IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 23 September 2023, accepted 11 October 2023, date of publication 16 October 2023, date of current version 23 October 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3324880

RESEARCH ARTICLE

Laryngeal Cancer Detection and Classification Using Aquila Optimization Algorithm With Deep Learning on Throat Region Images

FADWA ALROWAIS[1], KHALID MAHMOOD[2], SAUD S. ALOTAIBI[3], MANAR AHMED HAMZA[4], RADWA MARZOUK[5], AND ABDULLAH MOHAMED[6]

[1]Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[2]Department of Information Systems, College of Science and Art at Mahayil, King Khalid University, Abha 61421, Saudi Arabia
[3]Department of Information Systems, College of Computing and Information System, Umm Al-Qura University, Mecca 21421, Saudi Arabia
[4]Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Al-Kharj 11671, Saudi Arabia
[5]Department of Information Systems, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University, Riyadh 11671, Saudi Arabia
[6]Research Centre, Future University in Egypt, New Cairo 11845, Egypt

Corresponding author: Khalid Mahmood (kasgr@kku.edu.sa)

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/112/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R77), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

ABSTRACT Laryngeal cancer detection on throat area images is a vital application of medical image diagnosis and computer vision (CV) in the healthcare domain. It contains the analysis and detection of cancerous or abnormal tissues from the larynx, an essential part of the respiratory and vocal systems. Several machine learning (ML) and deep learning (DL) systems are executed for classifying the extraction features as both cancerous and healthy tissue. Convolutional Neural Networks (CNNs) and recurrent neural networks (RNNs) have shown promise in this context. With this motivation, this study designs a new Laryngeal Cancer Detection and Classification using the Aquila Optimization Algorithm with Deep Learning (LCDC-AOADL) technique on neck region images. The purpose of the LCDC-AOADL technique is to examine the histopathological images for the recognition and classification of Laryngeal Cancer. In the presented LCDC-AOADL technique, the Inceptionv3 model is used for the feature extraction process. Besides, the LCDC-AOADL technique employed a deep belief network (DBN) model for the identification and classification of LC. Moreover, the AOA is utilized for the hyperparameter tuning of the DBN model which results in improved detection rate. The simulation analysis of the LCDC-AOADL method is validated on the benchmark Laryngeal dataset. The experimental results pointed out the enhanced detection results of the LCDC-AOADL technique over other recent approaches with a maximum accuracy of 96.02%, precision of 92.10%, recall of 91.87%, and F-score of 91.86%.

INDEX TERMS Computer-aided diagnosis, ENT, Laryngeal cancer, deep learning, Aquila optimizer.

I. INTRODUCTION

In recent years, Laryngeal cancer (LC) has become a significant risk in human life in the advanced era around the world. It is found to be the most serious malignancy within the neck and head part in patients [1]. The condition of LC is rapid expansion between the people each year globally. The clinical treatment substitutes and prognoses depend on the stages of cancer in the case of diagnoses of LC [2]. The early stage of LC and pre-malignant is related to a greater degree of laryngeal protection together with a clear diagnosis in actual time, whereas extreme levels of LC require a multimodal diagnostic strategy that leads to poor life quality and crucial toxicities [3]. As regards an enhanced diagnostic strategy, current research determines a high rate of frequency with total survival of 34 to 62% [4]. The initial identification and accurate diagnosis of cancer are crucial for achieving great medical results [5]. To identify Laryngeal Squamous

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

Cell Carcinoma (LSCC), physicians examine the throat with endoscopy or laryngoscope [6].

