

Intelligent Deep Learning Based Disease Diagnosis Using Biomedical Tongue Images

V. Thanikachalam^{1,*}, S. Shanthi², K. Kalirajan³, Sayed Abdel-Khalek^{4,5},
Mohamed Omri⁶ and Lotfi M. Ladhar⁷

¹Sri Sivasubramaniya Nadar College of Engineering, Chennai, 603110, India

²Department of Information Technology, Department of Electronics and Communication Engineering, CARE College of Engineering, Tiruchirappalli, 620009, India

³Department of Electronics and Communication Engineering, KPR Institute of Engineering and Technology, Coimbatore, 641407, India

⁴Department of Mathematics and Statistics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

⁵Mathematics Department, Faculty of Science, Sohag University, Sohag, Egypt

⁶Deanship of Scientific Research, King Abdulaziz University, Jeddah, Saudi Arabia

⁷Department of Electrical and Computer Engineering, Faculty of Engineering, King Abdulaziz University, Jeddah, Saudi Arabia

*Corresponding Author: V. Thanikachalam. Email: thanikachalamv@ssn.edu.in

Received: 16 June 2021; Accepted: 30 July 2021

Abstract: The rapid development of biomedical imaging modalities led to its wide application in disease diagnosis. Tongue-based diagnostic procedures are proficient and non-invasive in nature to carry out secondary diagnostic processes ubiquitously. Traditionally, physicians examine the characteristics of tongue prior to decision-making. In this scenario, to get rid of qualitative aspects, tongue images can be quantitatively inspected for which a new disease diagnosis model is proposed. This model can reduce the physical harm made to the patients. Several tongue image analytical methodologies have been proposed earlier. However, there is a need exists to design an intelligent Deep Learning (DL) based disease diagnosis model. With this motivation, the current research article designs an Intelligent DL-based Disease Diagnosis method using Biomedical Tongue Images called IDLDD-BTI model. The proposed IDLDD-BTI model incorporates Fuzzy-based Adaptive Median Filtering (FADM) technique for noise removal process. Besides, SqueezeNet model is employed as a feature extractor in which the hyperparameters of SqueezeNet are tuned using Oppositional Glowworm Swarm Optimization (OGSO) algorithm. At last, Weighted Extreme Learning Machine (WELM) classifier is applied to allocate proper class labels for input tongue color images. The design of OGSO algorithm for SqueezeNet model shows the novelty of the work. To assess the enhanced diagnostic performance of the presented IDLDD-BTI technique, a series of simulations was conducted on benchmark dataset and the results were examined in terms of several measures. The



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

resultant experimental values highlighted the supremacy of IDLDD-BTI model over other state-of-the-art methods.

Keywords: Biomedical images; image processing; tongue color image; deep learning; squeezeNet; disease diagnosis

1 Introduction

For a long-known time, human tongue is assessed visually for signs and symptoms of diseases, since it is a non-invasive approach. Conventionally, medical experts observe its features while the diagnosis outcomes depend upon their years of experience [1]. But, subjectivity and ambiguity affect the results of such visual diagnosis. In order to eliminate such qualitative factors, tongue images are examined objectively for which a novel method is provided for disease diagnosis. Such model can also reduce the physical damage caused to persons (related to other medical investigations). In advanced computerized tongue image analyses, texture and colour features are more commonly observed [2]. There has been no or only few studies conducted upon tongue image analyses utilizing geometry features. However, conventional medicines like Traditional Chinese medicine (TCM) assess the shape of tongue to determine a person's disease. The study conducted earlier [3] presented a method to manually identify the shapes of tongue depending on geometric features. With the help of support decision tool, it weighs the comparative effects of geometric features and categorizes the tongue image under one out of six tongue shapes such as, rectangle, hammer, obtuse triangle, acute triangle, round, and square (on the basis of TCM). The combination of signal decomposition with dimension reduction techniques along with neural networks also provide good accuracy for both subject dependent and independent disease diagnosis using biomedical images [4].

Few disease diagnosis systems make decisions based on distinct features extracted during feature extraction phase [5]. So, the quality of output is straightaway compared with the segmented tongue image present in the source. Thus, it is crucial to perform accurate and quick tongue image segmentation. So far, several automated tongue image segmentation methods have been presented based on detailed systems [6]. Especially, Bi-Elliptical Deformable Contour (BEDC) integrates module-based methods and Active Contour Models (ACM). In comparison with other techniques, this technique attained significant segmentation outcomes. Its segmentation quality was heavily based on few preceding knowledge and is sensitive to location and early curves present in the tongue. In order to confront this, Ning et al. [7] substituted module-based methods with the help of region merging approach in order to obtain coarse segmentation outcomes. ACM was utilized in this study for post-processing phase to attain improved segmentation efficiency related to BEDC. Though, earlier researchers [8] depend mostly on existing knowledge, the location data of early markers must be defined manually. Wu et al. [9] implemented an edge and region-based combinational method which could additionally eliminate the impact of nearby artefact in tongue and accordingly increase the results and strength of boost segmentation. Hence, this technique displayed weak strength in several situations. Particularly, the segmentation result was poor, if tongue poses are uneven and if the tongue is nearly shadowed by the lips [10,11].

