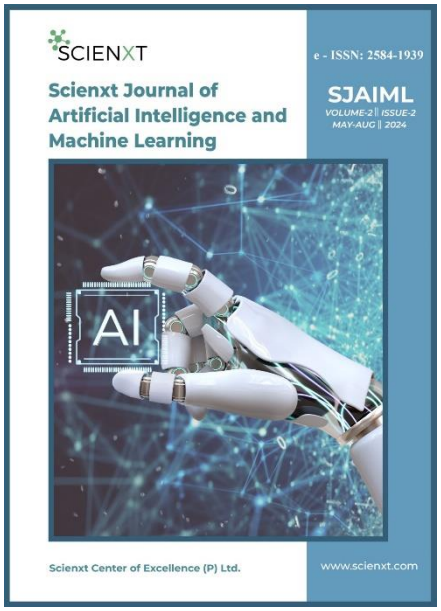


**Scienxt Journal of Artificial Intelligence and Machine Learning**
**Volume-2 || Issue-2 || May-Aug || Year-2024 || pp. 1-22**

# *"A review of methods for thyroid disease detection: examining hybrid meta-heuristic and lstm models in the literature"*

**Mr. Prajith S. A**
Student, Department of ECE, DSATM, Bengaluru
**Dr. Mallikarjun P. Y**
Professor & HOD, Department of ECE, DSATM, Bengaluru

**Corresponding Author: Mr. Prajith S. A*
*Email: prajith2908@gmail.com*

## Abstract:

Clinical research faces major challenges due to thyroid disease management and metabolic control complexity. Hypothyroidism and hyperthyroidism are two common thyroid disorders that are metabolically affected by the secretion of thyroid hormones. It is important to use data cleansing techniques to analyze the random data to assess patient risk accurately. Deep neural networks (DNNs) are an important and effective tool for predicting thyroid diseases, surpassing manual diagnostic methods in knowledge and time required. In this study, we present a new method for thyroid diagnosis and disease prediction, including long-term and short-term memory-based convolutional with a special architecture for disease detection. Let us include the neural network (LSTM-CNN). Feature selection incorporates bias field correction and uses a hybrid insufficiency technique, combining Black Widow optimization with a mayfly optimization approach (HBWO-MOA) to obtain disease classification using LSTM and Vgg-19 architectures in deep learning (DL). Framework in the. Using ultrasound images from the DDTI dataset, our method demonstrates skill in predicting and classifying thyroid diseases. The comparative analysis shows that the proposed Vgg-19-LSTM method, using metrics such as accuracy, sensitivity, precision, recall, and F1-score, outperforms current methods such as AlexNet-LSTM, ResNet-LSTM, and Vgg16-LSTM.

## Keywords:

Classification, HMOA-BWO, LSTM, pre-processing, segmentation, Vgg-19.

## 1. Introduction:

The healthcare industry is using advances in computational biology to collect patient data indicative of medical diseases. There are various methods of early detection. Intelligent applications for disease analysis i.e., medical technology information are not easy to store in the required data types. [1], [2]. However, in recent days, there has been a technology called machine learning (ML) optimization, which greatly helps to predict and solve nonlinear as well as serious problems. Emphasis is placed on the options available in any diagnostic approach, which makes it easier to categorize healthy individuals as much as possible from as much information as possible rather than a healthy individual exposed to unnecessary treatment due to misdiagnosis. Consequently, the accurate prognosis of any thyroid disease is a major concern [3], [4], [5], and [6].

The endocrine gland in the neck is also known as the thyroid gland. It grows under Adam's apple tree below. Identification and Classification of Thyroids Using the DNN part of the human throat helps with thyroid secretion hormone, which affects protein synthesis as well as metabolic rate [7]. The thyroid hormones in the body regulate the metabolism in a few ways, including heart rate and caloric expenditure. The thyroid gland's production contributes to the regulation of the body's metabolism. The thyroid glands actively secrete the thyroid hormones levothyroxine (T4) and triiodothyronine (T3). These hormones are crucial for controlling body temperature both during fabrication and the supervised construction process. T4 sometimes referred to as the two active hormones that the thyroid glands typically generate, thyroxin and T3. Across the body, these hormones play important roles in protein control, temperature regulation, energy storage, and transportation. Iodine is thought to be a crucial component of thyroid glands and its deficiency in certain, highly prevalent cases for these two thyroid hormones, T3 and T4 [8,], [9], [10], and [11].

Thyroid hormone deficiency contributes to hypothyroidism, among other things. Thyroid hormones contribute to hyperthyroidism. Thyroid dysfunction and hyperthyroidism are caused by various diseases. The remedy comes in various forms. Thyroid surgery requires patients to receive iodine, enzyme lac, ionizing radiation, and continuous thyroid dysfunction- producing thyroid hormones [12].

The diagnosis of thyroid disease is now made using ultrasound image technology. The patient's thyroid ultrasound standard plane (TUSP) was collected using an ultrasound device. Clinical examination, then the transverse plane of the thyroid Isthmus (TPTI) is used to modify the TUSP image. Gesture, the downside of the transverse plane of the right Partial thyroid resection

(DTPRT), advanced thyroid resection the isthmus (LPTI), the surface of the transverse plane of the right thyroid segment (UTPRT) is in the transverse plane. Left thyroid (LPLT), the large transverse center Left plane of the thyroid (MTPLT), and the middle transverse plane of the right thyroid (MTPRT), and the physician examines the ten groups afterward. TUSP Figures [13], [14]. Generally, the deep learning (DL) method is divided into two parts: Initially, the method, i.e., the Deep Neural Network (DNN), is used to train the model images and extract the deepest one. Parts, and finally, the trained neural network is used to classify or recognize sample images [15].

This study applies the CNN-LSTM method to predict and classify thyroid diseases. The Vgg-19 base network is used in the CNN model to classify diseases. A Vgg-19-based LSTM has been proposed to reduce computational complexity.

Moreover, it improves the accuracy of the LSTM. To overcome these challenges, a new lightweight LSTM model based on Vgg-19 is presented to provide efficient results in time-series data processing and sequential ultrasound imaging, many of which are found.

