AVE Trends in Intelligent Health Letters

Inventive Healthcare Monitoring and Harnessing Knowledge for Thyroid Disease Classification Using CTOA-Based TriAG-RDBNet

S. Venkatasubramanian[1,*]

[1]Department of Computer Science and Business Systems, Saranathan College of Engineering, Trichy, Tamil Nadu, India.
veeyes@saranathan.ac.in[1]

*Corresponding author

Abstract: Thyroid disease (TD) patients benefit greatly from early diagnosis and treatment. Ultrasounds and blood tests are two important diagnostic tools for evaluating thyroid function abnormalities. The seriousness of thyroid issues necessitates the use of reliable diagnostic methods. This paper introduces a new approach to TD classification by utilising cutting-edge feature extraction, classification, and picture preprocessing methods. Thyroid image noise is efficiently reduced by the Fast Non-Local Means Algorithm (FNLM), providing a clean input for further analysis. Histogram Features is a strategy for creating a comprehensive representation for disease classification that extracts discriminative features from pre-processed images. To enhance the model's ability to identify subtle patterns that might indicate thyroid issues, propose using a Triple Attention Guided Residual Dense (TriAG-RDBNet) and a BiLSTM-based connection model. The CTOA (Chaotic-Based Tumbleweed Optimisation) Algorithm is used to fine-tune the model's hyperparameters for optimal execution. Dynamic parameter optimisation helps the model achieve the highest possible classification accuracy. Extensive testing on a publicly available dataset demonstrates the efficiency of the proposed method, achieving an astounding 99.22% accuracy. Our integrated approach outperforms state-of-the-art models in comparison tests, suggesting it might be a helpful tool for TD classification.

Keywords: Bidirectional Long Short-Term Memory; Tumbleweed Optimisation; Diagnostic Tools; Thyroid Disease (TD); Histogram Features; Diagnosis and Treatment; Classification Accuracy.

Cite as: S. Venkatasubramanian, "Inventive Healthcare Monitoring and Harnessing Knowledge for Thyroid Disease Classification Using CTOA-Based TriAG-RDBNet," *AVE Trends in Intelligent Health Letters*, vol. 2, no. 2, pp. 72–84, 2025.

Journal Homepage: https://avepubs.com/user/journals/details/ATIHL

Received on: 21/08/2024, **Revised on:** 08/12/2024, **Accepted on:** 24/01/2025, **Published on:** 05/06/2025

DOI: https://doi.org/10.64091/ATIHL.2025.000167

1. Introduction

Thyroid disease (TD) refers to a collection of diseases affecting the thyroid, a small gland located at the centre of the throat, shaped like a butterfly. The hormones secreted by this gland have far-reaching effects on development, metabolism, and other bodily functions [1]. TD can cause a wide variety of symptoms and health issues, and can take many different forms, from hypoactivity to hyperactivity. Hyperthyroidism is a condition in which there is an abnormally high production of thyroid hormones. Hyperthyroidism is a common complication of Graves' disease, an autoimmune condition [2]. Graves' illness is characterised by an excess of hormone production due to a mistaken immune system attack on the thyroid gland. Hyperthyroidism causes fatigue, irritability, rapid heart rate, rapid weight loss, and intolerance to heat. Without therapy, this illness might cause serious complications, including osteoporosis and heart disease [3]. When compared to other kinds of cancer, thyroid carcinoma is an uncommon and deadly illness that can emerge from thyroid nodules. Papillary thyroid carcinoma, a kind of thyroid cancer, has a good prognosis if caught early. Surgical removal, radioactive iodine therapy, and

hormone replacement therapy are all viable alternatives for treatment [4]. Because the thyroid gland plays an essential role in regulating multiple physiological systems, detecting TD is crucial. Hormones produced by the thyroid gland regulate the organ's overall function. Hyperthyroidism and hypothyroidism are two of the many diseases linked to thyroid hormone abnormalities. Physical and emotional health can be severely impacted by TDs, making early identification essential [5].

Thyroid problems can cause a variety of symptoms, including lethargy, irritability, rapid or slow heartbeat, and changes in body weight. Early detection reduces the likelihood of complications and enables prompt medical intervention, thereby preventing disease progression [6]. Both the cardiovascular and reproductive systems are vulnerable to TD's effects [7]. Rapid diagnosis allows doctors to tailor treatment plans, which may include anti-thyroid medicines or hormone imbalance [8]. The thyroid interacts with the endocrine system in complex ways and has far-reaching effects on the body. Screening and early diagnosis programmes for TD have been shown to improve patient outcomes, quality of life, and the burden on the healthcare system [9]. A subfield of Machine learning (ML), Deep learning (DL), has achieved unprecedented success in several areas, including medical diagnosis, image identification, and natural language processing [10]. Because it can automatically recognise patterns and features in large datasets, it is particularly diagnostic. Large collections of thyroid ultrasound images may be used to train DL algorithms for pattern and anomaly detection in the context of TD diagnosis [12]. CNNs, or convolutional neural networks, are a type of DL architecture that has shown exceptional promise in image classification tasks and are well-suited for thyroid imaging analysis [13]; [11].

1.1. Motivation

The pressing need to improve TD detection beyond the scope of current diagnostic methods motivates our work. Due to frequent inaccuracies and inefficiencies, the healthcare industry desperately needs a paradigm shift. Our ultimate objective is to introduce a ground-breaking tool that, by using DL, has the potential to improve the accuracy and timeliness of TD diagnoses. Addressing the challenges posed by the intricacy of thyroid disorders, this unique approach helps move healthcare closer to its ultimate goal of enhancing patient outcomes through technological innovation [31].

1.2. Main Contributions

- **Effective Noise Reduction:** Noise in thyroid pictures is efficiently reduced using the Fast Non-Local Means Algorithm (FNLM), ensuring a clean input for subsequent analysis.
- **Feature Extraction:** The Histogram Features method is practical for preprocessed images, extracting unique features and constructing a holistic representation for accurate illness categorisation.
- **Advanced Model Classification:** It introduces the TriAG-RDBNet, a classical model that helps detect subtle features that may indicate thyroid issues.
- **Hyperparameter Optimisation:** It uses the Algorithm (CTOA) to dynamically optimise model parameters and modify hyperparameters for optimal classification accuracy.