Li et al. [12] proposed a simple and efficient tongue image segmentation system. At first, this technique extracts the primary tongue body areas with the performance of image threshold on transformed hue component in HSI colour space. Then, the image threshold results in red component of the actual tongue image which is chosen in an adaptive way to detect the gap between upper lip and tongue body root. At last, the primary body part of the tongue is

established by eliminating fake object areas, for instance upper lip, so as to attain the latter tongue image segmentation outcome. Zhang et al. [13] improved a disease diagnoses approach for diabetes in accordance with regularized tongue image by SVM. The texture and colour features of the tongue image are extracted and used as input variables. Then, diabetes diagnosis method is trained using SVM. The study assessed the impact of this combination in improving the incorporation of SVM kernel parameter and input variable.

Fan et al. [14] studied a number of tongue features in patients with gastric and diabetic symptoms. The researchers collected the images with the help of digital tongue imaging technique. During feature extraction, texture and four TCM tongue features were detected such as slenderness, coating colour, constitution colour, plumpness, and cracks. During classification, two distinct classification methods were utilized such as SVM & RF to categorize the symptoms of TCM & DM gastric diseases. Wang et al. [15] presented an AI structure with the help of DCNN framework to recognize tooth-marked tongue. Firstly, they made use of massive dataset containing 1548 tongue images captured by several tools. Meng et al. [16] presented an innovative feature extraction module termed CHDNet to extract unbiased features and reduce the human effort upon tongue diagnosis in TCM. CNN model is known to have primarily concentrated on learning convolution filter and adjusting the weights amongst itself. So, this model has two challenges such as inadequate ability & redundancy in dealing with unbalanced sample distribution. So, the researchers presented local response normalization and high dispersal function to address these problems.

The current research article designs an Intelligent Deep Learning based Disease Diagnosis Model using Biomedical Tongue Images, called IDLDD-BTI model. The proposed IDLDD-BTI model incorporates fuzzy-based Adaptive Median Filtering (FADM) technique for noise removal process. Besides, SqueezeNet model is also employed as a feature extractor that tunes the hyper-parameters of SqueezeNet model using Oppositional Glowworm Swarm Optimization (OGSO) algorithm. At last, Weighted Extreme Learning Machine (WELM) classifier is employed to allocate proper class labels to input tongue color images. In order to assess the improved diagnostic results of the presented IDLDD-BTI model, extensive experimental analysis was performed on benchmark dataset.

2 The Proposed IDLDD-BTI Model

Fig. 1 shows the working principle involved in the presented IDLDD-BTI model. The presented IDLDD-BTI model incorporates FADM-based noise removal technique, SqueezeNet-based feature extractor, OGSO-based parameter tuning, and WELM-based classification. The detailed working processes of these modules are discussed in succeeding sections.

2.1 FADM Based Preprocessing

During pre-processing stage, FADM technique is employed to remove the noise. Binary decision map $B(u, v)$ performs as 'switch' to turn on the filters, if a noise pixel is distinguished. For instance, $B(u, v) = 0$. Then, the filtering act is skipped, if $B(u, v) = 1$ and

- (1) Define the amount of noise-free pixel $G(u, v)$ and calculate the number of '1s' in $B(u, v)$
- (2) Develop $W_f(u, v)$ with 1 pixel at one of its each four sides (i.e., $L_f = Lr + 1$), if $G(u, v) < 1$. Repeating steps 1 and 2, still the condition $G(u, v) \geq 1$ is fulfilled.
- (3) Estimate the median pixel $M(u, v)$ with the help of every noise-free pixel from current (u, v) . The median pixel $M(u, v)$ is provided herewith.

$$m(u, v) = \text{median} \{x(u+p, v+q)\} \text{ with } B(t)(u+p, v+q) = 1$$

4. Compute the variance between central pixel and the neighboring pixels from filtering window [17].

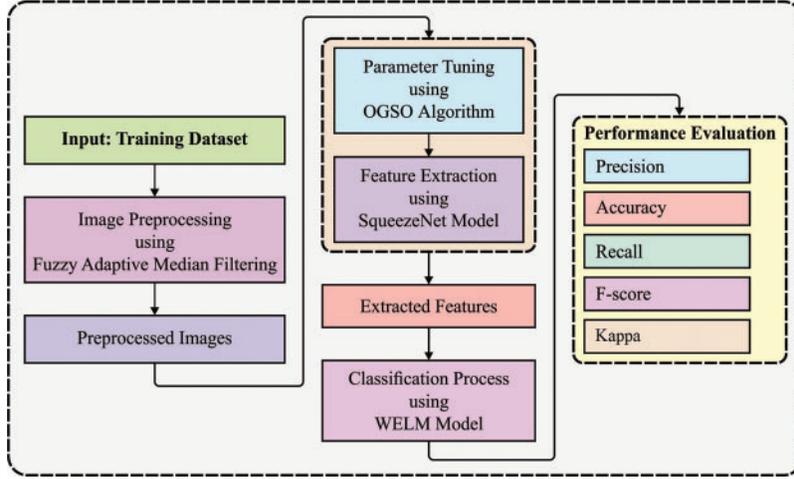


Figure 1: Overall process of IDLDD-BTI model

Based on:

