

Received June 20, 2021, accepted June 28, 2021, date of publication July 2, 2021, date of current version July 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3094226

Internet of Things and Synergic Deep Learning Based Biomedical Tongue Color Image Analysis for Disease Diagnosis and Classification

ROMANY F. MANSOUR¹, MAHA M. ALTHOBAITI², AND AMAL ADNAN ASHOUR³

¹Department of Mathematics, Faculty of Science, New Valley University, El-Kharga 72511, Egypt

²College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

³Department of Oral and Maxillofacial Surgery and Diagnostic Sciences, Faculty of Dentistry, Taif University, Taif 21944, Saudi Arabia

Corresponding author: Romany F. Mansour (romanyf@sci.nvu.edu.eg)

Present research work is supported by Taif University Researchers Supporting Project number (TURSP-2020/130), Taif University, Taif, Saudi Arabia.

ABSTRACT In recent times, internet of things (IoT) and wireless communication techniques become widely used in healthcare sector. Biomedical image processing is commonly employed to detect the existence of diseases using biomedical images. Tongue diagnosis is an efficient, non-invasive model to perform auxiliary diagnosis any time anywhere that is support the global necessity in the primary healthcare system. Conventionally, medical practitioners investigate the tongue features based on their expert's knowledge comes from experience. In order to eradicate the qualitative aspects, tongue images can be quantitatively examined, offering an effective disease diagnostic process in such a way that the physical harm of the patients can be minimized. Numerous tongue image analysis approaches exist in the literature, it is required to develop automated deep learning (DL) models to diagnose the diseases using tongue image analysis. In this view, this paper designs an automated IoT and synergic deep learning based tongue color image (ASDL-TCI) analysis model for disease diagnosis and classification. The proposed ASDL-TCI model operates on major stages namely data acquisition, pre-processing, feature extraction, classification, and parameter optimization. Primarily, the IoT devices are used to capture the human tongue images and transmitted to the cloud for further analysis. In addition, median filtering based image pre-processing and SDL based feature extraction techniques are employed. Moreover, deep neural network (DNN) based classifier is applied to determine the existence of the diseases. Lastly, enhanced black widow optimization (EBWO) based parameter tuning process takes place to enhance the diagnostic performance. For assessing the effectual performance of the ASDL-TCI model, a set of simulations take place on benchmark tongue images and examined the results under distinct dimensions. The simulation outcome verified the enhanced diagnostic performance of the ASDL-TCI model over the compared methods with the maximum precision, recall, and accuracy of 0.984, 0.973, and 0.983.

INDEX TERMS Biomedical image processing, tongue color image analysis, deep learning, deep neural network, parameter tuning, synergic deep learning, Internet of Things.

I. INTRODUCTION

Internet of Things (IoT) and deep learning (DL) have broad range of applications, healthcare is one among them. Due to the advancements in the internet, the classical schemes

The associate editor coordinating the review of this manuscript and approving it for publication was Chi-Hua Chen¹.

for patient services weakened and replaced by electronic healthcare systems. The utilization of IoT technologies provide healthcare professionals and patients the most modern medical device environment. At the same time, the worldwide demand for major healthcare support and the development of techniques enables point of care (POC) diagnosis. In spite of the current advancement in automatic

disease diagnosis equipment, the necessity of blood serum, reliability, non-professionals, detection time, accuracy, and need for additional confirmatory testing are the difficulties to conquer [1]. Therefore, tongue diagnosis, temperature, facial expressions, retinopathy, surface, and skin color can be some of the major variables for forthcoming smartphone based medical proficient system for attaining immediacy, simplicity, automated analysis, and non-invasiveness. Tongue diagnoses is an efficient non-invasive method for evaluating the state of a person's interior parts in oriental medication, i.e., Japanese traditional herbal medicine, traditional Korean medicine (TKM), traditional Chinese medicine (TCM), [2]. The human tongue comprises several features which are utilized for diagnosing disease, with colour features being significant. Conventionally, medicinal physicians will investigate these colour features depending upon extensive knowledge. But, subjectivity and ambiguity are often followed with their diagnoses outcome. For removing these qualitative factors, tongue colour analyses are objectively examined by its colour features that provide a novel method for diagnosing disease, one that reduces the physical injury caused to person (related to another medicinal investigation).

The diagnosis procedure is based on professional's view according to visual examination including form, substance, color, motion, and coating the tongue [3], [4]. Rather than tongue anomalous occurrence and disease, conventional tongue diagnoses are highly motivated for recognizing the disease. For instance, the tongue coating white greasy and yellow dense occurrence indicates hot and cold syndromes, correspondingly that are connected with health states like immune, inflammation, infection, endocrine, or stress disorders [5]–[8]. It contains 2 parallels, still interrelated syndromes in TCM. Removing the dependence on experience and subjective based analysis of tongue diagnoses might significantly raise the possibility of using tongue diagnoses in the world comprising Western medicine. Computerized tongue assessment involving color correction, light estimation, image analysis, geometry analysis, tongue segmentation, and so on can be an efficient device to diagnose disease that aims for addressing this concern [9]–[11].

Though several modules are depending upon approach of individual features that were presented and attained effective outcomes, this kind of technique uses lower level features [12]–[14]. Thus multi features are useful for detecting abnormal and normal tongue images. The study utilized multi features (like integration of shape, color, and texture) for identifying and matching tongue images. [15] introduced a coating separation phase beforehand feature extraction. [16] presented colour texture operators named PDSLBP for handling the tongue image corresponding problems. In [17], a TCoM is depending upon quantitative measurement which includes textural and chromatic features was presented for diagnosing appendicitis. In [18], multi labeled learning was utilized for classifying tongue images afterward extraction of texture and colour features. Indeed, the above-mentioned

techniques utilized lower level features whether it is individual /multi features that could not fully describe the features of tongue. It is required for integrating an architecture which could create whole features from tongue images. Therefore, higher level features have been required for CAD tongue analyses. One of the present publications was defined and employed DL modules for extracting higher level depictions to extensive vision analysis processes like object recognition, hand-written digit identification, and face recognition. But, there occurs slight or no study on CAD tongue image analyses by DL modules, where CAD professional systems with objectivity and unambiguity tongue analysis outcomes are utilized for facilitating the Western and TCM practice diagnosis outcomes.

This paper designs an automated IoT and synergic deep learning based tongue color image (ASDL-TCI) analysis model for disease diagnosis and classification. At the initial stage, the IoT devices are used to capture the human tongue images and transmitted to the cloud for further analysis. Besides, median filtering (MF) based image pre-processing is applied to enhance the quality of the tongue image. In addition, SDL based feature extraction model is used for deriving a helpful set of feature vectors. Moreover, deep neural network (DNN) based classifier is applied to determine the existence of the diseases. The DNN model is chosen due to the following features: It has high capacity to execute feature engineering on its own. Besides, it scan the data to search for features that correlate and combine them to enable faster learning without being explicitly told to do so. Finally, enhanced black widow optimization (EBWO) based parameter tuning process is utilized to boost the diagnostic results. In order to assess the proficient results of the ASDL-TCI model, a series of experimentations are performed on benchmark tongue images and examined the results under distinct dimensions.