1.3. Organisation of Work

The rest of the research is structured similarly to a shadow: Section 2 describes the related literature, Section 3 provides an impression of the suggested replica, Section 4 presents the study's findings and corroborating investigation, and Section 5 delivers a conclusion and summary [32].

2. Related Works

The purpose of the overview undertaken by Naeem et al. [14] was to compare presentations to improve the health index through a more precise analysis. This resulted in the collection of significant data, which contributed to the formulation of the study's conclusion. Researchers determined whether the model was beneficial by analysing secondary hypothyroid data alongside the study results. Academics have recently presented a wide variety of mining approaches to the healthcare industry for the first time. The potential application of this approach may be limited by the difficulty of establishing appropriate data formats. In this particular study, the SVM machine-learning classifier was able to accurately identify the symptoms of hypothyroidism in 84.72% of the individual patients who were diagnosed. A diagnostic paradigm for thyroid problems was developed by Prathibha et al. [15]. This paradigm utilised DL techniques. A one-of-a-kind CNN-based ResNet architecture was used to identify hypergonadism, hypothyroidism, thyroid nodules, and other conditions. During the proposed research's training phase, dual optimisers were used to enhance both efficiency and precision. A high-level Python library called Keras, part of the TensorFlow framework, was used to develop the deep learning algorithms. When the presentation measures used to determine the type of TD were compared with past research, there was a considerable improvement in the performance measures. Following retraining, the modified ResNet model achieved an accuracy of approximately 97%, an increase from the first ResNet design's 94%.

In addition, a web-based framework was established to determine the type of TD from a scanned image provided by the input system. By increasing the number of pixels in the image's outer regions using Kirsch's edge detector, Shankarlal et al. [16] proposed a method for recognising cancer. This method would be used to examine the image. After the picture had been refined, the Tree Contourlet Transform was applied to it to extract coefficients that corresponded to the refined picture. Following this, features were developed using the updated thyroid image. These features were then used to train and classify a model using the Co-Active Adaptive Neuro segmentation approach. The thyroid picture was then used to divide the cancerous areas. The CNN method was then utilised to assign tumour grades. To determine the likelihood of hypothyroidism and hyperthyroidism, Hossain et al. [17] utilised a range of machine learning techniques. Important qualities necessary for accurate diagnoses of TDs were also identified during the research. Multiple classification models were used in conjunction with both recent and historical data to select the best model for predicting thyroidism. The Random Forest (RF) algorithm produced the highest assessment score across all domains in the sample, while Bayes produced scores below average. Across all performance metrics, the RF approach to feature selection produced the best results. The research conducted by Alnaggar et al. [18] resulted in the development of an enhanced multiclass classification model. This model used XGBoost to assign individuals to groups based on the type of thyroid disease they had.

The most important contributions were improved feature identification accuracy in the dataset and multiclass classification that differentiated among three distinct forms of thyroid disease. Based on the criteria analysed, the exceedingly selective algorithm XGBoost demonstrated the best categorisation performance. After the hyperparameters were optimised, the model achieved an accuracy that was 99% higher than that of the most cutting-edge models. Dhamodaran et al. [19] wanted to determine whether or not it would be possible to classify thyroid datasets into a sum of different categories by employing support vector machines (SVM), kernel neural networks (KNN), and Naive Bayes by contrasting several diverse machine learning tactics, to determine which one provided the most accurate disease prediction. Using the Expert organisation model, it was demonstrated that predictions of future TD instances, in addition to estimates of pretentious rate, were accurate. In terms of accuracy (98.53 per cent) and throughput (98.34 per cent), the complex model came close to meeting the specified performance levels. An investigation by Brindha and Muthukumaravel [20] evaluated the accuracy of two classification algorithms for diagnosing thyroid conditions. For this study, the CNN and Support Vector were evaluated to determine how effectively they could identify hypothyroidism in addition to hyperthyroidism. During training, the model was trained on data from the UCI Machine Learning Repository. It was found that the CNN classifier was more accurate than the SVM classifier, achieving 89% and 87%, respectively.

3. Proposed Method

Figure 1 illustrates the workflow presented in this paper, showing the connections among all procedures.

Figure 1: Block diagram

3.1. Dataset Explanation

In the deep learning era, data is so valuable that many new companies are providing picture annotation services. The datasets used for the proposed model assessment may be available on various platforms [21]-[24], including Kaggle. The proposed system is evaluated using both an existing thyroid dataset from the UCI repository and a brand-new thyroid dataset containing ultrasound images for hypoTDs. The number of datasets used to evaluate the proposed system is shown in Table 1.

Table 1: Dataset compilation

Disease Type	Amount of Images
Thyroid-nodules	147
customary thyroid	89
Thyroiditis	74
Hypothyroid	110
Excited-thyroid	78
Thyroid cancer	99

3.2. FNLM for Preprocessing

The NLM denoising method is an efficient strategy for reducing background noise [25]. In addition, this technique can fix issues with regular denoising distortion of the signal. For this reason, the NLM denoising method is preferable to other noise-reduction algorithms, such as the Gaussian, Wiener, and median filters, which use a sliding window approach over the entire image. Euclidean distance is used as a measure to evaluate the overall geometric composition. Equation (1) provides a meaning of the NLM denoising technique:

$$NL[I](a) = \sum_{N \in I} \omega(a,b)I(b) \quad (1)$$

Wherein the weight $\omega(a,b)$is described as going after within the equation (2):

$$\omega(a,b) = \frac{1}{Z(a)} \sum \quad e^{-\frac{G_\sigma(\tau)\|(M+\tau)-I(n+\tau)\|_2^2}{d^2}} \quad (2)$$

where τthe number of pixels; $G_\sigma(\tau)$is used to describe backdrop pixels; $\| I(M+\tau) - I(b+\tau) \|_2^2$is the intensity difference surrounded by pixels that are strongly influenced by Euclidean coldness measurements; $Z(a)$choose the setting in equation (3):

$$Z(a) = \sum_b e^{-\frac{G_\sigma(\tau)\|(a+\tau)-I(b+\tau)\|_2^2}{d^2}} \quad (3)$$