$$D_l(u, v) = \max \{D_1(m)\} \quad (1)$$

where

$$D_1(m) = |x|(u+k, v+l) - x(u, v) \quad (2)$$

$$f(u, v) = \begin{cases} (D(u, v) - T_1) & D(u, v) < T_1 \\ \frac{(D(u, v) - T_1)}{T_2 - T_1} & T_1 D(u, v) < T_2 \\ D(u, v) T_2 & \end{cases}$$

where T_1 and T_2 are two existing thresholds.

6. Calculate the restoration term $y(u, v)$ as follows:

$$y(u, v) = F(u, v) \cdot M(u, v) + [F(u, v) \cdot x(u, v)] \quad (3)$$

2.2 SqueezeNet Based Feature Extraction

SqueezeNet is a convolutional network that applies optimum efficiency than AlexNet with $50\times$ smaller parameter. It has 15 layers with five distinct layers as two convolutional layers, three max pooling layers, eight fire layers, one global average pooling layer, and one output layer softmax. The structure of network is shown in Fig. 2. As demonstrated in the figure, $K\times K$ notations signify the receptive field size of filter, 's' implies the stride sizes and l represents the length of the feature map, correspondingly. The input of networks consists of 227×227 dimensions with RGB channel.

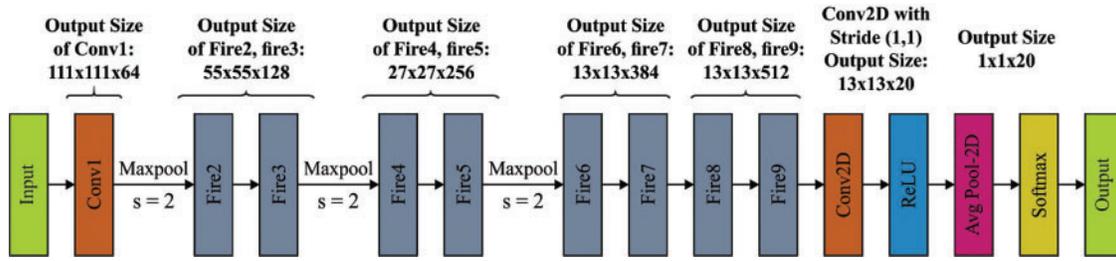


Figure 2: SqueezeNet architecture

The generalization of input images by convolutional and max-pooling layers is executed. The convolutional layers get convoluted amongst weights as well as small regions from input volume, with 3×3 kernels. All the convolutional layers carry out an elementwise activation purpose as a positive part of their arguments [18]. The outcome tensor scale and input of fire depend upon each other. The squeeze stage utilizes the filter sized 1×1 , but the development utilizes the filter sized 1×1 and 3×3 . Initially, an input tensor $H \times W \times C$ permits with squeeze and the amount of convolutional layers is equivalent to $C/4$ of the number of input tensor channels.

Eventually, expansion outcomes are stacked from in-depth dimensions of input tensors by concatenating function. Fig. 3 shows a review on fire layers and sub operations. Considering that FM and C determine the feature maps and channels, the resultant layers $f\{y\}$ of squeeze function with kernel w are written as follows

$$f\{y\} = \sum_{fm1=1}^{FM} \sum_{c=1}^c w_c^f x_c^{fm1} \tag{4}$$

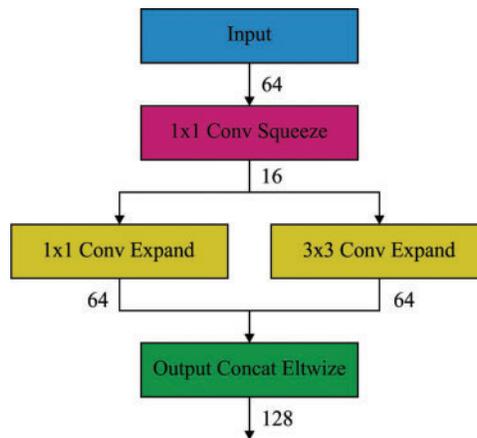


Figure 3: Structure of fire layer

At this point, $f\{y\} \in \mathbb{R}^N$ and $w \in \mathbb{R}^{C \times 1 \times FM^2}$. The squeeze outcomes are determined as weighted combination of feature maps of distinct tensors. In the network, max-pool layers apply a down sampling function in addition with spatial dimensional and global average pool that altogether changes the feature maps of classes as to one value.