This paper is organized as follows: Section IV describes the proposed research methodology; Section V. Section V: Parameter optimization, experimental design, Performance descriptions, research results, and comparisons as well. Additionally, it includes a brief discussion of the diagnosis. Section 6 presents conclusions and future work technical knowledge is provided.

## 2. Related work:

In 2019, Liu et al. [16] also presented automatic node classification and ultrasound detection-specific projects, i.e. drawings using DL-based technology of CAD. The CAD system had two modes that demonstrated the use of a multi-scale region search network to detect nodules & characterize pyramid dal at different scale features. Prior knowledge of the actual node rules was also estimated using magnitude distributions and shape distributions. Thyroid differentiation was a nodule simulation effective.

In 2019, Ma et al. [17] performed thyroid diagnosis using SPECT images using computer-aided CNN optimization. The CNN DenseNet system was trained and developed. SPECT images were accurate in detecting thyroid diseases & its performance was superior to other conventional techniques.

In 2020, Song et al. [18] provided the throne- reducing CNN-based hybrid feature crops to establish a position region of the same smooth image placement and local has two datasets, the

proposed method performed & achieved 93.24 % accuracy, 97.18 % recall, and 95.17 % F1-measurement.

In 2020, Moussa et al. [19] proposed a CNN technology based on ResNet-50 from a DL fine-tuning perspective accuracy for classifying thyroid nodules in ultrasound images. Images from 814 ultrasounds were used for the analysis according to the Vgg-19 method to increase the classification efficiency significantly.

In 2020, Abdolali et al. [20] developed a mask- based multi-service model to integrate the DL method with basic, network so-parameters optimized with special DNN networks to improve detection without refinement-based complicated post-processing steps and regularization of function loss was provided by Off R-CNN Technology. The performance of the loss algorithm, the separation of the prioritized prediction methods the character of the character. A validation method was developed for 821 ultrasound frames from 20 patients. This presented technology detected thyroid glands. The detection of thyroid glands was superior compared to other techniques.

In 2020, Shankarlal et al. [21] classifier used the detection of Kirsch's edge to generate thyroid images in which the pixels of the edge region were enhanced by Dual Tree Contourlet transform (DTCT) to obtain coefficients who used the Co-Active Adaptive Neuro Expert System (CANFES) features to create a modified image of the thyroid. The abnormal thyroid image is then classified by morphologic features to classify the tumor sites. Finally, the segmented tumor segment was analyzed using the CNN method to analyze the mild, severe, and non-malignant features.

In 2021, Hosseinzadeh et al. [22] obtained thyroid diseases from translation reports using the ANN network to enhance the evaluation efficiency of the IoMT system to avoid overfitting in the training set and to enhance the generalization neural network features. The potential was announced IoMT systems compared the number of class variables in MMLP to page width, increasing the accuracy by 4.6% with a final accuracy of 99%. In 2021, Namdeo et al. [23] introduced a new method for thyroid detection that includes two steps, i.e., feature extraction and classification. Initially, principal component analysis (PCA) and neighborhood-based gradient features were used to extract & organize data features. A classification system was also established. Finally, the features of the given method, i.e. WF-CS, compared with PSO, FFA, CS, and ABC, and thus proved superior for the detection of thyroid presence.

In 2021, Wan et al. [24] introduced the exception Dynamic CEUS for imaging diagnosis is the thyroid nodules network of hierarchical temporal attention (HiTAN) which enhances the

classification of hierarchical nodules and detects dynamic features in deep profile. Gated recurrent units (GRUs) were modeled to investigate problem dependence. Simulation analysis improved the performance of the systematic diagnosis tool while also studying dynamic models.

In 2021, Chu et al. [25] presented a web-based signal-guided deep ultrasound model for the thyroid gland. U-Net presented a study on the thyroid to examine the characteristics of computerized technology for thyroid nodule removal to save time the accuracy of thyroid classification was 3% higher than any other convolutional network.

### 2.1. Review:

The methods described for the prediction of thyroid disease are used successfully. In [18], hybrid deep learning is used for thyroid disease detection and classification schemes.

In the detection and classification scheme presented in [19], the state of the diseases is not included in the given dataset. Therefore, this model is not suitable for complex models. In [20], DNN was used for the thyroid disease classification system, but it did not improve the overall quality of the criteria. Additionally, Edge does not apply to the cloud computing model. In [17], a convolution neural network method is used to classify thyroid disease [26]. Because the method presented in [27] for minimizing the maximum possible loss function does not give effective results.

## 3. Scope and methodology:

Thyroid illness classification is essential for assessing the condition and choosing a course of treatment [28] based on their classes. Different imaging modalities are utilized to identify thyroid illness [29]. MRI is widely used, nevertheless, due to its better image quality and lack of reliance on ionizing radiation [30]. Deep learning [31] is an area of machine learning that has demonstrated noteworthy performance lately, particularly in problems related to segmentation and classification [32].