By extending the NLM denoising method from two, as in the FNLM denoising method, to many computations (a,b). Equation (4) defines the modified (a,b):

$$\omega(a,b) = \frac{1}{Z(a)} H_i(I(a+s) - I(a-s)), \quad (4)$$

Where τ is distinct as $b-a$, s is definite as $a+\tau$, and $H_i$is described as follows in equation (5):

$$H_i(s) = \sum_{q=0}^{s} e^{-\frac{\|(q)-I(q+\tau)\|_2^2}{d^2}} \quad (5)$$

3.3. Feature Extraction from Histogram Features (H)

Histogram features can be used by selecting item relative to the columns and rows [26]; [27]. For feature extraction, this binary object serves as a unique image mask. The characteristics. First-order histogram and statistical characteristics are two other names for them. Equation (6) describes the probability of the P(h) first-order histogram:

$$P(h) = \frac{K(h)}{N} \quad (6)$$

The number N symbolises the total number of images, while the function K(h) shows all possible grayscale values for h. The mean, or average value, determines how light or dark a picture is:

$$\bar{\imath} = \sum_{h=0}^{p-1} hP(h) = \sum_a \sum_b \frac{k(a,b)}{k} \quad (7)$$

In this case, the grayscale value range from 0 to 255 is characterised by q. The resulting values are used to represent the pixel. SD labels the dissimilarity of a photograph. The following equation (8) demonstrates this:

$$\sigma_h = \sqrt{\sum_{h=0}^{P-1} (h-\bar{h})^2 P(h)} \quad (8)$$

The degree of asymmetry can be measured by comparing the value to a mean or median. Negative skewness is defined by Equation (10) and may be found in Equation (9):

$$\text{Skewness} = \frac{1}{\sigma_h^3} \sum_{h=0}^{P-1} (h-\bar{h})^3 P(h) \quad (9)$$

$$-\text{ Sketwness} = \frac{\bar{h} - \text{mode}}{\sigma_h} \quad (10)$$

To refer to the distribution of the grey equation (11) and explain it:

$$\text{Energy(E)} = \sum_{h=0}^{P-1} [P(h)]^2 \quad (11)$$

Entropy is a measure of how chaotic a set of image data is. It is distinct as follows by Equation (12):

$$\text{Entropy}(\Delta S) = -\sum_{i=0}^{q-1} q(i)\log_2[q(i)] \quad (12)$$

3.4. Use of TriAG-RDBNet for Classification

3.4.1. Triple-Attention Mechanism (TAM)

TriAG-RDBNet is a classification model that follows the extraction process to achieve optimal prediction accuracy for TD. The attention device relies on a process of probability weighting, as widely recognised [28]. The process dimensionally hides-layer features by varying the timing of feature computations, so that more informative features receive larger weighting coefficients. To apply the attention approach, i.e., $h_i = h_{i1}, h_{i2}, \ldots, h_{ib}$, to a dynamic weighted sum of features to get a single local characteristic c_i, which is represented as follows in equation (13):

$$c_i = \sum_{j=1}^{b} m_{ij} h_{ij} \quad (13)$$

Wherein the input progression's weight feature is displayed by m_{ij}. Attentional methods will be used to evaluate the relevance. e_{ij} scores: PCA (principal component analysis) reduces the dimensionality of the original thyroid image, yielding more evocative low-dimensional features with less computational overhead. Finally, the weight distribution of characteristics is obtained by using the TAM procedure. The use of a residual dense feature mining is a secondary objective, as is the integration of deep features (DF) via BiLSTM. The symbol denotes the TAM, and the pink bars represent the remaining network's (RN) sub-modules. Represents the process of combining multiple features by summing their residuals:

$$e_{ij} = w_l^T \tanh\left(W_e h_{t-1} + U_e h_{ij} + z_e\right) \quad (14)$$

Where $h_j$characterises pixel's state information using vector features. It serves as the focal point of. last of all, summation functions were utilised for $a_{ij}$and the hidden state$h_j$producing the context vector c_i. anywhere w_l^T, W_e, U_e, z_e are shared parameters that our scheme should learn in each instance step, as an indication of attention$a_{ij}$is computed as follows in equation (15):

$$a_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{b} \exp(e_{ik})} \quad (15)$$

This study proposes a TAM approach with three different concentration subsystems. Traditional attention mechanisms keep half of the feature weights, while average pooling mechanisms use the function to keep local characteristics in the foreground. The third attention method, max-pooling, emphasises local features by giving them more weight than other features. The following equations (16)– (18) can be used to calculate any of the three attention modules:

$$o = \sum_{j=1}^{b} \frac{\exp(e_{ij})}{\sum_{k=1}^{b} \exp(e_{ik})} h_{ij} \quad (16)$$

$$g = \text{AveragePooling}\left\{\sum_{j=1}^{b} \frac{\exp(e_{ij})}{\sum_{k=1}^{b} \exp(e_{ik})} h_{ij}\right\} \quad (17)$$

$$m = \text{MaxPooling}\left\{\sum_{j=1}^{b} \frac{\exp(e_{ij})}{\sum_{k=1}^{b} \exp(e_{ik})} h_{ij}\right\} \quad (18)$$

The Triple-Attention Mechanism assigns unique weights to each of the thyroid's initial attributes, ensuring that unique features are prioritised over duplicates. Using Eq. (19), researchers can calculate the combined output of the TAM features:

$$F_{(o,g,a)} = \text{concatencate}\ (o \oplus g \oplus a) \quad (19)$$

Once the attention mechanism equation has computed and distributed the weights, three weight processes are used to update the features. Where o is the original feature's output, m is the average of all features, and g is the highest feature's output from the pool. The full $F_{(o,g,a)}$ is the TAM result. The symbol works. The RDN may benefit from the enhanced DF produced by attention-based fusion features.