2.3 OGSO-Based Parameter Tuning

OGSO algorithm is employed in this study to adjust the hyperparameters involved in WELM model in an optimal manner. GSO is a smart swarm optimization method which is utilized on the basis of luminescent features of fireflies. In GSO method, glowworm swarm is dispersed in solution space and Fitness Function (FF) of all the positions of glowworm. The powerful glowworm shows the maximal brightness and its optimum location is made up of maximal FF rate. Glowworm has a vital line of sight called decision domain that has the density range for adjacent nodes. On the other hand, the decision radius is constrained, when glowworm travels to same kind of robust fluorescence in a decision domain. After obtaining high iteration values, all the glowworms are located in optimal position. It is limited by five stages:

- Fluoresce in concentration
- Neighbor set
- Decision domain radius
- Moving probability
- Glowworm position

The fluorescence, in concentration upgrading technique, is given as follows.

$$I_u(z) = (1 - \alpha)I_u(r - 1) + \beta f(x_u(z)), \quad (5)$$

whereas $I_u(f)$ denotes the fluoresce in concentration of i^{th} glowworm at time f , α indicates the fluoresce in volatilization coefficients, β represents the fluoresce in enhancement factor, $f(x)$ signifies FF and $x_u(r)$ denotes the position of glowworm u at f time that is given by,

$$N_u(t) = \{v : \|x_v(r) - x_u(z)\| < r_d^u; I_u(z) < I_v(z)\}, \quad (6)$$

Whereas $N_u(f)$ denotes the neighbor group of u^{th} glowworm at time r and $r_d^u(r)$ represents the radius of decision field of u^{th} glowworm at moment f which is given by

$$r_d^u(z + 1) = \min \{r_s, \max \{r_d^u(z) + \gamma(n_u - |N_u(z)|)\}\}, \quad (7)$$

Whereas r_s represents the obtained radius of glowworm, γ indicates the value of decision domain, and n_i denotes the neighbor threshold. The moving likelihood of upgraded method is illustrated herewith.

$$p_{uv}(r) = \frac{I_v(z) - I_u(z)}{\sum_{k \in N_z} I_k(z) - I_u(r)}, \quad (8)$$

Whereas $p_{uv}(z)$ represents likelihood in which the glowworm u travels towards the glowworm v at r time which is given by,

$$x_u(z + 1) = x_u(z) + s \left(\frac{x_v(z) - x_u(z)}{\|x_v(z) - x_u(z)\|} \right), \quad (9)$$

Opposition Based Learning (OBL) is mainly employed from efficient optimization procedures to improve the convergence speed of distinct heuristic optimization modules. The efficient performance of OBL contributes in the evaluation of opposite and current populations in the same generation to identify an optimal candidate solution for the given problem. OBL module is efficiently employed in distinct metaheuristics to improve the speed of convergence. The module of the opposite amount should be described in OBL.

Consider $N \in N[x, y]$ to be a real number. The opposite number N^0 is given by

$$N^0 = x + y - N \tag{10}$$

In the event of d dimension search space, the explanation may be extended as follows:

$$N_i^0 = x_i + y_i - N_i \tag{11}$$

where (N_1, N_2, \dots, N_d) denotes the d-dimensional search space and $N_i[x_i, y_i]$, $i = 1, 2, \dots, d$. OBL technique is employed in the initiation procedure of GSO method and for all the iterations in the application of jumping rate.

2.4 WELM Based Classification

At the final stage, WELM classifier is applied to allocate proper class labels for the input tongue color images. ELM is utilized as the classifier of balanced dataset whereas WELM is utilized as a classifier for imbalanced dataset. Training dataset has N different instances, (x_i, z_i) , $i = 1, 2, \dots, N$. A single hidden layer NN with L hidden layer nodes are written as given herewith.

$$\sum_{i=1}^L \beta_i \cdot l(w_i \cdot x_j + b_i) = z_j, \quad j = 1, \dots, N \tag{12}$$

where w_i represents a single hidden layer input weight, $l()$ implies the activation functions, β_i refers to output weights, and b_i stands for single hidden layers bias. Eq. (12) simply represents the notion discussed above.

$$S\beta = T \tag{13}$$

where S implies the output matrix of single hidden layers

$$S(w_1, \dots, w_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{pmatrix} l(w_1 \cdot x_1 + b_1) \cdots l(w_L \cdot x_1 + b_L) \\ \vdots & \ddots & \vdots \\ l(w_1 \cdot x_N + b_1) \cdots l(w_L \cdot x_N + b_L) \end{pmatrix}_{N \times L} \tag{14}$$

In accordance with Karush-Kuhn-Tucker theory, Lagrangian factor has been recognized in altering the trained ELM as a dual issue. The output weights β is estimated as given herewith.

$$\beta = S^T \left(\frac{1}{C} + SS^T \right)^{-1} T \tag{15}$$

where C signifies the regularization coefficients [19]. Hence, the output functions of ELM classification is written as follows

$$F(x) = s(x)S^T \left(\frac{1}{C} + SS^T \right)^{-1} T = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{1}{C} + \chi \right)^{-1} T \tag{16}$$

where χ implies the kernel matrix which is computed as follows

$$\chi = SS^T = s(x_i)s(x_j) = K(x_i, x_j) \tag{17}$$

It is clear in (13) that the hidden layer feature maps $s(x)$ implies the independence in classifier outcome of ELM. The classifier outcome is only connected with kernel functions, $K(x, y) \cdot K(x, y)$ assumes an inner product procedure, and the amount of hidden layer nodes are no outcome on output resultant.