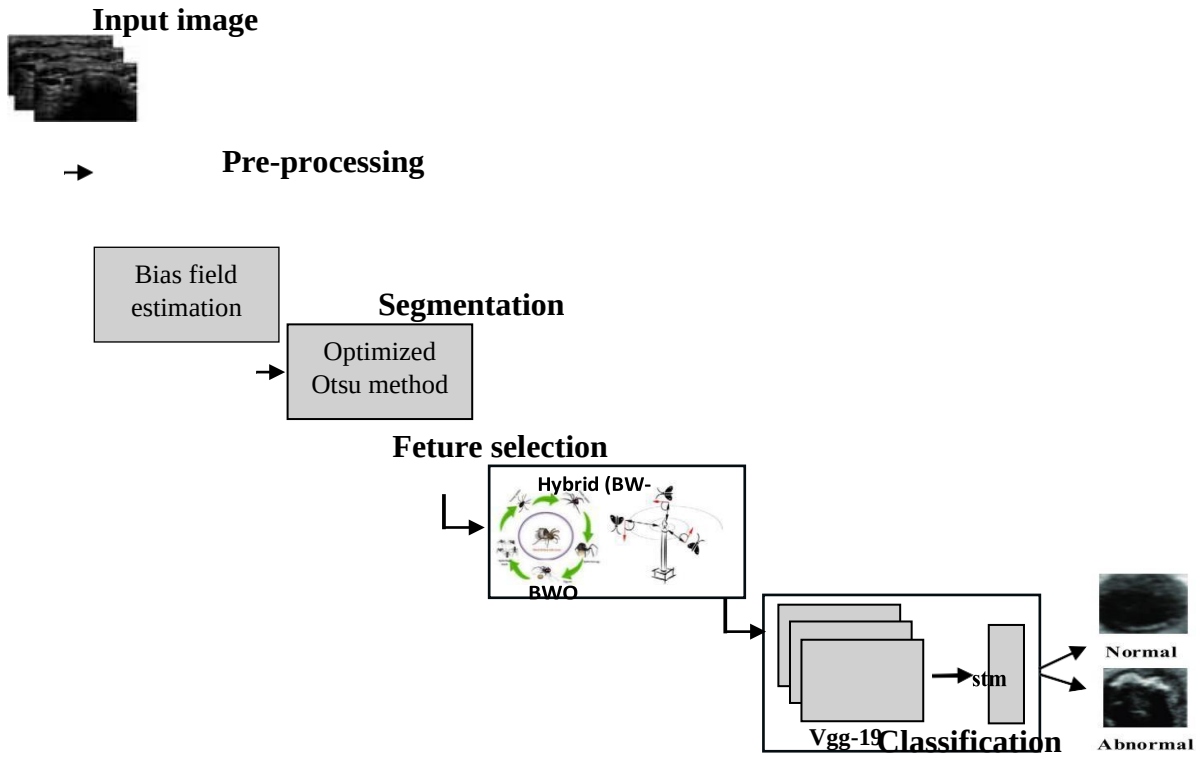


***Figure. 1: Overall flow of the proposed architecture***

Pre-processing, segmentation, feature selection, and classification are the four thyroid disease prediction and classification methods used in this work Thyroid ultrasound images are first acquired, then the preprocessing step using bias field estimation and correction method [33] and done. After the previous steps, the segmentation of the optimized Otsu method is used to separate the image from the MRI image. Once the image is classified, the classification results are fed into the feature selection process, which combines the Moth Flame Optimization Algorithm and Black Widow Optimization Algorithm to form a hybrid meta-heuristic approach.

## 3.1. Pre-processing:

Ultrasound images are affected by coil oscillations in the magnetic field in real-time applications, bias field estimation and correction are used to address this issue, which leads to intensity inhomogeneities and partial volume effects Known as multiplicative module of the image and therefore this paper uses the bias field

## 3.2. Thyroid segmentation image technique:

Here is an image of the thyroid with layers of I(X, Y) gray 0,1…ρ−1. An image interpolation that I can define as F={f0,f1.....fρ−1 }, where f0,f1.....fρ−1 represents the frequency of each image a is inserted I.

This classification method is an extension of the Otsu method. The method of segmentation

Represent the same noise and bias, respectively. The simulation model as listed in Eq. (1)

$P_t = b_{field} X_0 + n$ (1)

Then, the Gaussian filter G(x, y) with the size of the kernel as 3 × 3 is employed for image smoothing.

$$Is(G(x,y)) = \frac{1}{2\pi\sigma 2} e^{-\frac{x2+y2}{2\sigma 2}} \quad (2)$$

O represents the standard deviation.

### 3.3. Utilizing optimized otsu's approach:

Here is an image of the thyroid with layers of I(X, Y) gray 0,1…ρ−1. An image interpolation that I can define as F={f0,f1.....fρ−1 }, where f0,f1.....fρ−1 represents the frequency of each image a is inserted I.

This classification method is an extension of the Otsu method. The method of segmentation

Where, $N = \sum_{i=0}^{\rho} f_i$ the total number of thyroid images in pixels

The probability of $i^{th}$ greyscale level is calculated as

$Probi = \frac{fi}{N}$; Probi>0, $\sum Probi = 1$ (3)

of the thyroid image divides T + 1 into segments Sk = S0, S1. Sk.

Where, from the boundary set T T = {t [r], t [r+ 1], . . . . . ...t[r+1]−1}; Each set is segmented from grayscale level T onwards. Potential loads and intermediate loads can be calculated as follows by dividing η and S at each grayscale level.

$Wsk = \sum i \epsilon Sk Probi$ (4)

$$\eta = \frac{\sum i \epsilon Sk Probi.max\{i, g * log(i)\}; g \epsilon GC}{wsk} \quad (5)$$

The weighted grayscale intensity between components of the entire image is given by μW1 and the inner square variance $\sigma^2_{VS}$

$\mu W1 = \sum^{\rho=\rho} * Probi$ (6)

$\sigma 2 = \sum \rho$ ws . η2 − μ2 (7)

VS $i$=0 k W1

## 3.4. Thyroid image feature selection using hybrid:

Increase diversity to prevent premature meetings and speed up meeting speeds. As discussed

earlier, the proposed model applies to Vgg-19-LSTM with BWO and MFO. Here, the section mainly discusses the hybrid Meta heuristic algorithm with the hybridization of the Mayfly and Black Widow Optimization

## 3.5. Initial population:

In the Black Widow Optimization (BWO) algorithm, the population is initially killed randomly, requiring two different populations, such as male and female. Based on this beginning, you produce children for future generations. Healthy pricing is important in this process. The fitness function is denoted as f in the widow. The following is an estimate of $x_N$, is the black widow spider's population, N indicates population size, d denotes the number of decision variables of the problem, lb is the population lower bound, and ub is the population upper bound. The potential solution population's $x_N$, are utilized for minimizing or maximizing the following objective function, represented in Equation (2): the initial population of widowed blackbirds.

$$X_{N,d} = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,d} \\ x_{N,1} & x_{N,2} & x_{N,2} & \cdots & x_{N,d} \end{bmatrix} \quad (10)$$

$$lb \leq Xi \leq ub$$

## 3.6. Procreate:

Each couple is independently in a group that works at the same time to mate to produce a new generation. As discussed, matings are processed one at a time from other clients in the network. So in real time, about 10K eggs were produced. But only the strongest or

$$(\sigma 2_{VS})^* = \sum \max\{\sigma 2_{VS}(Si), \sigma 2_{VS}(Si)\} \quad (8)$$

The model determines the optimal threshold for each grayscale in the segmented area by maximizing the interclass contrast as much as possible.

$$\{t*[1], t*[2] \ldots t*[n]\} = \text{argmax}\ \sigma^2 * (t[1], t[2] \ldots .t[T] \quad (9)$$

The threshold method for automated images is presented in this example to deal with thyroid image segmentation. The over-segmentation procedure examines the between-class differences between background and foreground views and estimates noise levels

## 3.7. Meta-heuristic algorithm (bwo-mfo):

Algorithm discussed in detail. Population diversity can increase slowly against early

Assembly in the mayfly. In addition, it efficiently enables local optimization. The next section presents mathematical models of early population reproduction, presenting cannibalism, mutation, and convergence.