II. LITERATURE SURVEY

Zhou *et al.* [19] presented Tongue Net that is a specific and faster automated tongue segmentation scheme. The U-net is used as segmentation backbone employing a smaller scale image dataset. Moreover, a morphological layer is presented in the last stage of the framework. Wen *et al.* [20] proposed a novel constitution recognition technique depending upon zero shot learning with knowledge of TCM. To enhance the efficiency, a novel zero-shot learning technique is presented by combining features and learning discriminant latent attributes that could resolve the imbalance challenge of constitution classifications. Kamarudin *et al.* [21] proposed a 2 phase tongue multicolor classification depending upon SVM that is decreased by this presented k mean clustering detectors and red colour range to diagnose accurate tongue colour. Initially, k-means clustering is utilized for clustering a tongue image to 4 clusters of deep red region, image background (black), transitional region, and red or light red region. Next red or light red tongue images are additionally categorized to red tongue/light red tongue depending upon

red colour range acquire in this study. Li *et al.* [22] proposed light weighted framework is depending upon encoder and decoder structures. The TIFE model is implemented for generating the feature with huge responsive regions with no sacrifice spatial solution. The context model is utilized for increasing the efficiency by combining multi scale contextual data. The decoder is implemented as simple still effective feature up sampling model for fusing distinct depth features and improve the segmentation outcomes together with tongue borders.

Li *et al.* [23] presented an effective and simple tongue image segmentation technique. Particularly, the presented technique initially extracts the first tongue body parts with the execution of image threshold on converted hue element in HSI color space. Later, this one image threshold leads to red element of an actual tongue image is selected in adaptive manner for finding the gap among the tongue body root and upper lip. Lastly, the early tongue body area is developed with the removal of fake object areas like upper lip for obtaining last tongue image segmentation outcome. Zhang *et al.* [24] improve a diagnosis technique of diabetes depending upon normalized tongue image by SVM. Tongue coating and body have been divided using chrominance threshold and division merging methods. By extracting colour and texture features of tongue image as input parameters, the diagnoses module of diabetic with SVM is trained. When enhancing the integration of SVM kernel variables and input parameters, the effects of an integration on the module have examined.

Fan *et al.* [25] examined various tongue features in person with diabetic, gastric symptom utilize gathered images by digital tongue imaging. In feature extraction phase, texture and 4 TCM tongue features have been detected: coating colour, slenderness, cracks, constitution colour, and plumpness. In classification phase, 2 distinct classification techniques have been utilized, SVM and RF, to categorize TCM and DM gastric diseases symptom. Wang *et al.* [26] presented an AI architecture by DCNN to recognize tooth- marked tongue. Initially, they created comparatively huge dataset with 1548 tongue images taken using various tools. Later, they utilize ResNet34 CNN framework for extracting features and performs classification. Meng *et al.* [27] proposed a new feature extraction architecture named CHDNet for extracting unbiased features and decrease human labor to diagnoses tongue in TCM. Prior CNN modules are mainly concentrated on learning convolution filters and adapting weights among themselves, however, these modules contain 2 main problems: redundancy and inadequate ability in managing unbalanced sample distribution. They present higher dispersal and local response normalization function for addressing the problem.

In [34], a tongue color image analysis model is presented based on the shape of a human tongue and analyzing it by the use of geometry features by means of computerized methods. [35] investigates the design of new automated tongue diagnostic system on mobile enabled environment. Some

other tongue color image analysis techniques are available in the literature [24], [36]–[39].

III. THE PROPOSED ASDL-TCI MODEL

The proposed ASDL-TCI model operates on major stages namely data acquisition, median filtering (MF) based pre-processing, SDL based feature extraction, DNN based classification, and EBWO based parameter optimization. Primarily, the IoT devices are used to capture the human tongue images. With the improvement of hardware, preprocessing requirements are reducing day by day. The commercially available Camera OV7670 module possesses camera resolution as high 656×488 pixels with 30 frames per second (fps). Then, they are transmitted to the cloud for further analysis. Followed by, the remaining modules are detailed in the succeeding subsections.

A. IMAGE PRE-PROCESSING

Firstly, MF technique is applied to pre-process the input images and enhance the image quality. The MF is a non-linear signal modeling approach which depends upon data. The noisy value of digital images is exchanged with the median values of neighborhood (mask). The pixel of mask is ordered from sequences of its gray levels, and median values of set are saved for exchanging the noisy value. The MF output is $g(x, y) = \text{med}\{f(x-i, y-j), j, j \in W\}$, where $f(x, y)$, $g(x, y)$ are the original and output images correspondingly, W refers the 2D neighbourhood. As the MF is non-linear filtering, their mathematically analyses are comparatively difficult to an image with an arbitrary noise [28]. In order to image with zero mean noise in normal distributions, the noise variance of MF is given as:

$$\sigma_{med}^2 = \frac{1}{4nf^2(\bar{n})} \approx \frac{\sigma_i^2}{n + \frac{\pi}{2} - 1} \cdot \frac{\pi}{2}. \quad (1)$$

where σ_i^2 represents the input noise power (variance), n refers the size of MF neighborhood, $f(\bar{n})$ implies the function of noise density. With the noise variance of average filter as:

$$\sigma_0^2 = \frac{1}{n} \sigma_i^2. \quad (2)$$

Relating of Eqs. (1) and (2), the MF properties dependent upon 2 things: the size of mask, and the distribution of noise. The MF efficiency of arbitrary noise reduction is more than the average filter efficiency, however to the impulse noise, particularly narrow pulses are further away and pulse width is lesser than $n/2$, the MF is most effectual. The MF efficiency is enhanced when the MF technique, related to the average filtering technique, appropriately resizes the neighborhood based on the noise density.

B. FEATURE EXTRACTION

During feature extraction process, the pre-processed image is fed into the SDL model for deriving a helpful group of feature vectors [40], [41].

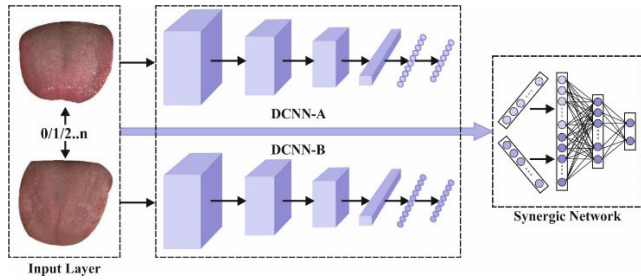


FIGURE 1. Structure of SDL network.

The SDL module extracts the beneficial group of features from the preprocessed image. It denotes SDL^k by three major components like input layer, C_k^2 synergic network (SN), and k DCNN component. All DCNN elements of network give an independent learning depiction from information with proper input data. The SN is consisting of a procedure with FC framework for ensures the input layer belongs to similar classes and provide remedial comments on the application of synergic error. Later, the SDL technique is separated into three sub-modules. Fig. 1 shows the structure of SDL network.

1) DCNN COMPONENT

Due to implied nature of significant residual network (ResNet), a *ResNet - 50* is utilized to initialize all DCNN elements $DCNN - a$ ($a = 1, 2, \dots, n$). Henceforth, it is stated that DCNN consist of GoogLeNet, VGGNet, and AlexNet, are matched in SDL technique as DCNN component. This element is trained by data series $X = \{x^{(1)}, x^{(2)}, \dots, x^{(M)}\}$ and final class label sequence $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(M)}\}$, aims to progress with group of parameter θ that endures cross entropy (CE) loss given by.

$$\log(\theta) = -\frac{1}{M} \left[\sum_{a=1}^M \sum_{b=1}^K 1 \{y^{(1)} = b\} \log \frac{e^{Z_b^{(a)}}}{\sum_{l=1}^K z_l^{(a)}} \right] \quad (3)$$

where n denotes number of classes, $Z^{(a)} = F(x^{(a)}, \theta)$ represents forward computing. An attained set of parameters for $DCNN - a$ denotes θ^a , and the parameters don't allocate huge $DCNN$ units.