$$K(x, y) = \exp(-\gamma \|x - y\|^2). \quad (18)$$

So, the KELM classifier efficiency is distributed into two parameters such as kernel function parameters, γ and penalty parameters, C . The original benefits of ELM can be achieved in WELM too by maintaining the weight for different instances to deal with imbalanced classifier issues. The situation of output purpose is computed as given herewith.

$$F(x) = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^T \left(\frac{1}{C} + W\chi \right)^{-1} WT, \quad (19)$$

$$W = \text{diag}(w_{ii}), i = 1, 2, \dots, N \quad (20)$$

where W represents the weight matrix. WELM has two weightage models as given below.

$$w_{ii} = \frac{1}{\#(z_i)}, \quad (21)$$

$$w_{ii} = \begin{cases} \frac{0.618}{\#(z_i)}, & \text{if } z_i > \bar{z} \\ \frac{1}{\#(z_i)}, & \text{otherwise} \end{cases} \quad (22)$$

where $\#(z_i)$ implies the amount of instances going to class $z_i, i = 1, \dots, m$. m represents the amount of classes. \bar{z} implies the average of complete instances of each class.

3 Performance Validation

The proposed model was validated using tongue image dataset that comprises of images under different class labels namely Healthy, Chronic Kidney Disease (CKD), nephritis, verrucous gastritis, pneumonia, nephritis syndrome, chronic cerebral circulation insufficiency, upper respiratory tract infection, erosive gastritis, Coronary Heart Disease, chronic bronchitis and mixed hemorrhoid. The dataset includes 936 images with each class containing 78 images. Fig. 4 depicts the sample test images from the dataset. The parameter setting of the SqueezeNet model is given as follows. Batch size: 500, max. epochs:15, learning rate: 0.05, dropout rate: 0.2, and momentum: 0.9.

Tab. 1 and Fig. 5 shows a brief classification result attained by IDLDD-BTI technique with a training size of 60%. The outcomes exhibit that the proposed IDLDD-BTI model accomplished better results under all the runs. For instance, under run-1, the IDLDD-BTI model obtained a precision of 99.43%, recall of 99.40%, accuracy of 98.13%, F1-score of 98.97%, and kappa of 97.63%. At the same time, under run-3, the IDLDD-BTI model accomplished a precision of 98.47%, recall of 98.03%, accuracy of 98.23%, F1-score of 98.28%, and kappa of 97.97%. Moreover, under run-5, the IDLDD-BTI model obtained a precision of 98.77%, recall of 99.45%, accuracy of 99.32%, F1-score of 98.16%, and kappa of 98.89%.



Figure 4: Sample tongue images

Table 1: Performance analysis of the proposed IDLDD-BTI method with 60% training size under different measures

No. of folds	Precision	Recall	Accuracy	F1-Score	Kappa
Training size = 60%					
Run-1	99.43	99.40	98.13	98.97	97.63
Run-2	98.41	98.00	98.14	99.89	98.55
Run-3	98.47	98.03	98.23	98.28	97.97
Run-4	99.29	98.10	98.64	99.72	98.66
Run-5	98.77	99.45	99.32	98.16	98.89
Average	98.87	98.60	98.49	99.00	98.34

[Tab. 2](#) and [Fig. 6](#) shows a brief classification results attained by IDLDD-BTI approach with 70% training size. The outcomes infer that IDLDD-BTI model accomplished optimum outcomes in all the runs. For instance, under run-1, the IDLDD-BTI model achieved a precision of 99.81%, recall of 95.75%, accuracy of 98.02%, F1-score of 99.03%, and kappa of 98.65%. Likewise, under run-3, the IDLDD-BTI model attained a precision of 98.88%, recall of 97.19%, accuracy of 98.54%, F1-score of 97.89%, and kappa of 97.49%. Also, under run-5, the IDLDD-BTI method accomplished a precision of 98.91%, recall of 96.56%, accuracy of 97.57%, F1-score of 98.69%, and kappa of 97.97%.

[Tab. 3](#) and [Fig. 7](#) shows a detailed classification result attained by IDLDD-BTI algorithm with 80% training size. The outcomes demonstrate that the proposed IDLDD-BTI model accomplished optimal outcomes under all the runs. For instance, under run-1, the IDLDD-BTI model obtained a precision of 97.56%, recall of 98.33%, accuracy of 97.64%, F1-score of 97.07%, and

kappa of 96.81%. Concurrently, under run-3, IDLDD-BTI model achieved a precision of 98.54%, recall of 99.31%, accuracy of 98.89%, F1-score of 96.25%, and kappa of 95.70%. In addition, under run-5, the IDLDD-BTI method yielded a precision of 96.55%, recall of 99.30%, accuracy of 98.23%, F1-score of 99.45%, and kappa of 98.15%.