$xN$, is the black widow spider's population, N indicates population size, d denotes the number of decision variables of the problem, lb is the population lower bound, and ub is the population upper bound. The potential solution population's $xN$, are utilized for minimizing or maximizing the following objective function, represented in Equation (2):

N– Total count of samples ti - True sample value
$t_i$ - Corresponds to the predictive value
$t_i$ - Corresponds to the predictive value

Strongest spider in the wing survives. In this algorithm, an array is considered for the breeding process, this array-based breeding is performed until a widow array with a random number occurs and then μ is determined for breeding based on the following equation.

X1 and x2 ➜Parents

y1 and y2 ➜Offspring

$$y1 = \mu \times x1 + (1 - \mu) \times x2 \qquad (12)$$

$$y2 = \mu \times x2 + (1 - \mu) \times x1 \qquad (13)$$

i and j can be represented in the range of 1 to N

μ can be determined in the random range of 0 and 1.

## 3.8. Cannibalism:

Cannibalism can be executed in three ways: sexual homicide, child predation, and fratricide. In sexual cannibalism, the female eats the male spiders during or after mating. Here, the fitness value is considered more in this process. The second type of cannibalism is a child eating its parents based on its fit value of whether the dogs are weak or strong. Similarly, a spider that eats its brother, if it is weak, eats its brother. In this algorithm, the homicide rate is determined to calculate the survival rate.

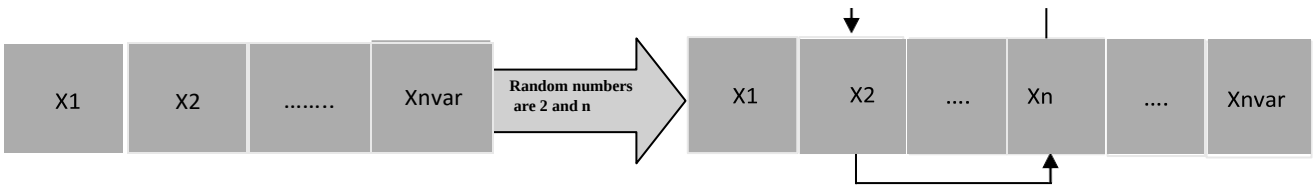


***Figure. 2: Mutation***

Fig. 2: Replacing two elements in a random array based on the selected solution. The mutation rate is used to calculate the mute pop

## 3.9. Mutation:

The mutation process is performed based on the random selection process to generate the population by choosing the mute pop number.

## 3.10. Convergence:

To increase population diversity, premature convergence must be prevented; The mayfly optimization algorithm should accelerate the convergence speed. Mayflies have the unique property of slowly increasing population diversity. This process helps to come out of a local optimum. Moreover, the method works well for detection and exploitation processes using MFOs. Here, the convergence process of the modified Mayfly optimization algorithm is included based on the use of equation (5). The mathematical scheme of the procedure is below.

$$\sigma^{i+1} = \sigma^{t} + \text{usign}[\text{rand} - 0.5] \oplus \text{Levy}(\beta) \qquad (14)$$

t,u ➜Random parameter is considered for uniform distribution

where, $\sigma_i^t$ i $^{th}$➜Solution vector or mayfly,

$X_i$ ➜No of iteration,

t,u ➜Random parameter is considered for uniform distribution

$\oplus$ ➜Dot product (entry wise multiplications)

rand ➜random initialization range is [0,1].

The position is given by the symbol [rand − 0.5] where you can consider only 3 values such as 0, 1, -1. Similarly, in equation (5) the sign of u integrates [rand − 0.5] and enables the mayfly to perform a random walk. Local minima can be reduced with this method, and global search can be improved by adding Modified Mayfly to the Black Widow Optimization algorithm Also, the mayfly algorithm is mainly based on random walks where step lengths contribute to detection steps in its process, are shown below

$$\text{Levy}(\beta) \sim \mu = t^{-1-\beta}, (0 \leq \beta \leq 2) \qquad (15)$$

The equation (6) is to compute the Levy random numbers.

$$\text{Levy}(\beta) \sim \frac{\theta \times \mu}{|v|^{1/\beta}} \qquad (16)$$

Where, μ and v →standard normal distributions

Γ→ a standard Gamma Function

β = 1.5, and φ is defined as follows:

$$((1 + \beta) \times \sin(\pi \times \beta /2) \times |\Gamma(((1 + \beta)/2) \times \beta \times 2^{(\beta-1/2)})^{1/\beta} \quad (18)$$

Global Search The power of the algorithm proposed here The MFO algorithm has been improved by adding Random Walk and Levy Flight to remove its weaknesses. And this is included structure can be reduced to minimum local structure particularly successful results in multi-model benchmark tasks and single- model tasks.

## 3.11. Pseudo-code of the bwo- mfo:

Input: maximum- number of iterations, number of cannibalization rate, procreate rate to num. ‘‘Nr’’ mutation rate at birth

**Initialization**

1. Initialize the population of BWO, D- dimensional problem, Chromosome’s D-dimensional array for each pop.
2. Fitness value (RMSE) evaluation until termination reached.
3. Determine nr and find the best solution in pop1 (population 1).