3.4.2. RDN

To build a new RDN to mine and participate in DF, in addition to effectively removing extraneous features. The RN adds a residual procedure to reduce the trait parameters and prevent gradient degradation and vanishing in deep networks. The RN scheming equation (20) is displayed below.

$$Y = W_i\delta(W_{i-1}X_{i-1}) + X \quad (20)$$

Ydepicts the layer's result; Wshow the weights matrix;δrepresents the ReLU's opening function; Xsymbolises the existing layer's input. The number of DF maintained by fusion increases because dense network (DN) a priori knowledge provides substantial information. By applying feature fusion procedures to both the initial features and the features, a double-dense fusion technique is used in this study. More DF can be obtained using the double-dense fusion method, which combines both the initial features and the deep residual features. Initial features are calculated using equations (21) and (22) in DN:

$$X_0 = H_\ell(X_{\ell-1}) + X_{\ell-1} \quad (21)$$

$$X_Y = H_\ell(Y_{\ell-1}) + Y_{\ell-1} \quad (22)$$

$X_Y$characterises the intensive utilisation of the RN output characteristics; X_0characterises the concentrated operation of the earliest features; and $H_\ell$characterises the feature fusion process. Our solution to the disappearance of deep network gradients and the explosion problem involves using the dense RN to remove superfluous features, while performing deeper feature removal and fusion on multidimensional remote images.

3.4.3. Bi-LSTM

To incorporate the RDN-produced DF, this work proposes employing a BiLSTM network. To improve the correlation flanked by DF at numerous locations, BiLSTM devised a feature that incorporates convolutional semainformation across D different scales. The final step is recognition, in addition to classification, which is handled by the Softmax function. To selectively manage information flow while also protecting and regulating the unit's state, the LSTM unit employs three threshold structures. The input gate and output gate are all structures. This is how they are calculated:

$$I_t = \sigma(X_tW_{ui} + H_{t-1}W_{hi} + b_i) \quad (23)$$

$$F_t = \sigma(X_tW_{uf} + H_{t-1}W_{hf} + b_f) \quad (24)$$

$$O_t = \sigma(X_tW_{uo} + H_{t-1}W_{ho} + b_o) \quad (25)$$

WhereinW_{ui}, W_{uf}, W_{uo} and W_{hi}, W_{hf}, W_{ho} are mass and b_i, b_f, b_o are prejudice coefficients. The mechanism of action is the application of the tanh function over the range (-1, 1). This results in the subsequent equation (26) at timestep t:

$$\tilde{C} = \tanh(X_tW_{uc} + H_{t-1}W_{hc} + b_c) \quad (26)$$

Wherein theW_{uc} and W_{hc}There are factors for weight and $b_c$is a gauge of bias. This part of the model mainly uses two DNs to incorporate initial features and deep residual information. When dense blocks are joined together, residual computation for shallow and DF is performed, thereby improving feature quality [29]. The output layer represents the context-semantic properties after DF integration. A Layer represents the LSTM's forward propagation. The Backward Layer represents the LSTM's backward propagation. ReLU symbolises an activation function. Two parameters of LSTMs: the forgetting parameter F_t, that describes what percentage of ancient memory cells $C_{t-1}$to keep, and the learning parameter I_t, which decides how much weight to give to new input via C_t. By using the pointwise multiplication technique, to obtain the following information, equation (27):

$$C_t = F_t \odot C_{t-1} + I_t \odot \tilde{C}_t \quad (27)$$

After that, to may direct how much data travels from the memory cell to the secret location. As the output gate $H_t$as can be shown in Equation (28), $H_t$as:

$$H_t = O_t \odot \tanh(C_t) \tag{28}$$

Upon setting the output gate to 1, all of the memory data is effectively transferred to the predictor factor. When the result is 0, stop processing and leave the data in the cell. A BiLSTM, based on LSTM, integrates input data in both directions. Regarding the outcome at t, the forward LSTM layer in the input sequence incorporates data from both the past and present. The reverse LSTM layer knows both the input series' time and the time t. LST represents the output at time $tM_{forward}$, and the result of the backward LSTM layer at occasion t is represented by $LSTM_{backward}$:

$$\overrightarrow{H_t} = \Phi\left(X_t W_{uh}^{(f)} + \overrightarrow{H_{t-1}} W_{hh}^{(f)} + b_h^{(f)}\right) \tag{29}$$

$$\overleftarrow{H_t} = \Phi\left(X_t W_{uh}^{(n)} + H_{t-1} W_{hh}^{(n)} + b_h^{(n)}\right) \tag{30}$$

Following that, $\overrightarrow{H_t}$besides$H_t \leftarrow$ to get the concealed state $H_t$mentioned to the output layer via equations (29) and (30). And last, the output $O_t$is calculated by the output layer as shown in equation (31):

$$O_t = H_t W_{hq} + b_q \tag{31}$$

In this case, the weight restriction W_{hq}besides bias limitb_q These are the model parameters of the output layer.

3.5. Tuning Hyperparameters with CTOA

3.5.1. TOA

As a next step after categorisation, TOA employs a grouping construction to use more conventional methods when tuning search parameters. That is, each subpopulation of a tumbleweed population contains many search individuals [30]. The Theory of Approximate (TOA) algorithm can benefit from this hierarchical grouping, and subgroups can help prevent local optima from forming. These two iterations represent the evolution and reproduction of a single tumbleweed. There are two phases to a tumbleweed's life cycle: the individual phase and the reproduction phase.

.