When suggested by doctors, a biopsy will take place and be observed under a proper microscope by a histopathologist for a final diagnosis [7]. These processes are also reliant on the histopathologist's expertness and human knowledge of morphological characteristics and histological textures [8]. The histopathologist is required to consider several characteristics and organs, namely nucleolus patterns, mitosis, special cells, the shape of the nucleus, the epithelium, the cytoplasm and the distance among cells [9]. An accurate and instantaneous diagnosis is required in the surgery but it can able to take a lot of time for pathologists to identify specific features to make the right decision. In addition, to gain expertise, pathologists need long-time training, and in most parts of the world, experienced histopathologists are insufficient [10]. Moreover, the Convolutional Neural Networks (CNNs) model can obtain the capability to make diagnoses in a very short period than medical specialists. Therefore, this technique may be supported by the diagnoses of LSCC in the future [11]. Deep learning (DL) is referred to as a Machine Learning (ML) approach, which is based on a Neural Network (NN) method with various levels of data format [12]. The CNNs are feed-forward NNs with convolution calculation and deep structure. A technique can handle detection and classification problems. The comparison with the traditional image processing methods, CNN has strong potential for analysis and feature extraction [13]. Currently, Artificial Intelligence (AI) based on the deep CNN (DCNN) technique is used for Magnetic Resonance Imaging (MRI), pathology, congenital cataracts, skin cancer classification and diabetic retinopathy diagnosis. By using these cutting-edge DL approaches, the AI model can directly offer an accurate diagnosis based on image data that can support the earlier identification of diseases and improve the patient's survival rates [14].

This study designs a new Laryngeal Cancer Detection and Classification using the Aquila Optimization Algorithm with Deep Learning (LCDC-AOADL) technique on neck region images. The major intention of the LCDC-AOADL technique is to investigate the histopathological images for the identification of LC. In the presented LCDC-AOADL technique, the Inceptionv3 model is used for the feature extraction process. Besides, the LCDC-AOADL technique employed a deep belief network (DBN) model for the identification and classification of LC. Moreover, the AOA is utilized for the hyperparameter tuning of the DBN model which results in improved detection rate. The simulation analysis of the LCDC-AOADL method is validated on the benchmark dataset

II. RELATED WORKS

The authors in [15] proposed a new and improved DL-based Mask RCNN model for the detection of LC and its symptoms by using different CT images and image datasets in the real world. Moreover, the proposed method can able to capture and detect minor malignancies of the larynx part in a faster and more significant way in the practical screening of the patient and save a lot of time for the doctors, which allows for more screening of patients daily. In [16], the study analyzed the DL method for in-vivo hyperspectral LC diagnosis. For these purposes, we evaluate and design a CNN model with 2D or 3D spatio-spectral convolution integrated with the advanced DenseNet structure. For assessment, an in-vivo dataset was used with HSI of the oropharynx or oral cavity.

Sachane and Patil [17] presented an effective method for the classification of LC using the presented Adaptive Spotted Hyena Optimizer-based Deep Quantum Neural Network (ASHO-based Deep QNN). At last, classification can be obtained by the Deep QNN, where the presented method is utilized for tuning the network classifiers. The ASHO is proposed by inheriting the advantages of the Adaptive concept with Spotted Hyena Optimizer (SHO). Tayade [18] presented a CAD powered by the DL algorithm. To accomplish this, a new Multiple Instance Learning (MIL) technique was developed that categorizes healthy and unhealthy/cancerous tissues. Furthermore, the classical CNN and transfer learning DenseNet121 models are combined on the datasets for the assessment and better comparison. The model was estimated by the standard metrics which are specific to the biomedical field.

An LC classification network (LPCANet) based on CNN and attention modules was developed by Zhou et al. [19]. Firstly, the original HSI are successively cropped into patches. Next, the patches are inputted into ResNet50 for extracting the local features. Later, channel and position attention modules are simultaneously added to capture the spatial and channel dependency, correspondingly. Both models generate fusion feature maps for improving the classification performance of the network and enhancing the feature representation. Esmaeili et al. [20] proposed a DCNN model for the automated classification of CE-NBI images into benign and malignant classes with minimal human interference. A pretrained ResNet50 architecture fused with the cut-off layer method was chosen as a DCNN structure. Maskeliūnas et al. [21] introduced a popular CNN model for image classification to analyse voice audio signals. The method takes input from the DNN model as input from the Mel-frequency spectrogram (MFCC). Kwon et al. [22] aimed to enhance the classification performance by comparing the outcomes attained by using DT ensemble learning which is a method for increasing the classification performance for comparatively smaller datasets, with the outcomes attained using the CNN model for the diagnoses of the glottic tumour.