2) SDL MODEL

The DCNN unit with synergic label of pairs is employed for the embedding, input layers, and FC learning. Consider (Z_A, Z_B) , as a data pair provided by the input for two DCNN components like $(DCNNa, DCNNb)$, correspondingly. The concluding outcome from successive FC layers in $DCNN$ embodies deep data feature which is learned by $DCNN$ recovered from forwarding compute, given by.

$$f_A = \mathcal{F}(Z_A, \theta^{(a)}) \quad (4)$$

$$f_B = \mathcal{F}(Z_B, \theta^{(a)}) \quad (5)$$

Later, the deep features from entire data are embedded by $f_{A \circ B}$, and specific results with synergic label as follows.

$$y_S(Z_A, Z_B) = \begin{cases} 1 & \text{if } y_A = y_B \\ 0 & \text{if } y_A \neq y_B \end{cases} \quad (6)$$

To conquer this problem, the percentage data pair in class is high and simpler to 0 views the synergic signal with alternative sigmoid layer and following binary CE loss is denoted by

$$l^S(\theta^S) = y_S \log \hat{y}_S + (1 - y_S) \log(1 - \hat{y}_S) \quad (7)$$

where θ^S indicates SN attribute, \hat{y}_S denotes SN forward computing. It authenticates when the input data pairs belong to similar class and provide remedial respond in the presence of synergic error.

3) TRAINING AND TESTING

When the training is made, the feature of DCNN components and SN is increased.

$$\begin{cases} \theta^{(a)}(z+1) = \theta^{(a)}(z) - \eta(z) \cdot \Delta^{(a)} \\ \theta^{S(a)}(z+1) = \theta^{S(a)}(z) - \eta(z) \cdot \Delta^{S(a,b)} \end{cases} \quad (8)$$

where $\eta(z)$ and $S(a, b)$ denotes rate of learning and SN between $DCNNa$ and $DCNNb$ is given by.

$$\Delta^{(a)} = \frac{\partial l^{(a)}(\theta^{(a)})}{\partial \theta^{(a)}} + \lambda \sum_{b=1, b \neq a}^n \frac{\partial l^{S(a)}(\theta^{S(a,b)})}{\partial \theta^{S(a,b)}} \quad (9)$$

$$\Delta^{S(a)} = \frac{\partial l^{S(a)}(\theta^{S(a,b)})}{\partial \theta^{S(a,b)}} \quad (10)$$

and λ indicates trade-off among sub-models of classification and synergic error. The association of trained procedure of SDL^2 model has been increased. In the application of trained SDL^k , the test data x is categorized by $DCNN$ unit $DCNN - u$ whereas it gives predictive vector $P^{(a)} = (p_1^{(a)}, p_2^{(a)}, \dots, p_k^{(a)})$, activation in resulting FC layer. The class label of test data is expected by,

$$y(Z) = \underset{v}{\operatorname{argmax}} \left\{ \sum_{u=1}^k p_1^{(u)}, \dots, \sum_{u=1}^k p_v^{(u)}, \dots, \sum_{u=1}^k p_K^{(u)} \right\} \quad (11)$$

C. IMAGE CLASSIFICATION

At the time of image classification, the derived feature vectors are applied to the DNN model to allocate the class labels. The DL technique implements in this study is a feed-forward ANN, training with stochastic gradient descent (SGD) utilizing backpropagation (BP). Now, several layers of hidden units are executed among the inputs as well as outputs of model. All hidden units, j , classically utilizes the logistic function β and the nearly connected hyperbolic tangent is also frequently utilized. Besides, some function with orderly derivative is employed for mapping their outputting y_j entire input in x_j :

$$y_i = \beta(x_j) = \frac{1}{1 + e^{-x_j}} \quad (12)$$

For the multiclass classifier of the tongue diagnosis problem, the output unit j changes their entire input, x_j , as to class probabilities, P_j , by utilizing a normalized exponential function called “softmax”:

$$P_i = \frac{\exp(X_j)}{\sum_h \exp(X_h)} \quad (13)$$

where h implies the index on every class. The DNNs are selectively training with BP derivatives of cost function which evaluate the discrepancy among the target and actual outputs caused to all trained cases [29]. If utilizing the softmax output function, the natural cost function C refers the CE among the target probability d and softmax outputs are P :

$$C = - \sum_i d_j \ln P_j \quad (14)$$

where the target probability, usually getting values of 1 or 0, is the supervised data given for training the DNN technique.

D. PARAMETER OPTIMIZATION

Finally, the parameter optimization of the DNN model takes place by the use of EBWO algorithm in such a way that the classification performance can be increased. The western black widow spider (*Latrodectus hesperus* or *L. hesperus*) is a poisonous spider species made in western Canada to southern Mexico. The venom, existing in female black widow, comprises a potent neurotoxin active towards an extent of animals. Additionally, the assumed venom is highly risky for humans provided that just single bite could lead to mortality. This spider feeds on insects like butterflies, cockroaches, and beetles; it is weave their web from trees and inhabit forests and swamps. Males, utilize sex pheromone to distinguish female mating position, are well-known to demonstrate no concern in mating with poorly and starving fed females, since females could exhibit cannibalistic performance. More detail regarding black widow spider’s performance is made in [30]. Fig. 2 illustrates the flowchart of BWO.

1) MOVEMENT

Based on moving approaches, the spider movement with the web have been demonstrated by spiral and linear, as defined by Eq. (11) whereas $\vec{x}_i(t+1)$ represents novel location of search agent, representing motion of the spider, and $\vec{x}_*(t)$ indicates optimum search agent made from prior iteration. The parameter m denotes float amount created arbitrarily in the interval [0.4, 0.9], r_1 represents arbitrary integer amount made from the interval in one to maximal size of search agent, and $\vec{x}_{r_1}(t)$ indicates r_1 th search agent chosen, with $i \neq r_1$. Lastly, β denotes arbitrary float amount made in the interval [-1.0, 1.0], and $\vec{x}_i(t)$ indicates present search agent.

$$\vec{x}_i(t+1) = \begin{cases} \vec{x}_*(t) - m\vec{x}_{r_1}(t), & \text{if } \text{rand}() \leq 0.3, \\ \vec{x}_*(t) - \cos(2\pi\beta)\vec{x}_i(t), & \text{in othercase,} \end{cases} \quad (15)$$

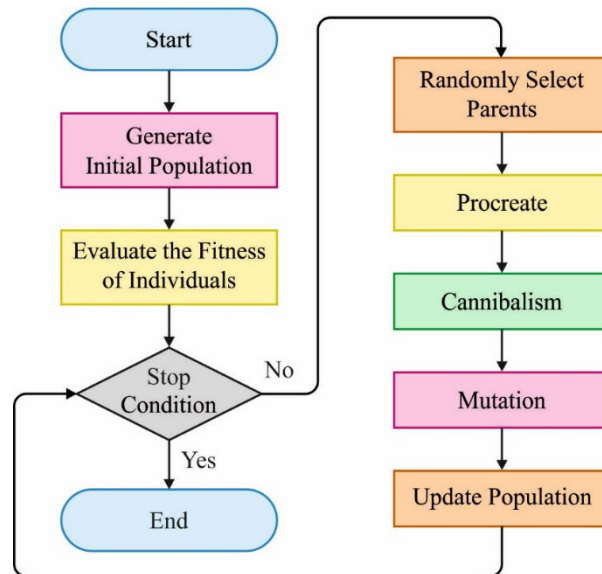


FIGURE 2. Flowchart of BWO algorithm.

