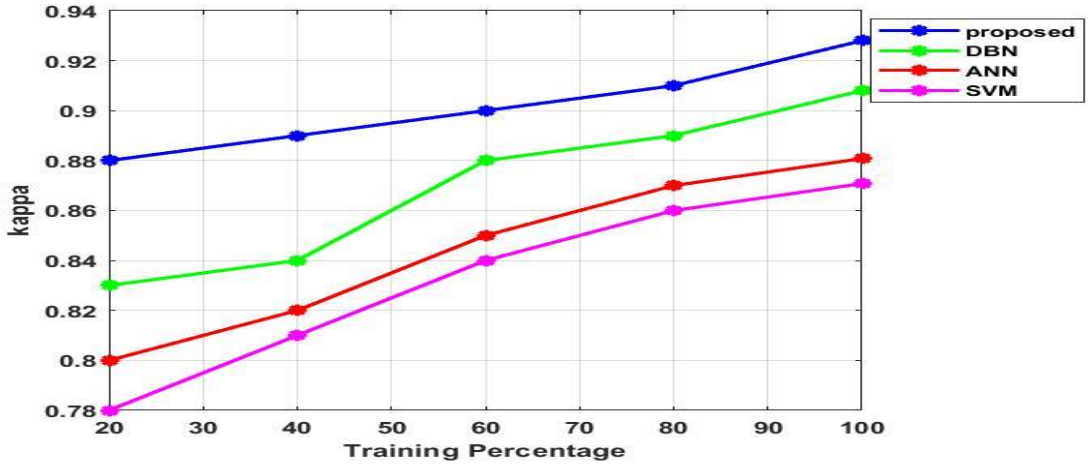

Figure 11: Comparative evaluation of Kappa

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Table 6: Comparative evaluation of different methods

S. No	References	Method	Accuracy	Precision
1	Tahir Alyas *et al.*, [11]	empirical technique	0.948	0.89
2	Hafiz Abbad Ur Rehman *et al.*, [12]	KNN	0.91	0.91
3	Mehdi Hosseinzadeh *et al.*, [13]	multiple multi-layer perceptron (MMLP)	0.7	0.85
4	Rajasekhar Chaganti *et al.*, [14]	deep learning method	0.94	0.92
5	K. Shankar *et al.*, [15]	multi-kernel SVM approach	0.97	0.90
6	-	Proposed approach	0.998	0.93

Table 7: Comparative evaluation different methods

S. No	Method	Accuracy	Precision
1	Whale Optimization Algorithm (WOA)	0.931	0.94
2	Particle swarm optimization (PSO)	0.65	0.71
3	Grey wolf optimization (GWO)	0.85	0.82
4	Black widow optimization (BWO)	0.93	0.90
5	Harmony search algorithm (HSO)	0.96	0.92
6	Proposed approach	0.998	0.93

To substantiate the effectiveness of the proposed thyroid detection model, the evaluation metric of recall is employed, presented in Figure 8. The performance of the proposed method is juxtaposed with conventional techniques such as DBN, ANN, and SVM. Impressively, the proposed technique achieves a recall of 0.91. In comparison, DBN, ANN, and SVM attain recall scores of 0.908, 0.88, and 0.87 respectively. The superiority of the proposed technique is evident from Figure 8. Moreover, sensitivity serves as a pivotal metric to reinforce the merits of the proposed thyroid detection model, illustrated in Figure 9. The proposed approach is compared against established techniques including DBN, ANN, and SVM. Notably, the proposed technique attains a sensitivity of 0.99. In contrast, DBN, ANN, and SVM yield sensitivity scores of 0.918, 0.89, and 0.88 respectively. This underscores the superior performance of the proposed technique, as highlighted by Figure 9. Likewise, specificity is employed to further affirm the proposed thyroid detection model's prowess, displayed in Figure 10. A comparative analysis is conducted against customary methods, specifically DBN, ANN, and SVM. Impressively, the proposed approach secures a specificity of 0.95. Conversely, DBN, ANN, and SVM exhibit specificity scores of 0.94, 0.92, and 0.885 respectively. The evident superiority of the proposed technique is evident from the suggested model for hypothyroidism monitoring is Figure 10. The assessment of the proposed thyroid detection model is further bolstered by the use of the kappa measure, depicted in Figure 11. The proposed method is compared against traditional techniques like DBN, ANN, and SVM. Impressively, the proposed technique achieves a kappa value of 0.93. In comparison, DBN, ANN, and SVM achieve kappa scores of 0.91, 0.918, and 0.889 respectively. The proposed technique's consistent superior performance is visually apparent in Figure 11. This comprehensive evaluation, encompassing diverse performance metrics, solidifies the proposed thyroid detection model's position as an exceptional approach within the realm of thyroid disease classification.