An important research gap in the field of Laryngeal Cancer Detection lies in the effectual tuning of hyperparameters in the recognition approaches. While DL and optimizer systems depict the ability to enhance detection accuracy, the intricacies of hyperparameter settings frequently remain underexplored. The better configuration of hyperparameters namely rate of learning, layer architectures, and batch sizes

has a major impact on the performance and generalization of recognition approaches. So, there is a compelling require for research that concentrates on emerging specialized hyperparameter tuning approaches tailored to laryngeal cancer detection approaches. These approaches take into account the unique features of the medical images and the particular requirements of the task, ultimately aiming to improve recognition accuracy and robustness in a medical context. Addressing this research gap is vital for enhancing the capability of automated laryngeal cancer detection methods and making sure it is reliable in real-world medical applications.

III. THE PROPOSED MODEL

In this study, we have developed the LCDC-AOADL technique. The main purpose of the LCDC-AOADL technique is to examine the histopathological images for the recognition and classification of LC. In the presented LCDC-AOADL technique, three stages of operations are involved such as Inceptionv3-based feature extraction, DBN classification, and AOA-based hyperparameter tuning. Fig. 1 demonstrates the workflow of the LCDC-AOADL algorithm.

A. FEATURE EXTRACTION: INCEPTIONV3 MODEL

Primarily, the Inceptionv3 model is used for the generation of feature vectors. The InceptionV3 is selected in the initial feature extraction model, due to its capability to extract higher-level features with different variations of the 277 filters [23]. Furthermore, without compromising the model efficacy, the architecture of InceptionV3 presents the features of reduction dimension by applying two (3×3) layers rather than a single (5×5) convolutional layer because with a similar amount of filters $a(5 \times 5)$ convolutional layer is $25/9 = 2.78$ times computational expensive than $a(3 \times 3)$ convolution layers. Consequently, using two (3×3) convolution layers, an overall gain of 28% is possible. Furthermore, large convolution was factorized as shown in Fig. 2, the consequence of asymmetric convolution leads to time reduction. Rather than (3×3) convolution layers, (1×3) and $a\,(3 \times 1)$ convolution layers decrease the computational burden. For the (3×3) convolution layers, the overall amount of parameters is fixed to 9. At the same time, a (1×3) and $a\,(1 \times 3)$ convolution layers lead to an overall amount of parameters of $(3 \times 1) + (1 \times 3) = 6$ which also decreases the overall amount of parameters by about 33%. Concerning the computation cost, the InceptionV3 model further applied an effective grid reduction method to overcome the issues of the pre-trained model. The typical pretrained method exploits the convolution layer along with max-pooling which is too greedy. After the convolution layer, the usage of max−pooling can be very expensive. To overcome these challenges, InceptionV3 implemented robust grid reduction using pooling and convolution operations followed by the last concatenation.

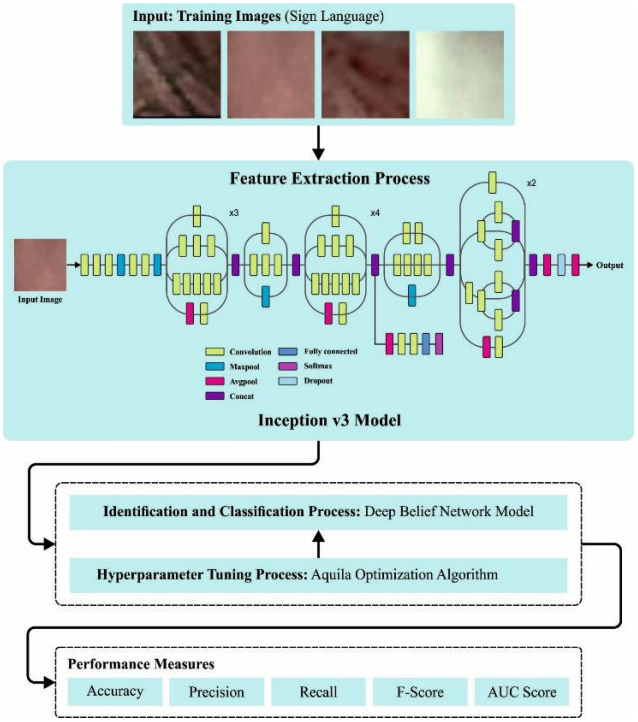

FIGURE 1. Workflow of LCDC-AOADL approach.