3.5.2. Personage Growth-Local Search

When carrying out a search, crash of the environs on the mth person throughout the cycle (v_l^m) is represent by P_l^m in equation (32):

$$P_l^m = \frac{fit(v_l^m)}{sum\left(fit(X^k)\right)+\xi} \tag{32}$$

Where ξ is a digit flanked by 0 and 1, and $X^k$is a matrix in which every element stands for a sole. The higher P_l^m, The seeds x_l^minside this setting. In TOA, construct.$G_i(i = 1,2,3 \ldots, b)$. The top 50% $G_i$reserved to engage in combat with other subpopulations:

$$\text{Factor} = \begin{cases} \frac{s_1*\left(Xbest_k - v_l^m\right)+c_2*\left(Ybest_k - v_l^m\right)+c_3*\left(Ybest_{l+1} - v_l^m\right)}{3}, m = 1 \\ \frac{s_1*\left(Ybest_l - v_l^m\right)+s_2*\left(Ybest_{l-1} - v_l^m\right)}{2}, m = K = M \\ \frac{s_1*\left(Ybest_{l-1} - v_l^m\right)+s_2*\left(Ybest_l - v_l^m\right)+s_3*\left(Ybest_{l+1} - v_l^m\right)}{3}, m = \text{ or else.} \end{cases} \tag{33}$$

Then, the numerical expressions for all individuals in the subset are shown in equation (34):

$$v_{l+1}^m = v_l^m + r_1 * \text{ Factor} \tag{34}$$

where s_1, s_2, s_3 are all chance integers flanked by the principles of 0 and 2. The residual half is not capable of evolution, and the formula is represented by equation (35); low environmental adaptation subpopulations cannot vie, but they can stop intra-evolution:

$$v_{l+1}^{m} = v_{l}^{m} + r_1 * \left(s_{4K} * (\text{Ybest}_{k} - v_{l}^{m}) + s_5 * \left(v_{\text{select},l}^{m} - v_{l}^{m}\right)\right) \quad (35)$$

where s_4, s_5 is an accidental numeral from 0- 1, and r_1 symbolises the influence of surroundings, through repetition.

3.5.3. Individual Reproduction-Global Search

Tumbleweeds only reproduce after they achieve reproductive maturity, which happens during their own reproductive cycle. The formula for the development of this process is (36):

$$v_{l+1}^{m} = \text{Xbest}_{l} + V * \frac{\text{Max\_itertion} - gc}{\text{Max\_itertion}} \quad (36)$$

Where V is the vector for the seed's chance of falling.

3.5.4. CTOA

The population in the TOA is initialised using a Gaussian distribution of random numbers, as described by equation (37).

$$pop = ln + (un - ln) * \text{rand} \quad (37)$$

Where rand is a random matrix with values among 0 and 1, thus, these two TOA processes are linked to the local and global searches conducted by the population. The sequences generated following the TOA were not uniformly distributed, which could affect results across different sequences. This phenomenon will diminish the optimisation algorithm's resilience. Here, to have the option of using a method to determine the residents' random shape. First, a preliminary x_1u`:

$$X = \{u_1, u_2, \ldots, u_i\}, u_i = f(u_{i-1}), i = |pop| \quad (38)$$

Wherein the selected chaotic map is depicted as f(x). In a chaotic sequence, the next generation solution is to feed the preceding generation's solution into the function. Next, using equation (38), pop_chaotic with chaotic features is constructed:

$$pop\_chaotic = ln + (un - ln) * X \quad (39)$$

whereinlb is the smallest value in the space, and besidesub is the greatest value in the space. Similar to the chaotic. The equations (38) and (39). When an algorithm requires an accidental succession, equations (38) and (39) are utilised in conjunction with chaotic maps to substitute the chaotic sequence. As a result, the chaos-based tumbleweed follows:

- **Step 1:** generate the primary dataset arbitrarily.
- **Step 2:** Make a chaotic succession by frequently iterating selected chaotic mapX.
- **Pace 3:** Make an inhabitant of chaos and label it pop chaotic; then, control its boundaries with Equations (38).
- **Pace 4:** Utilising the population from action 3, conclude the person search piece of CTOA.
- **Pace 5:** Terminate the CTOA's global investigation section.
- **Pace 6:** discover a workable key.

The pseudocode for CTOA's total optimisation process is displayed in Procedure 1.

Algorithm 1: Pseudo- CTOAcode

Input: input: ps： variables: max_gen, population size, search dimension, growth cycle, and K, the maximum number of smaller groups. First, divide the tumbleweed population into K smaller groups. Second, use the chaotic map as the beginning point for

6. if grow_iter $< \frac{gc}{2}$ then
7. Calculate P_k^i using Equation (33);
8. Update r_1;
9. if rank(i) $< \frac{pS - g}{2} \mathcal{E}\mathcal{E}$ rand $<$ using P_k^i then
10. x_i^k using Equation (34);
11. else
12. Utilise the roulette wheel to select a certain;
13. Calculate x_i^k using Equation (35);
14. end
15. else
16. Calculate X_(new,k) using Equation (36);
17. end
18. Determine every entity's equal of fitness groupG_k;
19. Update pbest, pbest_val,gbest besides gbest_val;
20. end
21. if grow_iter==gc-1 then
22.Redivide K groups populaces;
23. end
24. End

4. Results and Discussions

4.1. Experimental Setup

It was determined whether the DL models were effective by utilising the WEKA 3.8.6 configuration. There is a data mining program called WEKA, licensed under the GNU General Public License. In addition to its enormous model library, it offers a wide range of capabilities, including data preparation, visualisation, and more.

4.2. Performance Metrics

The success of the proposed work is measured by the level of acceptance it receives. The portion of TD in the data set that was correctly identified indicates the precision. It reads as follows:

$$\text{Accuracy (ACC)} = \frac{\text{No.of correctly expressions}}{\text{Total no.of images}} \times 100 \quad (40)$$

Accuracy, f-measure, recall, and specificity are calculated using the following equations: (41) -(44).

$$\text{precision (PR)} = \frac{TP}{TP+FP} \times 100 \quad (41)$$

$$\text{F1} - \text{score (F1)} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \quad (42)$$

$$\text{Recall (RC)} = \frac{TP}{TP+FN} \times 100 \quad (43)$$

$$\text{Specificity (SP)} = \frac{TN}{TN+FP} \times 100 \quad (44)$$

Table 2: The extraction of features investigation

First Order Histogram likelihood	Image	Entropy	Energy	Mean	Standard Deviation	Skewness
0.81	Image 1	0.22	0.89	120.6	10.3	0.06
0.79	Image 2	0.19	0.76	113.3	8.8	-0.13
0.82	Image 3	0.24	0.93	129.7	12.2	0.08
0.80	Image 4	0.21	0.83	117.2	9.56	-0.07
0.78	Image 5	0.25	0.91	125	11.4	0.03

The outcomes of feature extraction for TD images are shown in Table 2 and Figure 2. First-order histogram probabilities range from 0.78 to 0.82, indicating the likelihood of unique pixel intensities across the five images. The visual complexity indicator, known as entropy, can range from 0.19 to 0.25. Between 0.76 and 0.93, the energy shows that the pixel values are fairly stable.