2) PHEROMONES

Pheromones act as a significant part of courtship mating of *L. hesperus* spiders. The connection among the spider’s diet and modification in pheromone signals affect the quantity and quality of silk are demonstrated. Similarly, healthy female spiders generate additional silks compared to hungry females. Male spiders are highly responsible for sex pheromones from healthy females since they give advantage of having greater fertility, however, it mainly avoids the cost of risk mating attempt with probable hungry cannibal female. Specifically, male black widow spiders choose for avoiding cannibalism instead of search for extra fertile females. Sex pheromones only provide a perception to current feeding history of females, feasibly decreasing the cost for males exposing their selection in the region. Thus, female spiders with lower pheromone rates aren’t chosen by male spiders. In this study, black widow spider pheromones rate values are determined by:

$$pheromone(i) = \frac{fitness_{max} - fitness(i)}{fitness_{max} - fitness_{min}}, \quad (16)$$

where $fitness_{max}$ and $fitness_{min}$ indicates worst and optimum fitness values from present generation, correspondingly, where the $fitness(i)$ indicates present fitness values of i th search agent. The pheromone vectors, Eq. (16), comprises the standardized fitness from interval of zero and one. For lower pheromones rate values equivalent/lesser compared to 0.3, Eq. (17) is employed. The lower pheromone level in female spider represents hungry cannibal spider. Thus, they are existing, the assumed female spiders won’t be selected however it would be substituted to other ones:

$$\vec{x}_i(t) = \vec{x}_*(t) + \frac{1}{2} [\vec{x}_{r_1}(t) - (-1)^\sigma * \vec{x}_{r_2}(t)], \quad (17)$$

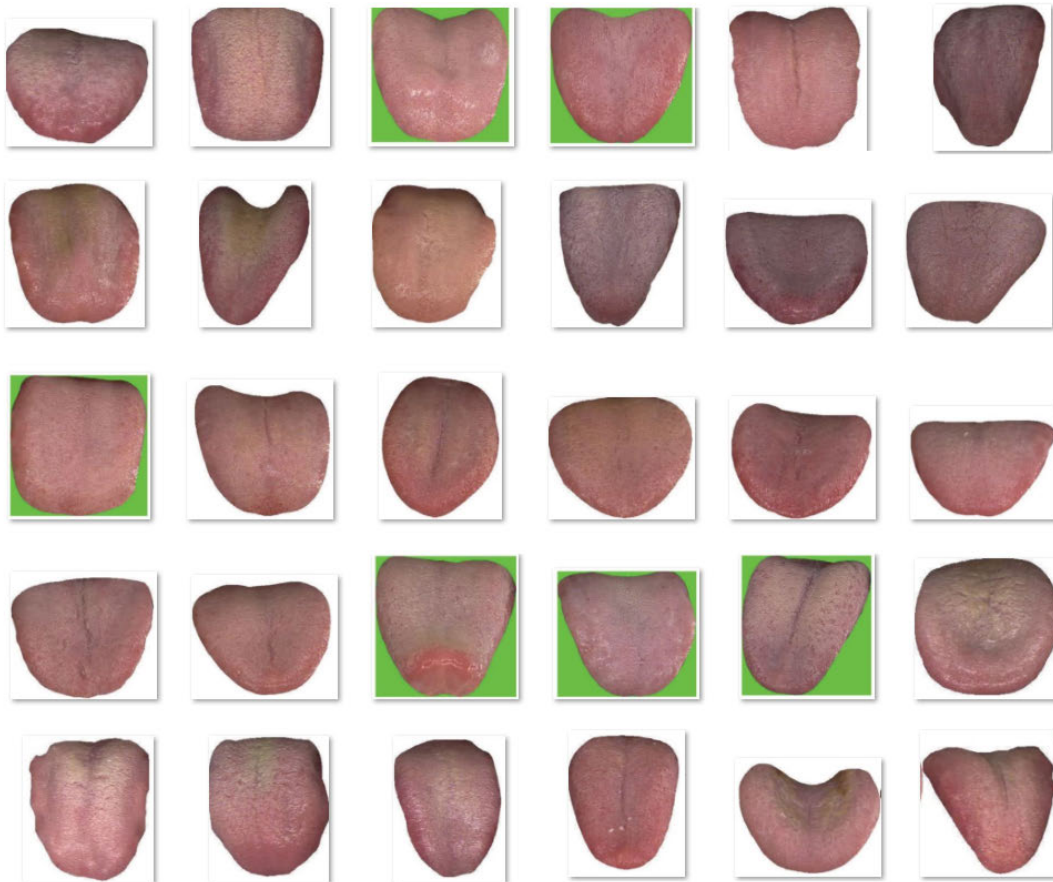


FIGURE 3. Sample disease affected tongue images.

where $\vec{x}_i(t)$ represents search agent (i.e. female spider) with lower pheromone rate would be upgraded. r_1 and r_2 indicates arbitrary integer number created in the interval from one to the maximum size of search agent (i.e. spiders), with $r_1 \neq r_2$, where $\vec{x}_{r_1}(t)$ and $\vec{x}_{r_2}(t)$ denotes r_1, r_2 th search agent chosen, $\vec{x}*(t)$ represents optimum search agent made from prior iteration, and σ indicates binary amount arbitrarily made, $\sigma \in \{0, 1\}$.

In order to avoid the local optimal problem of the BWO algorithm, the Levy flight concept is integrated to design the EBWO algorithm, which contains robust stochastics for improving the position update the black widow spider. The Levy flight boosts the jumping capability of black widow spiders and also extends the searching area of the swarm.

Levy's flight is an arbitrary stage defining the Levy distribution. It is a kind of random step model which is inspired by the flight path of levy. The step length is generally low; however, it irregularly occurs in large pulsation. The levy flight can be defined using Eq. (18):

$$Levy \sim u = t^{-\lambda}, \quad 1 < \lambda \leq 3 \quad (18)$$

The Levy random step by Mantegna can be defined using Eq. (19):

$$s = \frac{\mu}{|v|^{1/\beta}} \quad (19)$$

where $\beta = 1.5, \mu = N(0, \sigma_\mu^2)$ and $v = N(0, \sigma_v^2)$ denotes gamma function [31]. The variance of the parameter can be defined using Eq. (20):

$$\sigma_\mu = \left[\frac{\Gamma(1 + \beta) \times \sin(\pi \times \beta/2)}{\Gamma[(1 + \beta)/2] \times \beta \times 2^{(\beta-1)/2}} \right]^{1/\beta}, \quad \sigma_v = 1 \quad (20)$$

IV. EXPERIMENTAL VALIDATION

The proposed model is simulated utilizing PC i5-8600k processor, GeForce 1050Ti, 4GB RAM, 16GB OS Storage, and 250GB SSD File Storage. The simulation tool used is Python 3.6.5 tool along with some packages namely tensorflow, keras, numpy, pickle, matplotlib, sklearn, pillow, and opencv-python. The proposed model is experimented with using a benchmark dataset, comprising 936 images with 78 images under distinct 12 class labels. The information related to the dataset is provided in Table 1 [36]. Besides, 10 fold cross validation technique is employed. Fig. 3 illustrates some of the sample test images. The parameter setting of the DNN model is given as follows: mini batch size: 200, dropout: 0.5, number of hidden layers:3, number of hidden units: 1024, and activation function: softmax.

Fig. 4 illustrates some preprocessed images from the applied dataset. The figure depicted that the quality of the images gets increased in the applied test images.