B. IMAGE CLASSIFICATION: DBN MODEL

The DBN model is used for the recognition and classification of LC. DBN is a kind of generalization module that uses a processing layer for capturing complex abstractions and structures in data [24]. It includes a set of trained stacked RBMs on top of one another. In an unsupervised way, RBM in a DBN is trained from which the training process initiates with an unsupervised phase. Fig. 2 depicts the framework of the DBN technique.

Typically, the ''visible'' and the ''hidden'' layers are two processing layers in RBM. The visible layer (VL) is the noticeable features or entities of the data, whereas the hidden layer (HL) captures hidden or latent representation. The training of the DBN model includes training RBM in a layer-wise manner. The observable layer as input trained the initial RBM, and the HL activation becomes the VL for the following RBM. This procedure is repeated until RBM is trained. This layer-wise pretraining model assists in fixing a problem that occurs once the network is originally set up without trained, random connection weight. The generative stochastic neural network was learned from probability models using unsupervised learning models. The RBM model comprises two different processing steps, labelled ''visible'' and ''hidden''. The units within the similar layer have no connection to each other, however, the constructed and reconstructed processes are provided with the interconnection of this layer. An enormous amount of hidden entities $(h_1, h_2, \ldots h_j)$ hidden nodes in the network have binary values, obtain data from the visible nodes and a considerable amount of observable

entities $(v_1, v_2, \ldots v_i)$ make up the VL (v) of the networks, which are trained on unlabelled pattern structure fed into it and are capable of reconstructing the patterns (h).

Each obvious node talks to each obvious node as a symmetric bi-directional matrix of weight (S_{ij}), along with the biases (b) and (a) that are already there.

$$R(v,h) = \sum_{i \epsilon vis} \frac{(v_i - b_i)^2}{2\lambda_i^2} - \sum_{j \epsilon hid} a_j h_j - \sum_{ij} \frac{v_i}{\lambda_i} h_j S_{ij} \quad (1)$$

In Eq. (1), λ shows the distribution of Gaussian noise at the i^{th}visible dimension.

The learning model becomes increasingly difficult if the hidden and visible components are Gaussian. The standard deviation of the noise level measures the coefficient of the quadratic ''containment'' term that keeps the activity within moderate limits:

$$R(v,h) = \sum_{i \epsilon vis} \frac{(v_i - b_i)^2}{2\lambda_i^2} + \sum_{j \epsilon hid} \frac{(h_j - a_j)^2}{2\lambda_i^2} - \sum_{ij} \frac{v_i h_j}{\lambda_i \lambda_j} S_{ij} \quad (2)$$

During the training dataset, data is used for guesses regarding the probability of the concealed unit and to graphically signify those predictions.

$$M(h_j = 1) = l\left(a_j + \sum v_i w_{ij}\right) \quad (3)$$

With the sample of h, invisible parameter v' was reconstructed at the exposed level. Then, a fresh set of h' latent activation was collected as follows:

$$M(v_i = 1) = l\left(b_i + \sum h_j s_{ij}\right) \quad (4)$$

The results of multiplying v' with h' from the exterior is the key to these solutions.

$$\Delta Sij = \eta((v_i hj)_{data} - (v_i' \cdot 'j)_{model}) \quad (5)$$

In Eq. (5), learner speed η is considered. Make some modifications to b_i and h_j in Eqs. (6) & (7), where $(\cdot)$ denotes the logistic activation function, correspondingly.

$$b = b + l(v - v') \quad (6)$$

$$a = a + l(h - h') \quad (7)$$

$$\emptyset(x) = \frac{1}{1 + e^{-x}} \quad (8)$$

Eventually, a logistic activation function is utilized in all the nodes of processing which take the input value (x) and employs the logistic conversion to squeeze the output within [0, 1].