**Procreating and cannibalism**

4. For i = 1 tonrdo
5. Choose two solutions at random as parents from pop1.
6. Using equation 1, generate D children.
7. Choose two solutions randomly as parents from pop1.
8. Equation1 is considered for generating D children. Results:Objective function’s –RMSE,
9. Do cannibalism to for destroying father and some weakest children and generate new solution.
10. Generate new pop2 based on remaining solution.
11. End for

**Updating**

12. Modified-Mayfly optimization algorithm to update the population.
13. Return the best solution from pop;
14. The obtained best solution is given into the CNNLSTM classifier. The obtained best solution is given into the CNN-LSTM classifier.
15. The performance is evaluated to prove the best classifier.
16. Stop

## 3.12. Thyroid disease classification using vgg-19 with lstm:

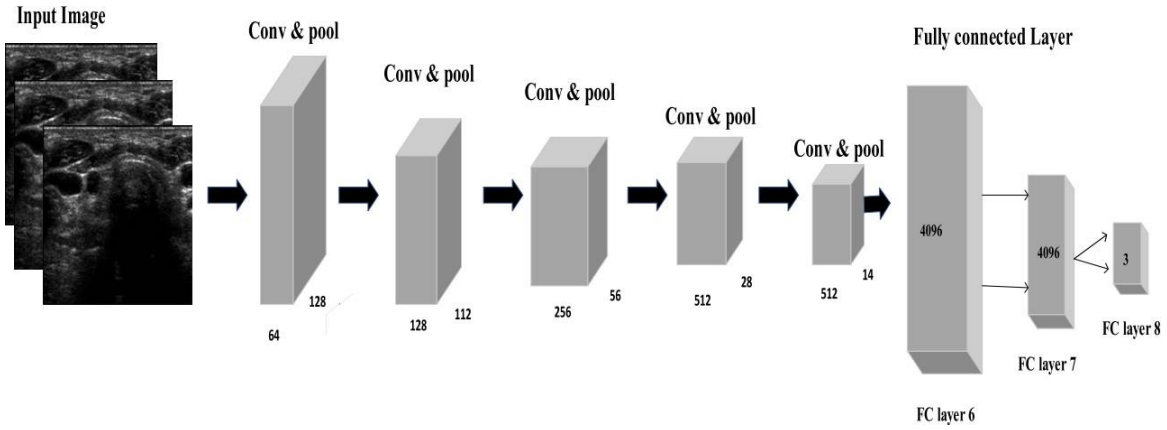


***Figure. 3: Vgg-19 architecture***

There are now several CNN models that improve performance and deeper construction. But it's deeper than that it is more difficult to train due to network requirements more data and larger values. This is the life that exists a large, structured dataset is critical to success and a generalizable system. Block-wise architecture Vgg-19 is shown in Fig.3.

The feature extraction procedure was developed in this paper via the Vgg- 19 network previously trained on Image Net.

These features were used as input signals for the 1 LSTM layer. The Vggnet has 16 convolutional layers with 3 × 3 filters With 1 step in the convolutional layer, 5 maximum pooling layers, and 1 flat 3 fully connected (FC) layers. Written by Wigg-19 There are a lot of smaller larger kernels stacked in filters that improve The trap is deep enough to remove the trap Hard features at low prices. One aspect of a general RNN is the short-term memory network (LSTM). The LSTM network can handle well the Problem of mold loss and long expansion time Replacement of primary secretory nerves and LSTM units in the RNN.

LSTM is based on three gates namely input, ''forget'', and output gate, where $x_t$ represents the current input; $c_t$ and $c_{t-1}$ refer to new & previous cell states, respectively; $h_t$ and $h_{t-1}$ are

present and past Exits anyway. The principle of the LSTM input gate has been shown that.

$$it = \sigma(Wi. [ht-1, xt] + bi) \quad (18)$$

$$C't = \tanh(Wi. [ht-1, xt] + bi) \quad (19)$$

$$Ct = ftCt-1 + itC't \quad (20)$$

Equation (24) is used to show that the information component is included by a sigmoid layer around $h_{t-1}$ and $x_t$.

After $h_{t-1}$ and $x_t$ intersect Equation (25) is used to provide additional data through the tanh layer. Equation (26) also relates the current moment data, $C'_t$ , to the long-term memory data, $C_{t-1}$&$C_t$ , where $W_i$ denotes the sigmoid output and $C'_t$ denotes the tanh output The weighting matrices define $W_i$, while $b_i$ denotes the bias input gate of the LSTM. Using dot product and sigmoid levels, LSTM's forgetting gate allows selective data transmission. Equation (27), where $W_f$ denotes the weighting matrix, $b_f$ is the offset, and σ is the sigmoid function, is set to the context from the previous cell with specific probabilities.

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f) \quad (21)$$

The conditions defining the LSTM's output gate, in equations (28), (29), are required for continuity by the $h_{t-1}$ and $x_t$ inputs.

State decision vectors, which carry additional information, are multiplied by the final $C't$ coming through the tanh layer.

$$O_t = \sigma(W_o. [h_{t-1}, x_t] + b_o) \quad (22)$$

$$h_t = O_t \tanh(C_t) \quad (23)$$

Where $W_o$ &$b_o$ are weighted matrices output gate & LSTM bias, respectively.

Hybrid algorithm and Vgg-19 - LSTM method developed in this paper for thyroid disease detection and classification using lateral full thyroid ultrasound dataset This algorithm was developed using hybrid optimization and LSTM mesh, where complex images are extracted and LSTM acts as a classifier. Network layers: one LSTM layer, two FC layers, and three dense layers depending on the activation function. LSTM level, followed by a

0.5 dropout. The ReLU function generates the convolutional with a size of 3 × 3 kernel used for feature extraction. The image stacking is dimensionally minimal, the size of the max-pooling layer is used with a 2 × 2 kernel. After analysis of construction time features, the ultrasound sorts the thyroid images by absolute overlap to determine whether they are in one of two groups (normal and abnormal).

## 3.13. Hyper-parameter tuning with hybrid (bwo-mfo):

Hyper-criteria selection is an important process in deep learning performance improvement; therefore, an appropriate strategy to optimize the tuning process should difficult, especially for a complex architecture like LSTM. So why, how effective customization is added has a huge impact on deep learning performance.