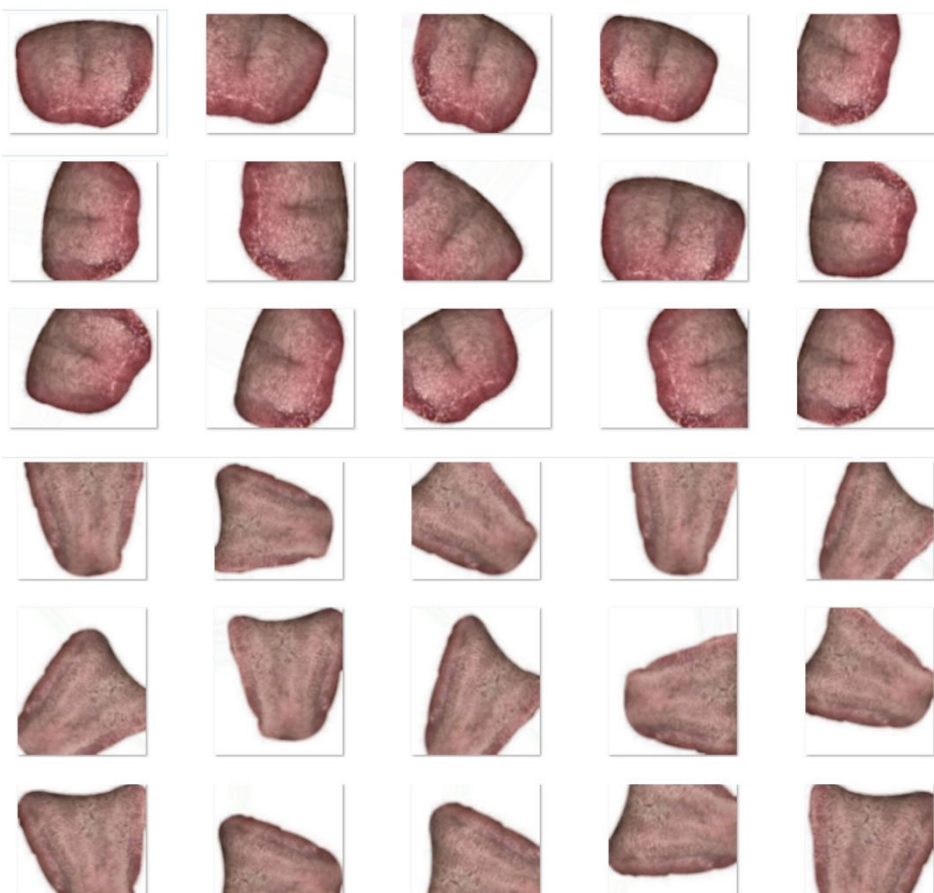


FIGURE 4. Sample preprocessed images.

TABLE 1. Dataset description.

| Diseases Name | No. of Samples |
|--|----------------|
| Chronic Kidney Disease | 78 |
| Nephritis | 78 |
| Verrucous Gastritis | 78 |
| Pneumonia | 78 |
| Nephritis Syndrome | 78 |
| Chronic Cerebral Circulation Insufficiency | 78 |
| Upper Respiratory Tract Infection | 78 |
| Erosive Gastritis | 78 |
| Coronary Heart Disease | 78 |
| Chronic Bronchitis | 78 |
| Mixed Hemorrhoid | 78 |
| Healthy | 78 |
| Total Images | 936 |

A brief classification results analysis of the ASDL-TCI model under different folds is provided in Table 2 and Figs. 5-6. From the resultant values, it is demonstrated that the ASDL-TCI model has obtained improved diagnostic outcomes under all the applied folds. For instance, under fold-1, the proposed ASDL-TCI technique has attained a prec. of 99.21%, rec. of 97.25%, accuracy of 99.58%, F1-score of 97.17%, and kappa of 97.13%. Similarly, under

TABLE 2. Performance analysis of proposed ASDL-TCI method in terms of different measures.

| No. of Folds | Precision | Recall | Accuracy | F1-Score | Kappa |
|----------------|--------------|--------------|--------------|--------------|--------------|
| Fold-1 | 99.21 | 97.25 | 99.58 | 97.17 | 97.13 |
| Fold-2 | 98.74 | 98.11 | 99.58 | 98.57 | 98.76 |
| Fold-3 | 97.59 | 96.47 | 96.62 | 99.76 | 98.18 |
| Fold-4 | 99.11 | 97.55 | 99.80 | 96.28 | 97.35 |
| Fold-5 | 97.51 | 96.50 | 96.34 | 96.62 | 96.60 |
| Fold-6 | 97.85 | 97.61 | 98.21 | 96.53 | 98.80 |
| Fold-7 | 99.80 | 96.30 | 96.24 | 98.47 | 98.10 |
| Fold-8 | 97.70 | 97.21 | 99.45 | 98.49 | 98.27 |
| Fold-9 | 97.36 | 97.70 | 98.73 | 96.53 | 96.58 |
| Fold-10 | 98.72 | 98.50 | 98.22 | 98.66 | 97.75 |
| Average | 98.36 | 97.32 | 98.28 | 97.71 | 97.75 |

fold-2, the proposed ASDL-TCI method has reached a prec. of 98.74%, rec. of 98.11%, accuracy of 99.58%, F1-score of 98.57%, and kappa of 98.76%. In line with, under fold-3, the proposed ASDL-TCI manner has attained a prec. of 97.59%, rec. of 96.47%, accuracy of 96.62%,

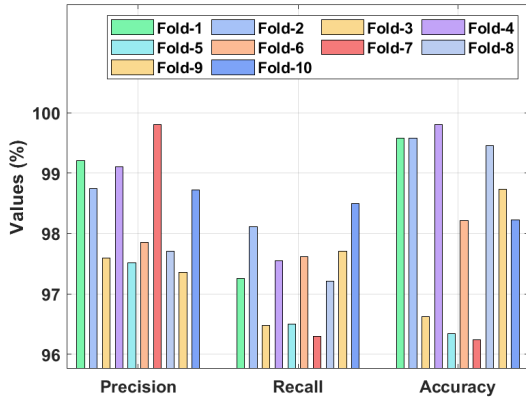


FIGURE 5. Results analysis of ASDL-TCI model under different folds-I.

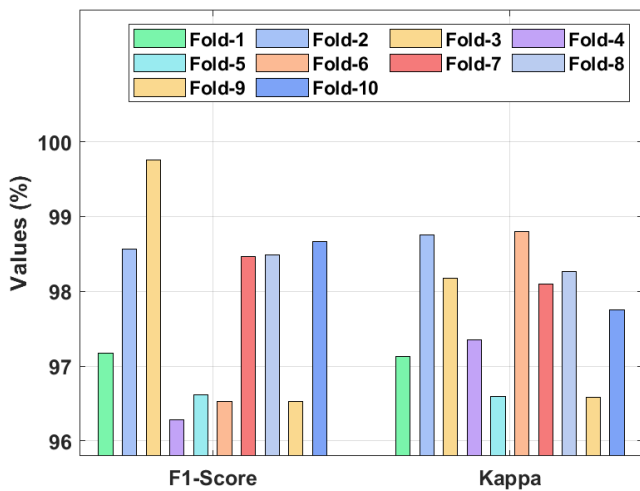


FIGURE 6. Results analysis of ASDL-TCI model under different folds-II.

F1-score of 99.76%, and kappa of 98.18%. Along with that, under fold-4, the proposed ASDL-TCI approach has reached a prec. of 99.11%, rec. of 97.55%, accuracy of 99.8%, F1-score of 96.28%, and kappa of 97.35%. Likewise, under fold-5, the proposed ASDL-TCI technique has obtained a prec. of 97.51%, rec. of 96.5%, accuracy of 96.34%, F1-score of 96.62%, and kappa of 96.60%. At the same time, under fold-6, the proposed ASDL-TCI algorithm has reached a prec. of 97.85%, rec. of 97.61%, accuracy of 98.21%, F1-score of 96.53%, and kappa of 98.8%. Concurrently, under fold-7, the proposed ASDL-TCI method has attained a prec. of 99.8%, rec. of 96.3%, accuracy of 96.24%, F1-score of 98.47%, and kappa of 98.1%. Eventually, under fold-8, the proposed ASDL-TCI technique has reached a prec. of 97.7%, rec. of 97.21%, accuracy of 99.45%, F1-score of 98.49%, and kappa of 98.1%. Besides, under fold-9, the proposed ASDL-TCI approach has attained a prec. of 97.36%, rec. of 97.7%, accuracy of 98.73%, F1-score of 96.53%, and kappa of 96.58%. Finally, under fold-10, the proposed ASDL-TCI methodology has obtained a prec. of 98.72%, rec. of 98.5%, accuracy of 98.22%, F1-score of 98.66%, and kappa of 97.75%.