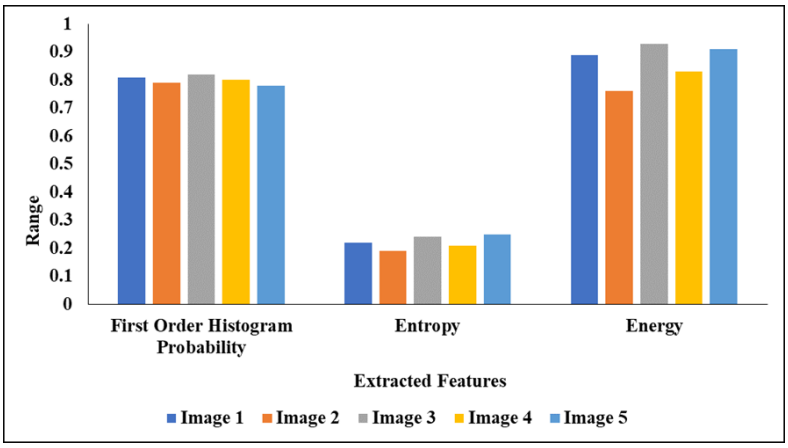

Figure 2: Feature extraction investigation

The central tendency, represented by the mean pixel intensity, varies from 113.3 to 129.7. The standard deviation, a measure of the spread of pixel values, ranges from 8.8 to 12.2. The asymmetry in the intensity distribution across pixels, as measured by skewness, ranges from -0.13 to 0.08. These numerical parameters capture the complication and consistency of TD images, which is important for further research and potential diagnostic applications.

Table 3: Presentation investigation of diverse classes of TD

Classes	PR (%)	ACC (%)	F1 (%)	RC (%)	SP (%)
Thyroid cancer	98.87	99.22	98.76	98.53	99.12
Hyper thyroid	97.94	97.91	97.98	97.67	97.53
Hypo thyroid	97.69	98.47	98.11	98.67	98.22
Thyroid nodules	98.40	98.70	98.26	98.68	98.66
Normal thyroid	98.68	97.92	98.67	98.21	98.79
Thyroiditis	99.31	99.11	99.16	99.24	99.43

Metrics for a model's success in TD classification are presented in Table 3 and Figure 3. The model has a good overall ACC, with accuracies ranging from 97.91% to 99.22% across classes. With PR at 98.87%, RC at 98.53%, and SP at 99.12%, the classical method accomplishes 99.22% ACC cancer. Hyperthyroidism has an ACC of 97.91 percent and corresponding PR, RC, and SP rates of 97.94 percent, 97.67 percent, and 97.5 percent, respectively. Hypothyroidism is diagnosed with an ACC of SP values of 97.69%, and 98.22%, respectively.

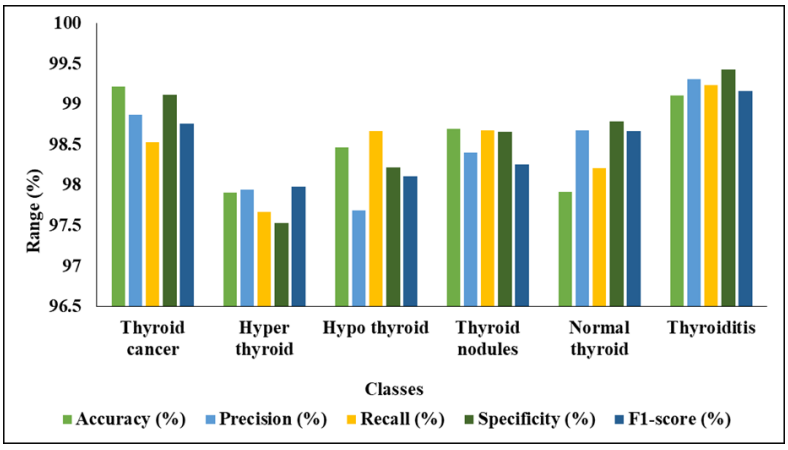

Figure 3: Presentation examination of kinds of classes

The model achieves values of 98.40%, 98.68%, and 98.66%, respectively. A normal thyroid classification has an ACC of 97.92%, a positive predictive value of 98.68%, a negative prognostic value (RC) of 98.21%, and a sensitivity (SP) of 98.79%. The thyroiditis ACC is 99.11 percent, the RC rate is 99.31 percent, the PR rate is 99.24 percent, and the SP rate is 99.43 percent. The F1, taken together, suggests a robust and accurate model for classifying TDs, achieving excellent results across a range of disease types.

Table 4: Classification investigation of a variety of models

Models	PR (%)	ACC (%)	F1 (%)	RC (%)	SP (%)
LSTM	95.34	95.62	95.46	95.23	95.14
Bi-LSTM	96.89	96.73	96.76	96.75	96.66
ResNet	97.46	97.45	97.39	97.33	97.42
DenseNet	98.98	98.21	98.82	98.87	98.75
TriAG-RDBNet	99.18	99.22	99.21	99.21	99.19

Table 4 and Figure 4 summarise the consequences of a classification test involving multiple models.