C. HYPERPARAMETER TUNING: AOA

In this work, the AOA is used for the hyperparameter tuning of the DBN model. AOA is a new intelligent optimization technique extensively used to resolve real-time optimization problems [25]. This technique has two different strategies

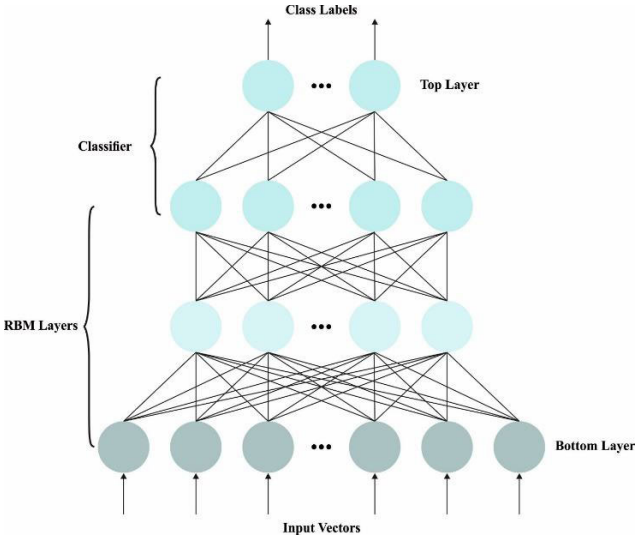

FIGURE 2. Architecture of DBN.

namely, exploration and development strategies. The AOA has obvious advantages compared to other metaheuristic approaches. The algorithm was based on the four different swarm behaviour of Northern Aquila birds at the time of predation. Firstly, using Eq.(9), the position matrix of the population can be randomly initialized:

$$X_{ij} = rand \times (UB_j - LB_j) + LB_j, \quad i = 1, 2, \ldots, Dim \quad (9)$$

where LB_j and UB_j indicate the j^{th} lower and upper constraints for the search problem and *rand* refers to the random vector.

1) The first stage is expanding the exploration while the Aquila is chasing the bird in the air. The bird uses vertical glide height for expanding the searching space and it can be mathematically expressed by using Eq. (10) and (11):

$$X_1(t+1) = X_{best}(t) \times \left(1 - \frac{t}{T}\right) + (X_M(t) - X_{best}(t) \times rand) \quad (10)$$

$$X_M(t) = \frac{1}{N} \sum_{i=1}^{N} X_j(t), \quad \forall j = 1, 2, \ldots, Dim \quad (11)$$

where $X_{best}(t)$ shows the optimum individual location attained by the algorithm at t^{hetth} iteration, $X(t)$ and $X(t+1)$ denote the individual position of the AO at the t^{th} and $t+1$ iterations, correspondingly.

2) The next phase is Downsizing while the Aquila find the prey from high in the air. It selects to spiral over the prey, ready for landing, and attack. This can be mathematically expressed by using Eq. (12):

$$X_2(t+1) = X_{best}(t) \times Levy(D) + X_R(t) + (y - x) \times rand \quad (12)$$

where Levy (D) denotes the Levy strategy, s shows the constant value set as 0.01, and u and v indicate randomly

generated integers within [0, 1].

$$Levy(D) = s \times \frac{u \times \sigma}{|v|^{\frac{1}{p}}} \tag{13}$$

$$\sigma = \left[\frac{\tau\,(1+\beta) \times sine\left(\frac{\pi\beta}{2}\right)}{\tau\left(\frac{1+\beta}{2}\right) \times \beta \times 2\left(\frac{\beta-1}{2}\right)} \right] \tag{14}$$

Now, θ indicates the helix angle, r refers to the search process, the radius of the helix. θ_1 shows the initial helix angle. $\tau(x)$ shows the Gamma function and β denotes the constant value set as 1.5. x and y indicate the shape of a spiral flight. D_1 denotes the integer matrix from 1 to the length of the search region.

$$y = r \times cos\,(\theta) \tag{15}$$

$$x = r \times sin\,(\theta) \tag{16}$$

$$r = r_1 + U \times D_1 \tag{17}$$

$$\theta = -\omega \times D_1 + \theta_1 \tag{18}$$

$$\theta_1 = \frac{3 \times \pi}{2} \tag{19}$$

where r_1 ranges from 1 to 20, ω takes the value 0.005, and$U$ take the value 0.00565.