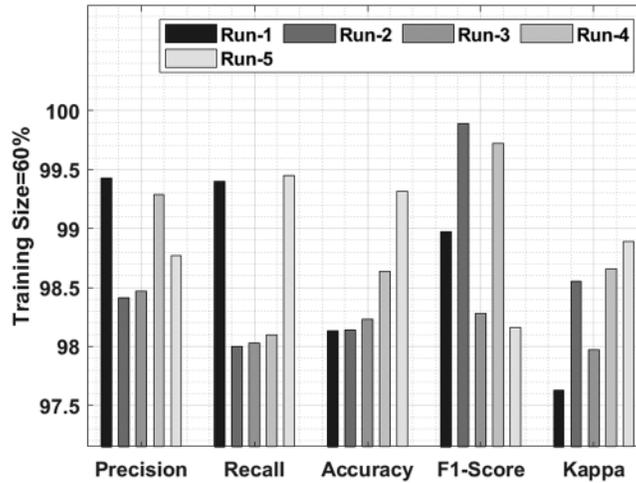


Figure 5: Results of the analysis of IDLDD-BTI model with 60% training size

Table 2: Performance analysis of the proposed IDLDD-BTI method with 70% training size under different measures

No. of folds	Precision	Recall	Accuracy	F1-Score	Kappa
Training size = 70%					
Run-1	99.81	95.75	98.02	99.03	98.65
Run-2	99.34	97.31	97.98	98.70	97.95
Run-3	98.88	97.19	98.54	97.89	97.49
Run-4	98.47	95.49	98.02	98.70	97.98
Run-5	98.91	96.56	97.57	98.69	97.97
Average	99.08	96.46	98.03	98.60	98.01

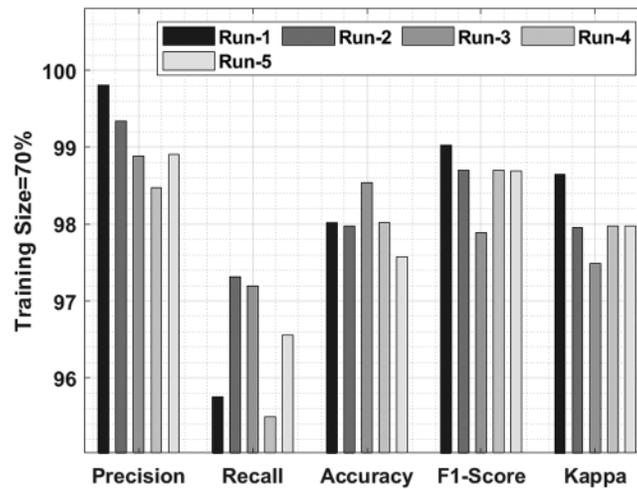


Figure 6: Result analysis of IDLDD-BTI model on training size of 70%

Table 3: Performance analysis of proposed IDLDD-BTI method on training size = 80% in terms of different measures

No. of folds	Precision	Recall	Accuracy	F1-Score	Kappa
Training size = 80%					
Run-1	97.56	98.33	97.64	97.07	96.81
Run-2	98.23	97.75	98.38	96.14	97.98
Run-3	98.54	99.31	98.89	96.25	95.70
Run-4	97.57	98.53	97.74	97.25	96.86
Run-5	96.55	99.30	98.23	99.45	98.15
Average	97.69	98.64	98.18	97.23	97.10

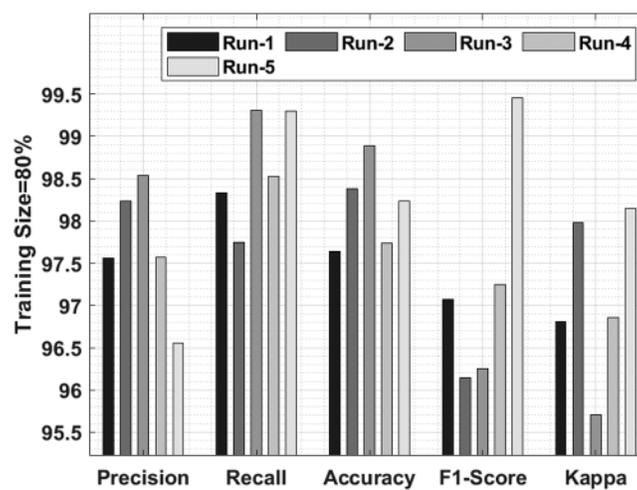


Figure 7: Result analysis of IDLDD-BTI model on training size of 80%

Tab. 4 and Fig. 8 shows the average results of analysis of IDLDD-BTI technique under varying TS. The experimental outcomes infer that the presented IDLDD-BTI technique produced a maximum average precision of 98.55%, recall of 97.90%, accuracy of 98.23%, F1-score pf 98.28%, and kappa of 97.82%.