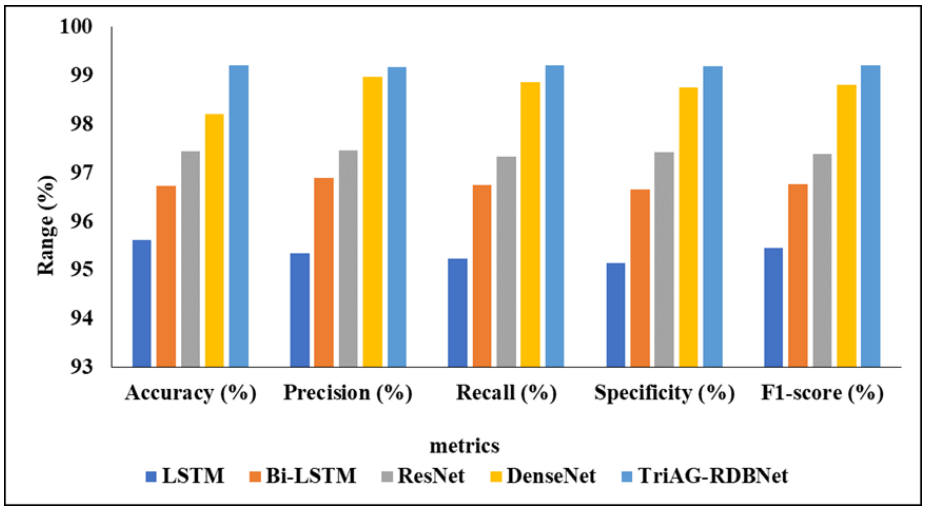

Figure 4: Classification investigation

The LSTM classical achieves an ACC of 95.62% with model has the highest accuracy (ACC: 96.73%) and the highest values of precision (RC: 96.75%), recall (SP: 96.66%), and false discovery rate The DenseNet model achieves the best results across altogether metrics, with ACC of 98.21besides F1 principles of 98.78%, 98.87%, besides 98.82%, respectively. The TriAG-RDBNet representation results in a competition ACC of 99.22% and excellent PR, RC, and SP, along with F1 scores of 98.18%, 99.19%, and 99.22%, respectively. After tuning its hyperparameters with TOA, the proposed TriAG-RDBNet representation proves to be the most precise classifier among the evaluated models.

5. Conclusion

Our findings underscore the imperative for accurate diagnostic techniques and stress the significance of early detection of thyroid diseases (TDs) to enhance patient outcomes. This study introduces a precise, methodical methodology for categorising TDs via a meticulously crafted feature-extraction and image-classification framework. The Fast Non-Local Means (FNLM) technique is used to remove noise from thyroid images, ensuring the input data is of the highest quality for later analysis. This preprocessing technique makes structural patterns clearer and helps extract useful features from the data. After the noise is reduced, the Histogram Features technique is applied to the preprocessed images. This method enables the identification of discriminative statistical traits, providing a strong basis for accurate classification of thyroid problems. The proposed TriAG-RDBNet model is added to the pipeline to improve the diagnostic system further. This model improves diagnostic accuracy by detecting subtle and complex patterns that may indicate the presence of different TDs. This makes the system more sensitive and reliable. The Classification-Tuned Optimisation Algorithm (CTOA) is also used to adjust the model's internal settings dynamically. This adaptive method ensures the network maintains strong generalisation and consistently delivers high classification performance across a wide range of image samples. Tests done on a publicly available dataset show that the proposed method works well. The system's 99.22% accuracy is very impressive, suggesting it could be a reliable tool for

diagnosing problems. Though these results are good, they could be much better if the dataset included a wider range of demographic groups and a greater variety of thyroid problems. This kind of growth would strengthen the model, make it more useful in real-world clinical settings, and, ultimately, help doctors diagnose thyroid disorders more accurately and comprehensively.

Acknowledgement: The author expresses sincere gratitude to all who provided guidance and support throughout this work. Their encouragement and valuable insights greatly contributed to the successful completion of the study.

Data Availability Statement: This study uses data on inventive healthcare monitoring and thyroid disease classification, trained on the CTOA-based TriAG-RDBNet model. All datasets used are integral to the analytical evaluation presented.

Funding Statement: This research work is not funded by any agency or organization.

Conflicts of Interest Statement: The author claims no conflict of interest.

Ethics and Consent Statement: Ethical approval was obtained, and informed consent was secured from the organization and all participants. All procedures adhered to established ethical guidelines and protocols.

References

1. L. Aversano, M. L. Bernardi, M. Cimitile, A. Maiellaro, and R. Pecori, "A systematic review on artificial intelligence techniques for detecting thyroid diseases," *PeerJ Computer Science,* vol. 9, no. 6, p. e1394, 2023.
2. H. H. Kumar, "A novel approach of SVM based classification on thyroid disease stage detection," in Proc. *3rd Int. Conf. Smart Syst. Inventive Technol. (ICSSIT),* Tirunelveli, India, 2020.
3. H. Abbad Ur Rehman, C. Y. Lin, Z. Mushtaq, and S. F. Su, "Performance analysis of machine learning algorithms for thyroid disease," *Arab. J. Sci. Eng.,* vol. 46, no. 1, pp. 9437–9449, 2021.
4. R. R. Sankar, C. S. Mahulikar, and V. Viswanatha, "Thyroid disease detection using machine learning approach," *J. Xi'an Univ. Archit. Technol.,* vol. 15, no. 7, pp. 327–334, 2023.
5. X. Luo, R. Wang, C. Lv, G. Chen, J. You, and F. Yu, "Detection of selenocysteine with a ratiometric near-infrared fluorescent probe in cells and in mice thyroid diseases model," *Anal. Chem.,* vol. 92, no. 1, pp. 1589–1597, 2019.
6. L. Agilandeeswari, I. Khatri, J. Advani, and S. M. Nihal, "An efficient thyroid disease detection using voting based ensemble classifier," in Proc. *Int. Conf. Intell. Syst. Des. Appl., Springer Int. Publishing,* Cham, Switzerland, 2021.
7. K. Dharmarajan, K. Balasree, A. S. Arunachalam, and K. Abirmai, "Thyroid disease classification using decision tree and SVM," *Indian J. Public Health Res. Dev.,* vol. 11, no. 3, pp. 224–229, 2020.
8. X. Zhang, V. C. Lee, J. Rong, J. C. Lee, J. Song, and F. Liu, "A multi-channel deep convolutional neural network for multi-classifying thyroid diseases," *Comput. Biol. Med.,* vol. 148, no. 9, p. 105961, 2022.
9. X. Zhang, V. C. Lee, J. Rong, J. C. Lee, and F. Liu, "Deep convolutional neural networks in thyroid disease detection: a multi-classification comparison by ultrasonography and computed tomography," *Comput. Methods Programs Biomed.,* vol. 220, no. 6, p. 106823, 2022.
10. L. Ma, C. Ma, Y. Liu, X. Wang, and W. Xie, "Diagnosis of thyroid diseases using SPECT images based on convolutional neural network," *J. Med. Imaging Health Inform.,* vol. 8, no. 8, pp. 1684–1689, 2018.
11. L. Ma, C. Ma, Y. Liu, and X. Wang, "Thyroid diagnosis from SPECT images using convolutional neural network with optimization," *Comput. Intell. Neurosci.,* vol. 2019, no. 1, pp. 1–11, 2019.
12. R. B. Namdeo and G. V. Janardan, "Thyroid disorder diagnosis by optimal convolutional neuron based CNN architecture," *J. Exp. Theor. Artif. Intell.,* vol. 34, no. 5, pp. 871–890, 2022.
13. S. Y. Ko, J. H. Lee, J. H. Yoon, H. Na, E. Hong, K. Han, I. Jung, and E. K. Kim, "Deep convolutional neural network for the diagnosis of thyroid nodules on ultrasound," *Head Neck,* vol. 41, no. 4, pp. 885–891, 2019.
14. A. B. Naeem, B. Senapati, A. S. Chauhan, M. Makhija, A. Singh, M. Gupta, P. K. Tiwari, and W. M. Abdel-Rehim, "Hypothyroidism disease diagnosis by using machine learning algorithms," *Int. J. Intell. Syst. Appl. Eng.,* vol. 11, no. 3, pp. 368–373, 2023.
15. S. Prathibha, D. Dahiya, C. R. Robin, C. V. Nishkala, and S. Swedha, "A novel technique for detecting various thyroid diseases using deep learning," *Intell. Autom. Soft Comput.,* vol. 35, no. 1, pp. 199–214, 2023.
16. B. Shankarlal, P. D. Sathya, and V. P. Sakthivel, "Computer-aided detection and diagnosis of thyroid nodules using machine and deep learning classification algorithms," *IETE J. Res.,* vol. 69, no. 2, pp. 995–1006, 2023.
17. M. B. Hossain, A. Shama, A. Adhikary, A. D. Raha, K. M. A. Uddin, M. A. Hossain, I. Islam, S. A. Murad, M. S. Munir, and A. K. Bairagi, "An explainable artificial intelligence framework for the predictive analysis of hypo and hyper thyroidism using machine learning algorithms," *Hum. -Centric Intell. Syst.,* vol. 3, no. 6, pp. 1–21, 2023.
18. M. Alnaggar, M. Handosa, T. Medhat, and M. Z. Rashad, "Thyroid disease multiclass classification based on