3) The third stage is expanding the development stage, while the Aquila bird is in the hunting region, prepared to land and then attack; they usually adopt the vertical drop technique. The mathematical equation can be given as follows:

$$X_3\,(t+1) = (X_{best}\,(t) - X_M\,(t)) \times \alpha - rand + ((UB - LB) \times rand + LB) \times \sigma \tag{20}$$

In Eq. (17), α and β denote the development adjustment parameter that is smaller than 0.1.

4) In this stage, there exists a particular arbitrariness because of attacking prey, walking, and capturing the target to diminish the development while the Aquila bird is closer towards the prey. This can be formulated by Eq. (21):

$$X_4\,(t+1) = QF \times X_{best}(t) - (G_1 \times X\,(t) \times rand) - G_2 \times Levy\,(D) + rand \times G_1 \tag{21}$$

In Eq. (21), G_1 shows the different trajectories of the motion of Aquila birds while escaping prey, *and* G_2 signifies the reducing values from 2 to 0 during escaping prey. QF denotes the average search process of the mass function, and.

$$QF\,(t) = t\frac{2 \times rand\,() - 1}{(1-T)^2} \tag{22}$$

$$G_1 = 2 \times rand\,() - 1 \tag{23}$$

$$G_2 = 2 \times \left(1 - \frac{t}{T}\right) \tag{24}$$

AOA is a method that could attain better outcomes for complex multiobjective problems. Firstly, the initial group calculates all the indexes. In this phase, based on the present optimum result, the algorithm creates a novel group where every individual gives a new variable. Based on the fittest individual in the existing population, the model constantly seeks new optimum solutions and also makes various choices based on the best individual. The optimum strategy was selected through the rotary table when the number of optimum solutions was not sufficient. The process terminates the search and returns to the first step after finding the optimum solution. Lastly, the model selects the best scheme and sorts the best scheme from them. The AOA technique derives an FF to attain high efficiency of classification and defines a positive integer to characterize the superior outcome of the solution candidate. The decline of the classification error rate is viewed as an FF.

$$fitness\,(x_i) = Classifier\ Error\ Rate\,(x_i) = \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} * 100 \tag{25}$$

TABLE 1. Details of database.

Name	Class	No. of Samples
Healthy Tissue	He	330
Hypertrophic Blood Vessels	Hbv	330
Leukoplakia	Le	330
Abnormal IPCL-like Vessel	IPCL	330
Total Number of Samples		1320

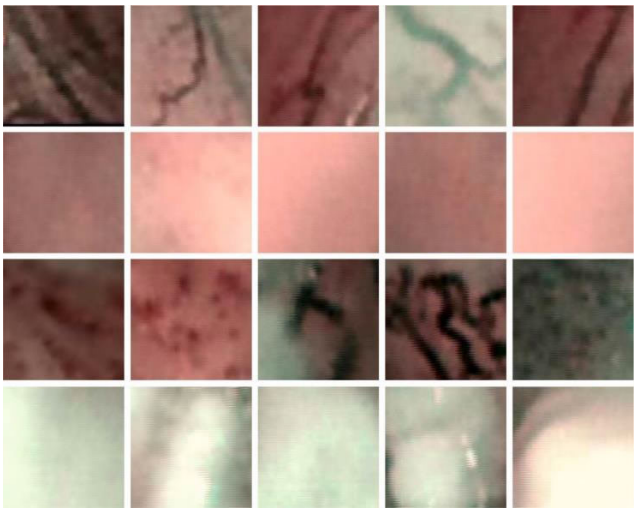

FIGURE 3. Sample images.

IV. RESULTS AND DISCUSSION

In this section, the cancer detection results of the LCDC-AOADL method are examined on the dataset [26], [27], comprising 1320 samples with four classes as depicted in Table 1. The database in the laryngeal dataset.tar comprises 1320 patches of healthy and early-stage cancerous laryngeal tissues. The patches (100 × 100 pixels) are manually extracted in 33 narrow-band laryngoscopic images of 33 distinct patients affected by laryngeal spinocellular carcinoma (analyzed after histopathological examination).