A hybrid meta-heuristic algorithm with Black Widow and Mayfly optimization hybrid BWO is proposed here. Also, it's hybrid.

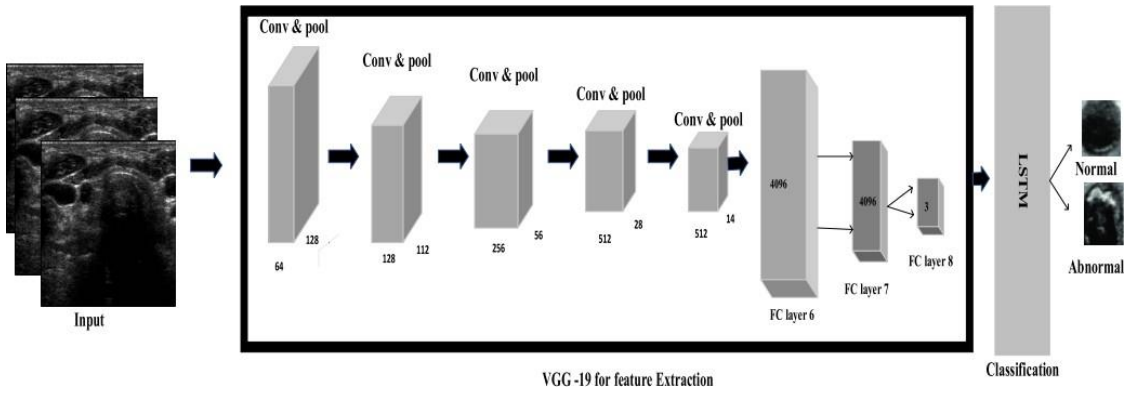


***Figure. 4: Classification structure***

Be considered. Tuning is very BWO method for solving optimization problems of choice and tuning the hyper- parameters makes a rather big difference other changes. Reasons for choosing a hybrid BWO should increase the convergence speed. The proposed algorithm, hybrid BWO, manages the entire population space, which is shared by BWO and MFO. To increase population diversity, premature convergence

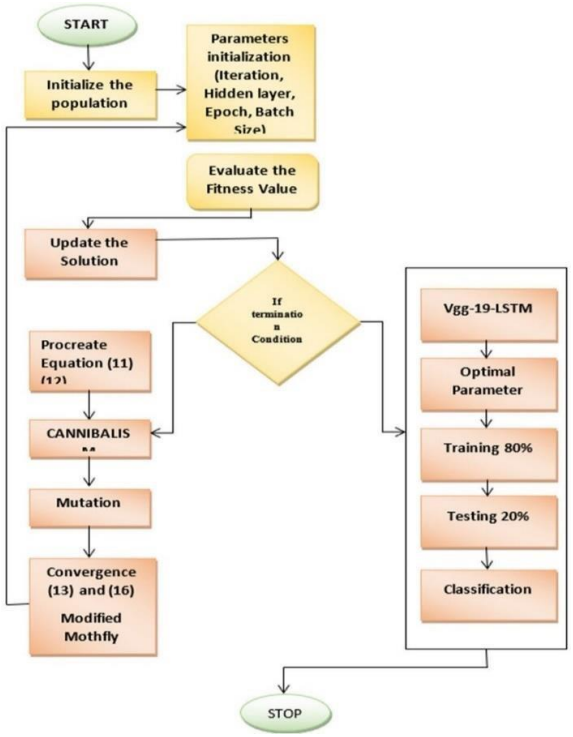


***Figure. 5: Flowchart of proposed Vgg-19 LSTM parameters***

Must be prevented; the mayfly optimization algorithm needs to accelerate the convergence speed. This algorithm provides high convergence speed; thus, performance and results can be better achieved. As discussed in the previous section, the LSTM model is enhanced by generating the parameters of the LSTM using the HM- BWO- scheme. The parameter types used for this function are Sequential Input Layer, LSTM Layer, Dropout Layer, Fully Connected Layer, and SoftMax Layers. The parameters used in this work are listed below. The basic parameters of the vgg- 19-LSTM are shown in Figure 8 illustrating the proposed workflow.

## 4. Experimental results and discussions:

### 4.1. Dataset description and evaluation metrics:

This study has 134 pictures and 99 cases of DDTI dataset to identify thyroid nodules. To develop tests, the data set was split into two sections, 20% and 80%, respectively. The average amount of time needed to find a thyroid nodule was 5.304 seconds, whereas the average training period was 102.93156 seconds. Furthermore, a comparison of the performance metrics of the specified approach with those of other approaches in use is used to assess the system's efficacy. A matrix of illusions is used to determine the range for business decisions. Performance measures for this test include recall, accuracy, precision,

$$\text{Accuracy} = \frac{(\text{Tp} + \text{Tn})}{(Tp + Tn + Fp + Fn)} \quad (24)$$

Accuracy: Accuracy is defined as the ratio of classified thyroid image quality to counted thyroid image quality and DNN-predicted incorrectly classified as positive when negative thyroid images were counted. Thyroid image segmentation.

$$\text{Precision} = \frac{\text{Tp}}{(Tp + Fp)} \quad (25)$$

Recall: Recall can be defined as the ratio of a set of positive thyroid images to counted positive thyroid images.

$$\text{Recall} = \frac{\text{Tp}}{(Tp + Fn)} \quad (26)$$

F1_Score: F1_Score is defined as the mean harmonic range between the precision and recall. (31)

$$\text{F1Score} = \frac{2(\text{Recall} \times \text{Precision})}{(Recall + Precision)} \quad (27)$$

Specificity: Specificity is distinct as the ratio of exactly the categorized count of negative thyroid images to the total count of negative thyroid images.

$$\text{Specificity} = \frac{\text{Tn}}{(Tn + Fp)} \quad (28)$$

FDR: FDR stands for False Discovery Rate. The FDR can be defined as the ratio of the total number of thyroid images classified as false positives to the total number of thyroid images classified as positive.

$$\text{FNR} = \frac{\text{Fn}}{(Tp + Fn)} \quad (29)$$

F1_score, specificity, detection time, training time, loss, FDR, FNR, FPR, MCC, and NPV.