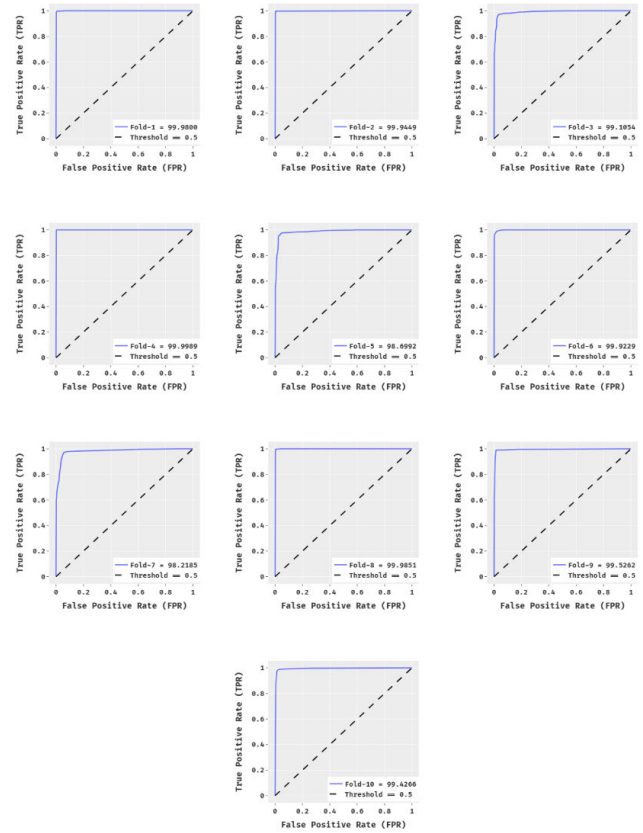


FIGURE 7. ROC analysis of ASDL-TCI model.

TABLE 3. Performance analysis of proposed ASDL-TCI method with recent methods in terms of different measures.

| Methods | Precision | Recall | Accuracy | F1-Score | Kappa |
|-----------|-----------|--------|----------|----------|-------|
| ASDL-TCI | 0.984 | 0.973 | 0.983 | 0.977 | 0.978 |
| SIFT-DTCA | 0.834 | 0.824 | 0.824 | 0.825 | 0.822 |
| SIFT-SVMA | 0.842 | 0.841 | 0.841 | 0.839 | 0.832 |
| HOG-DTCA | 0.898 | 0.889 | 0.889 | 0.891 | 0.878 |
| HOG-SVMA | 0.917 | 0.911 | 0.911 | 0.911 | 0.910 |
| VGG19-RF | 0.938 | 0.937 | 0.937 | 0.937 | 0.931 |
| VGG19-GNB | 0.945 | 0.925 | 0.925 | 0.928 | 0.922 |

Fig. 7 demonstrates the ROC analysis of the proposed ASDL-TCI model under varying folds. From these figures, it can be revealed that the ASDL-TCI algorithm has accomplished effective ROC values under all the applied folds.

Table 3 examines the comparative results analysis of the ASDL-TCI technique with other recent methods interms of different measures [32], [33].

Fig. 8 shows the result analysis of ASDL-TCI model interms of prec., rec., and accuracy. From the obtained values,

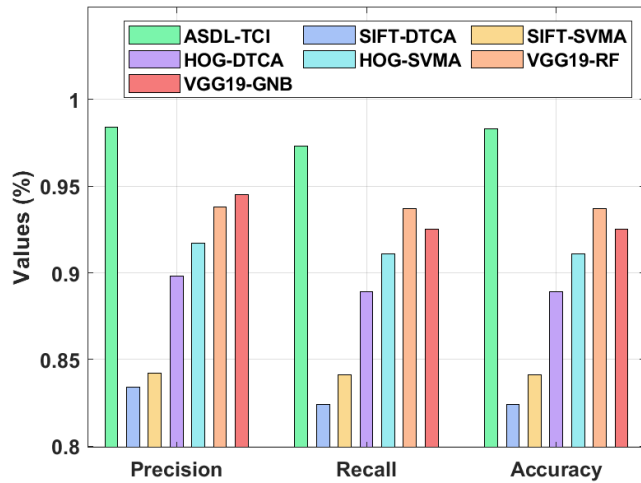


FIGURE 8. Result analysis of ASDL-TCI model with different measures.

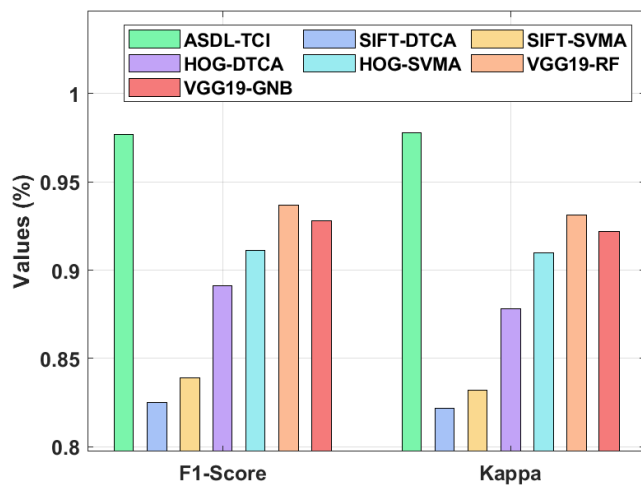


FIGURE 9. Result analysis of ASDL-TCI model interms of F1-score and kappa.

it can be clear that the SIFT-DTCA model has attained least performance with the lowest prec. of 0.834, rec. of 0.824, and accuracy of 0.824. Next, the SIFT-SVMA model has attained certainly increased outcomes with prec. of 0.842, rec. of 0.841, and accuracy of 0.841. At the same time, the HOG-DTCA technique has offered improved results over the other two methods with prec. of 0.898, rec. of 0.889, and accuracy of 0.889. Meanwhile, the HOG-SVMA model has resulted in a moderate result with prec. of 0.917, rec. of 0.911, and accuracy of 0.911. Eventually, the VGG19-RF model has gained somewhat enhanced performance with prec. of 0.938, rec. of 0.937, and accuracy of 0.937. Furthermore, the VGG19-GNB approach has portrayed competitive outcomes with prec. of 0.945, rec. of 0.925, and accuracy of 0.925. However, the ASDL-TCI model has gained maximum results over the other related manners with prec. of 0.984, rec. of 0.973, and accuracy of 0.983.