optimized gradient boosting model," *Egypt. J. Artif. Intell.*, vol. 2, no. 1, pp. 1–14, 2023.
19. S. Dhamodaran, B. B. Shankar, B. Balachander, D. Saravanan, and D. S. Kharate, "Estimation of thyroid by means of machine learning and feature selection methods," in *Artificial Intelligence for Smart Healthcare, Springer Int. Publishing*, Cham, Switzerland, 2023.
20. V. Brindha and A. Muthukumaravel, "Efficient method for predicting thyroid disease classification using convolutional neural network with support vector machine," in *Computational Intelligence for Clinical Diagnosis, Springer Int. Publishing*, Cham, Switzerland, 2023.
21. Official Dataset, "Thyroid hyper dataset," *Kaggle*, 2020. Available: https://www.kaggle.com/officialdataset/thyroid-hyper [Accessed by 18/06/2024].
22. Official Dataset, "Thyroid nodule dataset," *Kaggle*, 2020. Available: https://www.kaggle.com/officialdataset/thyroid-nodule [Accessed by 28/06/2024].
23. Official Dataset, "Thyroid-ditis dataset," *Kaggle*, 2020. Available: https://www.kaggle.com/officialdataset/thyroid-ditis [Accessed by 22/06/2024].
24. Official Dataset, "Thyroid cancer dataset," *Kaggle*, 2020. Available: https://www.kaggle.com/officialdataset/thyroid-cancer [Accessed by 11/06/2024].
25. B. G. Kim, S. H. Kang, C. R. Park, H. W. Jeong, and Y. Lee, "Noise level and similarity analysis for computed tomographic thoracic image with fast non-local means denoising algorithm," *Appl. Sci.*, vol. 10, no. 21, p. 7455, 2020.
26. S. Baswaraju, V. U. Maheswari, K. K. Chennam, A. Thirumalraj, M. P. Kantipudi, and R. Aluvalu, "Future food production prediction using AROA based hybrid deep learning model in agri-sector," *Hum. -Centric Intell. Syst.*, vol. 3, no. 10, pp. 1–16, 2023.
27. A. Ali, S. Qadri, W. K. Mashwani, W. Kumam, P. Kumam, S. Naeem, A. Goktas, F. Jamal, C. Chesneau, S. Anam, and M. Sulaiman, "Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image," *Entropy*, vol. 22, no. 5, p. 567, 2020.
28. W. Cai, B. Liu, Z. Wei, M. Li, and J. Kan, "TARDB-Net: triple-attention guided residual dense and BiLSTM networks for hyperspectral image classification," *Multimedia Tools Appl.*, vol. 80, no. 1, pp. 11291–11312, 2021.
29. A. Thirumalraj, V. Asha, and B. P. Kavin, "An improved Hunter-Prey optimizer-based DenseNet model for classification of hyper-spectral images," in AI and IoT-Based Technologies for Precision Medicine, *IGI Global*, Hershey, Pennsylvania, United States of America, 2023.
30. T. Y. Wu, A. Shao, and J. S. Pan, "CTOA: toward a chaotic-based tumbleweed optimization algorithm," *Mathematics*, vol. 11, no. 10, p. 2339, 2023.
31. A. Thirumalraj and T. Rajesh, "An improved ARO model for task offloading in vehicular cloud computing in VANET," *Research Square*, 2023. Available: https://assets-eu.researchsquare.com/files/rs-3291507/v1_covered_2b8de597-e0d0-42be-b5d7-f6de74e7356f.pdf [Accessed by 09/06/2024].
32. A. Khang, "AI and IoT-Based Technologies for Precision Medicine," *IGI Global*, Hershey, Pennsylvania, United States of America, 2023.