Accuracy: Accuracy can be distinct as, from the overall count of thyroid images, a percentage of the count accurately categorized thyroid images

FPR: FPR is the false-positive rate. FPR consists of thyroid images that were not incorrectly classified as positive when negative thyroid images were counted.

$$\text{FPR} = \frac{\text{Fp}}{(Fp + Tn)} \quad (30)$$

MCC: MCC stands for Matthew's Correlation Coefficient. The MCC is a correlation coefficient calculated by the following parameters Tp, Tn, Fp, and Fn.

$$\text{MCC} = \frac{((Tp \times Tn) - (Fp \times Fn))}{\sqrt{(\text{Tp} + \text{Fp})\,(\text{Tp} + \text{Fn})\,(\text{Tn} + \text{Fp})\,(\text{Tn} + \text{Fn})}} \quad (31)$$

NPV: NPV has a negative predictive value.

The NPV can be defined at any threshold, and Specificity: Specificity is distinct as the ratio of exactly the categorized count of negative the probability of a true non-binding variable is correctly classified as non-binding. Thyroid images to the total count of negative thyroid images.

$$NPV = \frac{Tn}{(Tn + Fn)} \quad (32)$$

The experimental development has been completed on MAT LAB Software on a FDR: FDR stands for False Discovery Rate. The FDR can be defined as the ratio of the total number of thyroid images classified as false positives to the total number of thyroid images classified as positive. Computer with 12 GB RAM, Intel ®core (7M) i3-6100CPU @ 3.70 GHz processor. MATLAB code was used to configure and train VGG-19 with LSTM and Deep Learning (DL) algorithms and to identify hyperparameters using the Mayfly Optimization Approach (HBWO-MOA) with Black Widow Optimization.

## 5. Conclusion:

The use of artificial intelligence in self- assessment of thyroid symptoms is considered to be an important part of the future. A unique DL approach was used in this study for the diagnosis and classification of thyroid disease. Ultrasound images are acquired sequentially, and LSTM provides multiple fields of view to capture this type of data sequentially. Thyroid detection and classification by DNN Therefore the proposed method using model for thyroid disease detection (Vgg-19-LSTM) performed well large thyroid image data is needed for normal or abnormal detection and for classification improved accuracy is a major issue. The high dimensional feature subset is required to implement accurate deep learning (Vgg-19-LSTM) with low error rate and time. Therefore, to increase the prediction rate of thyroid images, this work uses preprocessing, segmentation, hybrid (BWO-MOA) optimization for feature selection, and (Vgg- 19-LSTM) classifier for thyroid. For disease detection and classification, the proposed method has good accuracy and high computational efficiency. In the future, we will extend this work for classification tasks by developing models that can accurately predict pixel labels indicative of thyroid disease with minimal training data.

## 6. References:

(1) A. A. Seyhan and C. Carini, ‘‘Are innovation and new technologies in pre- cision medicine paving a new era in patients centric care?’’ J. Translational Med., vol. 17, no.

1, pp. 1–28, Dec. 2019.

(2) I. U. Din, A. Almogren, M. Guizani, and M. Zuair, ‘‘A decade of Internet of Things: Analysis in the light of healthcare applications,’’ IEEE Access, vol. 7, pp. 89967– 89979, 2019.

(3) R. Boutaba, M. A. Salahuddin, N. Limam, S. Ayoubi, N. Shahriar,

(4) Estrada-Solano, and O. M. Caicedo, ‘‘A comprehensive survey on machine learning for networking: Evolution, applications and research opportunities,’’ J. Internet Services Appl., vol. 9, no. 1, pp. 1–99, Dec. 2018.

(5) A. Angelopoulos, E. T. Michailidis, N. Nomikos, P. Trakadas, a Hatziefremidis, S. Voliotis, and T. Zahariadis, ‘‘Tackling faults in the industry 4.0 era—A survey of machine-learning solutions and key aspects,’’ Sensors, vol. 20, no. 1, p. 109, Dec. 2019.

(6) D. Gupta, S. Sundaram, A. Khanna, A. E. Hassanien, and H. C. de Albuquerque, ‘‘Improved diagnosis of Parkinson’s disease using optimized crow search algorithm,’’ Comput. Electr. Eng., vol. 68, pp. 412–424, May 2018.

(7) Y. Ito, A. Miyauchi, and H. Oda, ‘‘Low-risk papillary microcarcinoma of the thyroid: A review of active surveillance trials,’’ Eur. J. Surgical Oncol., vol. 44, no. 3, pp. 307– 315, Mar. 2018.

(8) D. Grimm, ‘‘Cell and molecular biology of thyroid disorders,’’ Int. J. Mol. Sci., vol. 20, no. 12, p. 2895, Jun. 2019.

(9) Y. Zekri, F. Flamant, and K. Gauthier, ‘‘Central vs. peripheral action of thyroid hormone in adaptive thermogenesis: A burning topic,’’ Cells, vol. 10, no. 6, p. 1327, May 2021.

(10) E. Fröhlich and R. Wahl, ‘‘Microbiota and thyroid interaction in health and disease,’’ Trends Endocrinol. Metabolism, vol. 30, no. 8, pp. 479–490, Aug. 2019.

(11) T. J. Visser, ‘‘Regulation of thyroid function, synthesis, and function of thyroid hormones,’’ in Thyroid Diseases (Endocrinology), P. Vitti and

(12) Hegedüs, Eds. Cham, Switzerland: Springer, 2018, doi: 10.1007/978- 3-319- 45013-1_1.

(13) M. V. Deligiorgi and D. T. Trafalis, ‘‘The intriguing thyroid hormones– lung cancer association as exemplification of the thyroid hormones–cancer association: Three decades of evolving research,’’ Int. J. Mol. Sci., vol. 23, no. 1, p. 436, Dec. 2021.

(14) A. Abdel-Moneim, A. M. Gaber, S. Gouda, A. Osama, S. I. Othman, and G. Allam, ‘‘Relationship of thyroid dysfunction with cardiovascular diseases: Updated review on

heart failure progression,'' Hormones, vol. 19, no. 3, pp. 301–309, Sep. 2020.