Fig. 9 shows the result analysis of ASDL-TCI model interms of F-measure and kappa. From the obtained values,

TABLE 4. Performance analysis of proposed ASDL-TCI method with traditional methods in terms of accuracy.

| Methods | Accuracy |
|-------------------|----------|
| ASDL-TCI | 0.983 |
| Geometry Features | 0.762 |
| Bayesian | 0.750 |
| K-NN | 0.734 |
| SVM | 0.765 |
| GA-SVM | 0.831 |
| VGG-SVM | 0.594 |
| GF+SRC | 0.792 |
| CADAE | 0.770 |

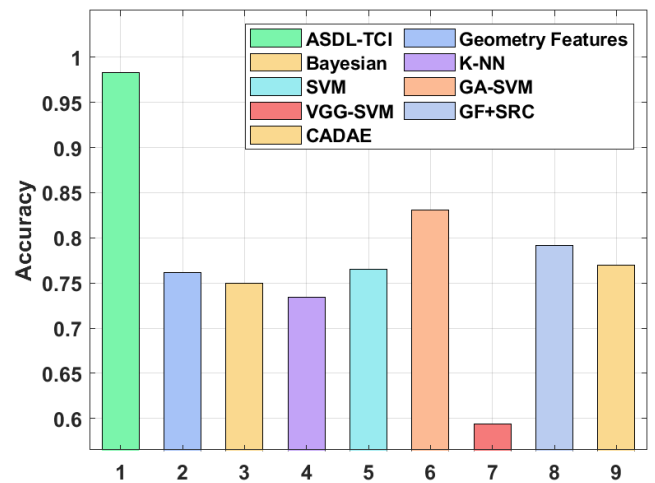


FIGURE 10. Accuracy analysis of ASDL-TCI model with existing techniques.

it can be obvious that the SIFT-DTCA model has attained least performance with the lowest F-measure of 0.825 and kappa of 0.822. Next, the SIFT-SVMA algorithm has attained certainly increased outcomes with F-measure of 0.839 and kappa of 0.832. Followed by, the HOG-DTCA manner has offered improved results over the other two methods with F-measure of 0.891 and kappa of 0.878. Meanwhile, the HOG-SVMA technique has resulted in a moderate result with F-measure of 0.911 and kappa of 0.910. Eventually, the VGG19-RF model has gained somewhat enhanced performance with F-measure of 0.937 and kappa of 0.931. Furthermore, the VGG19-GNB approach has showcased competitive outcomes with F-measure of 0.928 and kappa of 0.922. However, the ASDL-TCI methodology has gained superior outcomes over the other related approaches with F-measure of 0.977 and kappa of 0.978.

Table 4 and Fig. 10 examining the comparative outcomes analysis of the ASDL-TCI model with other existing methods [24], [34]–[39]. From the obtained values, it can be obvious that the VGG-SVM model has attained least performance with the lowest acc. of 0.594. Next, the K-NN model has attained certainly an increased outcome with an

acc. of 0.734. Likewise, the Bayesian technique has offered improved outcomes over the other two methods with an acc. of 0.75. Meanwhile, the Geometry features have resulted in a moderate result with an acc. of 0.762. Eventually, the SVM model has gained somewhat enhanced performance with an acc. of 0.765. Followed by, the CADA method has obtained moderate outcome with an acc. of 0.77. Moreover, the GF-SRC technique has accomplished reasonable outcomes with an acc. of 0.792. Also, the GA-SVM approach has showcased competitive outcomes with an acc. of 0.831. But, the ASDL-TCI model has gained superior results over the other related techniques with an acc. of 0.983.

From the above mentioned results, it is evident that the proposed ASDL-TCI model has obtained improved performance over the other methods due to the inclusion of SDL based feature extraction and EBWO based parameter optimization.

V. CONCLUSION

This paper has designed a new IoT enabled ASDL-TCI model for disease diagnosis and classification. The presented ASDL-TCI model aims to determine the existence of the diseases using tongue color images. The proposed ASDL-TCI model operates on major stages namely IoT based data acquisition, MF based pre-processing, SDL based feature extraction, DNN based classification, and EBWO based parameter optimization. The utilization of EBWO algorithm to tune the parameters of the DNN helps to considerably enhance the classification results to a maximum extent. For assessing the proficient results of ASDL-TCI model, a series of experimentations are carried out on benchmark tongue images and examined the results under distinct dimensions. The obtained results guaranteed the improved diagnostic performance of the ASDL-TCI model over the compared methods interms of different measures. As a part of future work, advanced DL architectures can be employed to further enhance the diagnostic performance.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

ETHICAL APPROVAL

This article does not contain any studies with human participants or animals performed by any of the authors.

CONSENT TO PARTICIPATE

Not applicable.

REFERENCES

- [1] M. H. Tania, K. T. Lwin, and M. A. Hossain, "Computational complexity of image processing algorithms for an intelligent mobile enabled tongue diagnosis scheme," in *Proc. 10th Int. Conf. Softw., Knowl., Inf. Manage. Appl. (SKIMA)*, Chengdu, China, 2016, pp. 29–36.
- [2] C. J. Jung, Y. J. Jeon, J. Y. Kim, and K. H. Kim, "Review on the current trends in tongue diagnosis systems," *Integrative Med. Res.*, vol. 1, no. 1, pp. 13–20, Dec. 2012.
- [3] S. K. Lakshmanaprabu, S. N. Mohanty, S. Krishnamoorthy, J. Uthayakumar, and K. Shankar, "Online clinical decision support system using optimal deep neural networks," *Appl. Soft Comput.*, vol. 81, Aug. 2019, Art. no. 105487.
- [4] U. Jayasankar, V. Thirumal, and D. Ponnuram, "A new objective image quality assessment metric: For color and grayscale images," *3D Res.*, vol. 9, no. 3, pp. 1–16, Sep. 2018.
- [5] R. Jahangir, Y. W. TEh, N. A. Memon, G. Mujtaba, M. Zareei, U. Ishtiaq, M. Z. Akhtar, and I. Ali, "Text-independent speaker identification through feature fusion and deep neural network," *IEEE Access*, vol. 8, pp. 32187–32202, 2020, doi: [10.1109/ACCESS.2020.2973541](https://doi.org/10.1109/ACCESS.2020.2973541).
- [6] M. S. Roomi, M. Imran, S. A. Shah, A. Almogren, I. Ali, and M. Zuair, "A novel de-ghosting image fusion technique for multi-exposure, multi-focus images using guided image filtering," *IEEE Access*, vol. 8, pp. 219656–219671, 2020, doi: [10.1109/ACCESS.2020.3043048](https://doi.org/10.1109/ACCESS.2020.3043048).
- [7] Aroosa, S. S. Ullah, S. Hussain, R. Alroobaea, and I. Ali, "Securing NDN-based Internet of health things through cost-effective signcryption scheme," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–13, Apr. 2021. [Online]. Available: <https://www.hindawi.com/journals/wcm/2021/5569365/>
- [8] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for Internet of Things (IoT) security," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 3, pp. 1646–1685, 3rd Quart., 2020.
- [9] L. W. Qin, M. Ahmad, I. Ali, R. Mumtaz, S. M. H. Zaidi, S. S. Alshamrani, M. A. Raza, and M. Tahir, "Precision measurement for industry 4.0 standards towards solid waste classification through enhanced imaging sensors and deep learning model," *Wireless Commun. Mobile Comput.*, vol. 2021, pp. 1–10, May 2021.
- [10] C.-H. Chen, F. Song, F.-J. Hwang, and L. Wu, "A probability density function generator based on neural networks," *Phys. A, Stat. Mech. Appl.*, vol. 541, Mar. 2020, Art. no. 123344.
- [11] C.-H. Chen, "A cell probe-based method for vehicle speed estimation," *IEICE Trans. Fundam. Electron., Commun. Comput. Sci.*, vol. E103.A, no. 1, pp. 265–267, Jan. 2020.
- [12] M. Pan, Y. Liu, J. Cao, Y. Li, C. Li, and C.-H. Chen, "Visual recognition based on deep learning for navigation mark classification," *IEEE Access*, vol. 8, pp. 32767–32775, Feb. 2020.
- [13] C.-H. Chen, F.-J. Hwang, and H.-Y. Kung, "Travel time prediction system based on data clustering for waste collection vehicles," *IEICE Trans. Inf. Syst.*, vol. E102.D, no. 7, pp. 1374–1383, Jul. 2019.
- [14] C.-H. Chen, "An arrival time prediction method for bus system," *IEEE Internet Things J.*, vol. 5, no. 5, pp. 4231–4232, Oct. 2018.
- [15] R. Kanawong, T. Obafemi-Ajayi, J. Yu, D. Xu, S. Li, and Y. Duan, "ZHENG classification in traditional Chinese medicine based on modified specular-free tongue images," in *Proc. IEEE Int. Conf. Bioinf. Biomed. Workshops (BIBMW)*, Philadelphia, PA, USA, Oct. 2012, pp. 288–294.
- [16] Z. Guo, "Tongue image matching using color and texture," in *Proc. Int. Conf. Med. Biometrics (ICMB)*, Hong Kong, 2008, pp. 273–281.
- [17] B. Pang, D. Zhang, and K. Wang, "Tongue image analysis for appendicitis diagnosis," *Inf. Sci.*, vol. 175, no. 3, pp. 160–176, Oct. 2005.
- [18] X. Zhang, J. Zhang, G. Hu, and Y. Wang, "Preliminary study of tongue image classification based on multi-label learning," in *Advanced Intelligent Computing Theories and Applications (Lecture Notes in Computer Science)*, vol. 9227. New York, NY, USA: Springer, 2015, pp. 208–220.
- [19] 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," *Appl. Sci.*, vol. 9, no. 15, p. 3128, Aug. 2019.
- [20] G. Wen, J. Ma, Y. Hu, H. Li, and L. Jiang, "Grouping attributes zero-shot learning for tongue constitution recognition," *Artif. Intell. Med.*, vol. 109, Sep. 2020, Art. no. 101951.
- [21] N. D. Kamarudin, C. Y. Ooi, T. Kawanabe, H. Odaguchi, and F. Kobayashi, "A fast SVM-based tongue's colour classification aided by k-means clustering identifiers and colour attributes as computer-assisted tool for tongue diagnosis," *J. Healthcare Eng.*, vol. 2017, pp. 1–13, Jan. 2017.