Table 4: Performance analysis of the proposed IDLDD-BTI method under different training sizes and different measures

Measures	Precision	Recall	Accuracy	F1-Score	Kappa
Training Size = 60%	98.87	98.60	98.49	99.00	98.34
Training Size = 70%	99.08	96.46	98.03	98.60	98.01
Training Size = 80%	97.69	98.64	98.18	97.23	97.10
Average	98.55	97.90	98.23	98.28	97.82

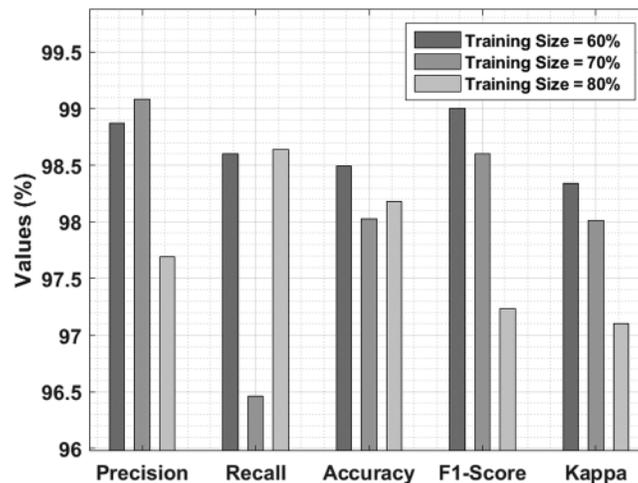
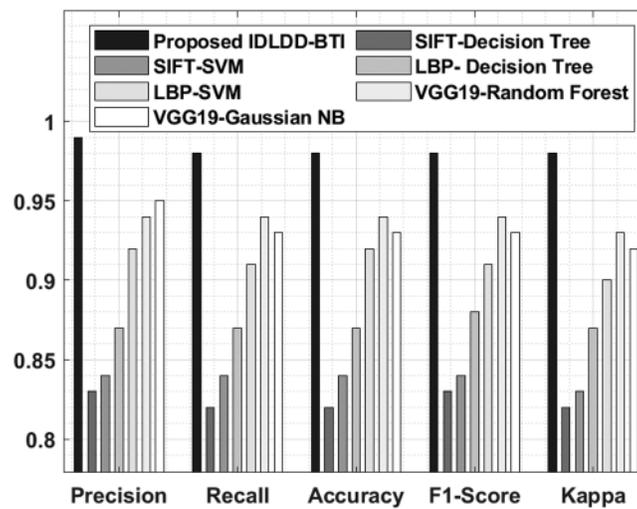


Figure 8: Results of the analysis of IDLDD-BTI model under different training sizes

A detailed comparative results of the analysis is shown in Tab. 5 and Fig. 9 to exhibit the improved diagnostic performance of IDLDD-BTI technique [20,21]. From the results, it is evident that SIFT-Decision Tree and SIFT-SVM models produced ineffective results with minimal classification outcomes. Moreover, the LBP-Decision Tree model accomplished slightly improved classifier results over earlier models. Eventually, the LBP-SVM, VGG19-Random Forest, and VGG19-Gaussian NB models demonstrated moderately closer performances. However, the proposed IDLDD-BTI method outperformed other models by accomplishing a superior performance with maximal precision of 0.99, recall of 0.98, accuracy of 0.98, F1-score of 0.98, and kappa of 0.98.

Table 5: Performance analysis of the proposed IDLDD-BTI model with recent methods under different measures

Methods	Precision	Recall	Accuracy	F1-Score	Kappa
Proposed IDLDD-BTI	0.99	0.98	0.98	0.98	0.98
SIFT-Decision tree	0.83	0.82	0.82	0.83	0.82
SIFT-SVM	0.84	0.84	0.84	0.84	0.83
LBP- Decision tree	0.87	0.87	0.87	0.88	0.87
LBP-SVM	0.92	0.91	0.92	0.91	0.90
VGG19-Random forest	0.94	0.94	0.94	0.94	0.93
VGG19-Gaussian NB	0.95	0.93	0.93	0.93	0.92

**Figure 9:** Comparative analysis of IDLDD-BTI model with different measures

4 Conclusion

The current research article developed a new IDLDD-BTI method to diagnose the diseases using tongue color images. The presented IDLDD-BTI model incorporates FADM-based noise removal technique, SqueezeNet-based feature extractor, OGSO-based parameter tuning, and WELM-based classification. OGSO algorithm classifier helps in fine tuning the hyperparameters of SqueezeNet model in such a way that the diagnostic performance gets considerably increased to a maximum level. In order to verify the improved diagnostic results of the presented IDLDD-BTI model, extensive experimental analysis was conducted using benchmark dataset. The resultant experimental values highlighted the supremacy of IDLDD-BTI model over existing state-of-the-art techniques. In future, the presented IDLDD-BTI model can be used as a smartphone application to diagnose patients from remote locations too.

Funding Statement: This paper was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, Saudi Arabia, under grant No. (D-79-305-1442). The authors, therefore, gratefully acknowledge DSR technical and financial support.