(15) Y.-T. Shen, L. Chen, W.-W. Yue, and H.-X. Xu, ''Artificial intelligence in ultrasound,'' Eur. J. Radiol., vol. 139, Jun. 2021, Art. no. 109717.

(16) M. Guo and Y. Du, ''Classification of thyroid ultrasound standard plane images using ResNet-18 networks,'' in Proc. IEEE 13th Int. Conf. Anti- Counterfeiting, Secur., Identificat. (ASID), Oct. 2019, pp. 324–328.

(17) S. Hassantabar, M. Ahmadi, and A. Sharifi, ''Diagnosis and detection of infected tissue of COVID-19 patients based on lung X- ray image using convolutional neural network approaches,'' Chaos, Solitons Fractals, vol. 140, Nov. 2020, Art. no. 110170.

(18) T. Liu, Q. Guo, C. Lian, X. Ren, S. Liang, J. Yu, L. Niu, W. Sun, and D. Shen, ''Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convo- lutional neural networks,'' Med. Image Anal., vol. 58, Dec. 2019, Art. no. 101555.

(19) 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, pp. 1–11, Jan. 2019.

(20) R. Song, L. Zhang, C. Zhu, J. Liu, J. Yang, and T. Zhang, ''Thyroid nodule ultrasound image classification through hybrid feature cropping network,'' IEEE Access, vol. 8, pp. 64064–64074, 2020.

(21) O. Moussa, H. Khachnaoui, R. Guetari, and N. Khlifa, ''Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network,'' Int. J. Imag. Syst. Technol., vol. 30, no. 1, pp. 185–195, Mar. 2020.

(22) F. Abdolali, J. Kapur, J. L. Jaremko, M. Noga, A. R. Hareendranathan, and Punithakumar, ''Automated thyroid nodule detection from ultrasound imaging using deep convolutional neural networks,'' Comput. Biol. Med., vol. 122, Jul. 2020, Art. no. 103871.

(23) B. Shankarlal, P. D. Sathya, and V. P. Sakthivel, ''Computer-aided detec- tion and diagnosis of thyroid nodules using machine and deep learning classification algorithms,'' IETE J. Res., vol. 69, no. 2, pp. 995–1006, Feb. 2023.

(24) M. Hosseinzadeh, O. H. Ahmed, M. Y. Ghafour, F. Safara, H. K. Hama, S. Ali, B. Vo, and H.-S. Chiang, ''A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagno- sis in the Internet of Medical Things,'' J. Supercomput., vol. 77, no. 4, pp. 3616–3637, Apr. 2021.

(25) R. B. Namdeo and G. V. Janardan, ‘‘Thyroid disorder diagnosis by optimal convolutional neuron based CNN architecture,’’ J. Experim. Theor. Artif. Intell., vol. 34, no. 5, pp. 871–890, Sep. 2022.

(26) P. Wan, F. Chen, C. Liu, W. Kong, and D. Zhang, ‘‘Hierarchical tem- poral attention network for thyroid nodule recognition using dynamic CEUS imaging,’’ IEEE Trans. Med. Imag., vol. 40, no. 6, pp. 1646–1660, Jun. 2021.

(27) C. Chu, J. Zheng, and Y. Zhou, ‘‘Ultrasonic thyroid nodule detection method based on U-Net network,’’ Comput. Methods Programs Biomed. vol. 199, Feb. 2021, Art. No. 105906.

(28) A. Prasanth, P. Surendran, D. John, and B. Thomas, ‘‘A hybrid approach for web traffic prediction using deep learning algorithms,’’ in Proc. 9th Int. Conf. Electr. Electron. Eng. (ICEEE), Mar. 2022, pp. 383– 386, doi: 10.1109/ICEEE55327.2022.9772575.

(29) K. Bakhti and M. El Amin Larabi, ‘‘Comparing deep recurrent learning and convolutional learning for multi-temporal vegetation classification,’’ in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Jul. 2021, pp. 4392–4395, doi: 10.1109/IGARSS47720.2021. 9553175.

(30) S. K. Prabhakar and S. Lee, ‘‘SASDL and RBATQ: Sparse autoencoder with swarm based deep learning and reinforcement based Q-learning for EEG classification,’’ IEEE Open J. Eng. Med. Biol., vol. 3, pp. 58–68, 2022, doi: 10.1109/OJEMB.2022.3161837.

(31) D. S. AbdElminaam, N. Ahmed, M. Yasser, R. Ahmed, P. George, and M. Sahhar, ‘‘DeepCorrect: Building an efficient framework for auto correction for subjective questions using GRU_LSTM deep learn- ing,’’ in Proc. 2nd Int. Mobile, Intell., Ubiquitous Comput. Conf. (MIUCC), May 2022, pp. 33–40, doi: 10.1109/MIUCC55081.2022. 9781766.

(32) E. E. Eryilmaz, D. Ö. Sahin, and E. Kiliç, ‘‘Filtering Turkish spam using LSTM from deep learning techniques,’’ in Proc. 8th Int. Symp. Digit. Forensics Secur. (ISDFS), Jun. 2020, pp. 1–6, doi: 10.1109/ISDFS49300.2020.9116440.

(33) S. Iqbal, G. F. Siddiqui, A. Rehman, L. Hussain, T. Saba, U. Tariq, and A. A. Abbasi, ‘‘Prostate cancer detection using deep learning and traditional techniques,’’ IEEE Access, vol. 9, pp. 27085–27100, 2021, doi: 10.1109/ACCESS.2021.3057654.

(34) A. Sbrana, A. G. de Almeida, A. M. de Oliveira, H. S. Neto, J. P. C. Rimes, and M. C. Belli, ‘‘Plastic classification with NIR hyperspectral images and deep learning,’’ IEEE

Sensors Lett., vol. 7, no. 1, pp. 1–4, Jan. 2023, doi: 10.1109/LSENS.2023.3234401.

(35) H. Jiang, M. Gao, H. Li, R. Jin, H. Miao, and J. Liu, ‘‘Multi-learner based deep meta-learning for few-shot medical image classification,’’ IEEE J. Biomed. Health Informat., vol. 27, no. 1, pp. 17–28, Jan. 2023, doi: 10.1109/JBHI.2022.3215147.