- [22] X. Li, D. Yang, Y. Wang, S. Yang, L. Qi, F. Li, Z. Gan, and W. Zhang, "Automatic tongue image segmentation for real-time remote diagnosis," in *Proc. IEEE Int. Conf. Bioinf. Biomed. (BIBM)*, Nov. 2019, pp. 409–414.
- [23] Z. Li, Z. Yu, W. Liu, Y. Xu, D. Zhang, and Y. Cheng, "Tongue image segmentation via color decomposition and thresholding," *Concurrency Comput., Pract. Exp.*, vol. 31, no. 23, Dec. 2019, Art. no. e4662.
- [24] J. Zhang, J. Xu, X. Hu, Q. Chen, L. Tu, J. Huang, and J. Cui, "Diagnostic method of diabetes based on support vector machine and tongue images," *BioMed Res. Int.*, vol. 2017, pp. 1–9, Jan. 2017.
- [25] S. Fan, B. Chen, X. Zhang, X. Hu, L. Bao, X. Yang, Z. Liu, and Y. Yu, "Machine learning algorithms in classifying TCM tongue features in diabetes mellitus and symptoms of gastric disease," *Eur. J. Integrative Med.*, vol. 43, Apr. 2021, Art. no. 101288.
- [26] X. Wang, J. Liu, C. Wu, J. Liu, Q. Li, Y. Chen, X. Wang, X. Chen, X. Pang, B. Chang, J. Lin, S. Zhao, Z. Li, Q. Deng, Y. Lu, D. Zhao, and J. Chen, "Artificial intelligence in tongue diagnosis: Using deep convolutional neural network for recognizing unhealthy tongue with tooth-mark," *Comput. Struct. Biotechnol. J.*, vol. 18, pp. 973–980, Jan. 2020.
- [27] D. Meng, G. Cao, Y. Duan, M. Zhu, L. Tu, D. Xu, and J. Xu, "Tongue images classification based on constrained high dispersal network," *Evidence-Based Complementary Alternative Med.*, vol. 2017, pp. 1–12, Mar. 2017.
- [28] Y. Zhu and C. Huang, "An improved median filtering algorithm for image noise reduction," *Phys. Procedia*, vol. 25, pp. 609–616, Jan. 2012.
- [29] M. S. Roodposhti, J. Aryal, A. Lucieer, and B. Bryan, "Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest," *Entropy*, vol. 21, no. 1, p. 78, Jan. 2019.
- [30] A. F. Peña-Delgado, H. Peraza-Vázquez, J. H. Almazán-Covarrubias, N. T. Cruz, P. M. García-Vite, A. B. Morales-Cepeda, and J. M. Ramírez-Arredondo, "A novel bio-inspired algorithm applied to selective harmonic elimination in a three-phase eleven-level inverter," *Math. Problems Eng.*, vol. 2020, pp. 1–10, Dec. 2020.
- [31] Z. Xing, "An improved emperor penguin optimization based multilevel thresholding for color image segmentation," *Knowl.-Based Syst.*, vol. 194, Apr. 2020, Art. no. 105570.
- [32] S. Rajakumaran and J. Sasikala, "An efficient machine learning based tongue color analysis for automated disease diagnosis model," *Int. J. Adv. Res. Eng. Technol.*, vol. 11, no. 12, pp. 718–734, Dec. 2020, doi: 10.34218/IJARET.11.12.2020.073.
- [33] J. Sasikala and S. Rajakumaran, "An automated tongue color image analysis for disease diagnosis and classification using deep learning techniques," *Eur. J. Mol. Clin. Med.*, vol. 7, no. 7, pp. 4779–4796, 2021.
- [34] B. Zhang and H. Zhang, "Significant geometry features in tongue image analysis," *Evidence-Based Complementary Alternative Med.*, vol. 2015, pp. 1–8, Jul. 2015.
- [35] M. H. Tania, K. Lwin, and M. A. Hossain, "Advances in automated tongue diagnosis techniques," *Integrative Med. Res.*, vol. 8, no. 1, pp. 42–56, Mar. 2019.
- [36] B. Zhang, X. Wang, J. You, and D. Zhang, "Tongue color analysis for medical application," *Evidence-Based Complementary Alternative Med.*, vol. 2013, pp. 1–11, Mar. 2013.
- [37] J. Ma, G. Wen, Y. Hu, T. Chang, H. Zeng, L. Jiang, and J. Qin, "Tongue image constitution recognition based on complexity perception method," 2018, *arXiv:1803.00219*. [Online]. Available: <http://arxiv.org/abs/1803.00219>
- [38] H. Zhang and B. Zhang, "Disease detection using tongue geometry features with sparse representation classifier," in *Proc. Int. Conf. Med. Biometrics*, May 2014, pp. 102–107.
- [39] S. Balu and V. Jeyakumar, "A study on feature extraction and classification for tongue disease diagnosis," in *Intelligence in Big Data Technologies—Beyond the Hype*. Singapore: Springer, 2020, pp. 341–351.
- [40] C. S. S. Anupama, M. Sivaram, E. L. Lydia, D. Gupta, and K. Shankar, "Synergic deep learning model-based automated detection and classification of brain intracranial hemorrhage images in wearable networks," *Pers. Ubiquitous Comput.*, pp. 1–10, Nov. 2020, doi: 10.1007/s00779-020-01492-2.
- [41] K. Shankar, A. R. W. Sait, D. Gupta, S. K. Lakshmanaprabu, A. Khanna, and H. M. Pandey, "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," *Pattern Recognit. Lett.*, vol. 133, pp. 210–216, May 2020, doi: 10.1016/j.patrec.2020.02.026.

• • •