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

References

- [1] B. Zhang and H. Zhang, "Significant geometry features in tongue image analysis," *Evidence-Based Complementary and Alternative Medicine*, vol. 2015, pp. 1–8, 2015.
- [2] K. Shankar, E. Perumal, P. Tiwari, M. Shorfuzzaman and D. Gupta, "Deep learning and evolutionary intelligence with fusion-based feature extraction for detection of COVID-19 from chest X-ray images," *Multimedia Systems*, pp. 1–13, 2021. <https://doi.org/10.1007/s00530-021-00800-x>.
- [3] B. Huang, J. Wu, D. Zhang and N. Li, "Tongue shape classification by geometric features," *Information Sciences*, vol. 180, no. 2, pp. 312–324, 2010.
- [4] M. T. Sadiq, X. Yu, and Z. Yuan, "Exploiting dimensionality reduction and neural network techniques for the development of expert brain–computer interfaces," *Expert Systems with Applications*, vol. 164, pp. 114031, 2021.
- [5] J. Zhou, Q. Zhang, B. Zhang and X. Chen, "Tonguenet: A precise and fast tongue segmentation system using u-net with a morphological processing layer," *Applied Sciences*, vol. 9, no. 15, pp. 3128, 2019.
- [6] J. Banumathi, A. Muthumari, S. Dhanasekaran, S. Rajasekaran, I. V. Pustokhina *et al.*, "An intelligent deep learning based xception model for hyperspectral image analysis and classification," *Computers, Materials & Continua*, vol. 67, no.2, pp. 2393–2407, 2021.
- [7] J. Ning, D. Zhang, C. Wu and F. Yue, "Automatic tongue image segmentation based on gradient vector flow and region merging," *Neural Computing and Applications*, vol. 21, no. 8, pp. 1819–1826, 2012.
- [8] J. Ning, L. Zhang, D. Zhang and C. Wu, "Interactive image segmentation by maximal similarity based region merging," *Pattern Recognition*, vol. 43, no. 2, pp. 445–456, 2010.
- [9] K. Wu and D. Zhang, "Robust tongue segmentation by fusing region-based and edge-based approaches," *Expert Systems with Applications*, vol. 42, no. 21, pp. 8027–8038, 2015.
- [10] K. Shankar, E. Perumal, M. Elhoseny and P. T. Nguyen, "An iot-cloud based intelligent computer-aided diagnosis of diabetic retinopathy stage classification using deep learning approach," *Computers, Materials & Continua*, vol. 66, no.2, pp. 1665–1680, 2021.
- [11] M. Elhoseny, M. M. Selim and K. Shankar, "Optimal deep learning based convolution neural network for digital forensics face sketch synthesis in internet of things (IoT)," *International Journal of Machine Learning and Cybernetics*, Jul. 2020. <https://doi.org/10.1007/s13042-020-01168-6>.
- [12] Z. Li, Z. Yu, W. Liu, Y. Xu, D. Zhang *et al.*, "Tongue image segmentation via color decomposition and thresholding," *Concurrency and Computation: Practice and Experience*, vol. 31, no. 23, pp. 1–9, 2019.
- [13] J. Zhang, J. Xu, X. Hu, Q. Chen, L. Tu *et al.*, "Diagnostic method of diabetes based on support vector machine and tongue images," *BioMed Research International*, vol. 2017, pp. 1–9, 2017.
- [14] S. Fan, B. Chen, X. Zhang, X. Hu, L. Bao *et al.*, "Machine learning algorithms in classifying TCM tongue features in diabetes mellitus and symptoms of gastric disease," *European Journal of Integrative Medicine*, vol. 43, pp. 101288, 2021.
- [15] X. Wang, J. Liu, C. Wu, J. Liu, Q. Li *et al.*, "Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark," *Computational and Structural Biotechnology Journal*, vol. 18, pp. 973–980, 2020.
- [16] D. Meng, G. Cao, Y. Duan, M. Zhu, L. Tu *et al.*, "Tongue images classification based on constrained high dispersal network," *Evidence-Based Complementary and Alternative Medicine*, vol. 2017, pp. 1–12, 2017.
- [17] V. Thirilogasundari, V. S. Babu and S. A. Janet, "Fuzzy based salt and pepper noise removal using adaptive switching median filter," *Procedia Engineering*, vol. 38, pp. 2858–2865, 2012.

- [18] F. Ucar and D. Korkmaz, "COVIDiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from X-ray images," *Medical Hypotheses*, vol. 140, pp. 109761, 2020.
- [19] H. Zhu, G. Liu, M. Zhou, Y. Xie, A. Abusorrah *et al.*, "Optimizing weighted extreme learning machines for imbalanced classification and application to credit card fraud detection," *Neurocomputing*, vol. 407, pp. 50–62, 2020.
- [20] S. Rajakumaran and J. Sasikala, "An efficient machine learning based tongue color analysis for automated disease diagnosis model," *International Journal of Advanced Research in Engineering and Technology*, vol. 11, no. 12, pp. 718–734, 2020.
- [21] J. Sasikala and S. Rajakumaran, "An automated tongue color image analysis for disease diagnosis and classification using deep learning techniques," *European Journal of Molecular & Clinical Medicine*, vol. 7, no. 7, pp. 4779–4796, 2021.