5. Conclusion

This paper introduces the Feature Extraction Based Recurrent Neural Network (FERNN) for computer-aided thyroid disease prediction, highlighting the increasing importance of computer-aided tools in medical diagnostics. The paper presents a method for thyroid disease prediction using a Gated Recurrent Unit within Recurrent Neural Networks, optimized with the COOT Optimization Algorithm.The methodology follows a systematic sequence of steps, beginning with the collection of databases from open-source platforms. A preprocessing protocol is then established to address missing values using the normalization method of min-max. The preprocessed dataset is subjected to feature extraction through the Two-Level Feature Extraction (TLFE) approach. The initial phase (Level 1) employs the method of ranking filter sets features, guided by expert recommendations and reinforced through label-driven validation. In Level 2, features are meticulously evaluated and selected using metrics like information gain, gain ratio, chi-square, and relief. Effective features are identified employing a combined indicator with a fuzzy basis. The process concludes with the implementation of the GRU-RNN model for thyroid disease classification, fine-tuning weight selection through the COOT Optimization Algorithm. Performance evaluation is executed using MATLAB, rigorously assessing metrics such as sensitivity, kappa, F-measure, recall, specificity, accuracy, and precision. The suggested approach is thoroughly benchmarked against industry standard methods, including ANN, SVM, and

DBNN. This comprehensive evaluation underscores the potential of the FERNN model as a cutting-edge solution for automated thyroid detection. The models advanced techniques and rigorous assessment contribute to the enhancement of medical diagnostic capabilities, holding promise for improved thyroid disease classification.

Author Contributions

Nagavali Saka: Conceptualization, investigation, reviewing and editing, investigation, methodology, writing an original draft S. Murali Krishna: research design, data analysis all authors have read and agreed to the published version of the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgements

Authors would like to thank the anonymous reviewers for their useful comments and suggestions

Disclosure Statement

No potential conflict of interest was reported by the authors

References

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

[2] Suresh Kumar, Vidhushavarshini, Satyabhama Balasubramanian, Vijayakumar Ravi, and Ajay Arunachalam.2022. "A hybrid optimization algorithm-based feature selection for thyroid disease classifier with rough type-2 fuzzy support vector machine." *Expert Systems* 39, no. 1 e12811.

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

[4] Chaubey, Gyanendra, Dhananjay Bisen, Siddharth Arjaria, and Vibhash Yadav.2021 "Thyroid disease prediction using machine learning approaches." *National Academy Science Letters* 44, no. 3, 233-238

[5] Li, Wenjun, Siyi Cheng, Kai Qian, Keqiang Yue, and Hao Liu.2021 "Automatic recognition and classification system of thyroid nodules in CT images based on CNN." *Computational Intelligence and Neuroscience.*

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

[7] Yadav, Dhyan Chandra, and Saurabh Pal. 2020."Prediction of thyroid disease using decision tree ensemble method." *Human-Intelligent Systems Integration 2, no. 1, 89-95*.

[8] Ahmad, Waheed, Ayaz Ahmad, Chuncheng Lu, Barkat Ali Khoso, and Lican Huang. 2018."A novel hybrid decision support system for thyroid disease forecasting." *Soft Computing* 22, no. 16, 5377-5383.

[9] Poudel, Prabal, Alfredo Illanes, Elmer JG Ataide, Nazila Esmaeili, Sathish Balakrishnan, and Michael Friebe. 2019."Thyroid ultrasound texture classification using autoregressive features in conjunction with machine learning approaches." *IEEE Access* 7, 79354-79365.

[10] Yadav, Dhyan Chandra, and Saurabh Pal.2020. "Prediction of thyroid disease using decision tree ensemble method." *Human-Intelligent Systems Integration* 2, no. 1, 89-95.

[11] Alyas, Tahir, Muhammad Hamid, Khalid Alissa, Tauqeer Faiz, Nadia Tabassum, and Aqeel Ahmad. 2022."Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach." *BioMed Research International* 2022.

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

[13] Hosseinzadeh, Mehdi, Omed Hassan Ahmed, Marwan Yassin Ghafoor, Fatemeh Safara, Saqib Ali, Bay VO, and Hsiu-Sen Chiang.2021. "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, 3616-3637.

[14] Chaganti, Rajasekhar, 2022. 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, 3914.

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

[16] Patro, S., and Kishore Kumar Sahu.2015."Normalization: A pre-processing stage." *arXiv preprint arXiv: 1503.06462*.

[17] Juneja, Kapil.2022. "Expanded and Filtered Features Based ELM Model for Thyroid Disease Classification." *Wireless Personal Communications* 1-38.

[18] Juneja, Kapil. "Expanded and Filtered Features." *Information Technology* 95, no. 9 (2017): 1996-2005

[19] Alrasheedi, Abdullah, and Abdulaziz Almalaq.2022. "Hybrid Deep Learning Applied on Saudi Smart Grids for Short-Term Load Forecasting." *Mathematics* 10, no. 15, 2666.

[20] Aldallal, Ammar.2022. "Toward Efficient Intrusion Detection System Using Hybrid Deep Learning Approach." *Symmetry* 14, no. 9, 1916

[21] Naruei, Iraj, and Farshid Keynia. 2021."A new optimization method based on COOT bird natural life model." *Expert Systems with Applications* 183, 115352.

[22] Huang, Yihui, Jing Zhang, Wei, Tao Qin, Yuancheng Fan, Xuemei Luo, and Jing Yang.2022 "Research on Coverage Optimization in a WSN Based on an Improved COOT Bird Algorithm." *Sensors* 22, no. 9, 3383.