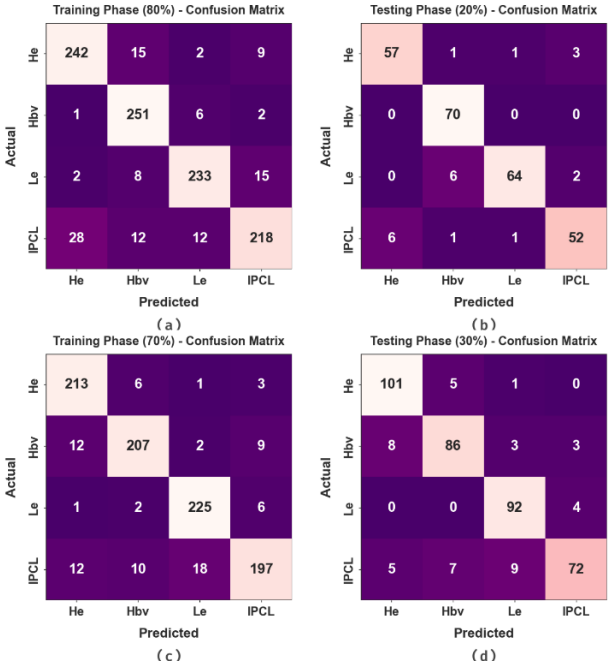

FIGURE 4. Confusion matrices of (a-b) 80:20 of TR set/TS set and (c-d) 70:30 of TR set/TS set.

TABLE 2. Cancer classifier outcome of LCDC-AOADL method on 80:20 of TR set/TS set.

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
Training Phase (80%)					
He	94.60	88.64	90.30	89.46	93.18
Hbv	95.83	87.76	96.54	91.94	96.07
Le	95.74	92.09	90.31	91.19	93.90
IPCL	92.61	89.34	80.74	84.82	88.72
Average	94.70	89.46	89.47	89.36	92.97
Testing Phase (20%)					
He	95.83	90.48	91.94	91.20	94.48
Hbv	96.97	89.74	100.00	94.59	97.94
Le	96.21	96.97	88.89	92.75	93.92
IPCL	95.08	91.23	86.67	88.89	92.11
Average	96.02	92.10	91.87	91.86	94.61

In detail, 4 tissue classes are assumed (330 patches/tissue class): He (healthy tissue), Hbv (tissue with hypertrophic vessels), Le (tissue with leukoplakia) and IPCL (tissue with intrapapillary capillary loops). All the subfolders comprise 4 folders comparative to the 4 tissue classes like Hbv, Le, He, and IPCL.

In Fig. 4, the confusion matrices of the LCDC-AOADL system on the cancer classification process are given. The result implies that the LCDC-AOADL method properly recognized four class labels effectively.

In Table 2 and Fig. 5, the cancer classification results of the LCDC-AOADL method are reported with 80:20 of the TR set/TS set. The results depict that the LCDC-AOADL technique reaches effectual performance. On 80% of the TR set, the LCDC-AOADL technique accomplishes an average $accu_y$ of 94.70%, $prec_n$ of 89.46%, $reca_l$ of 89.47%, F_{score} of 89.36%, and AUC_{score} of 92.97%. Besides, on 20% of the

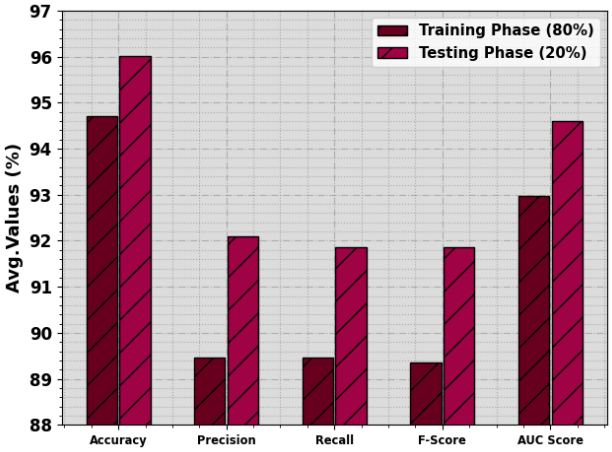

FIGURE 5. Average outcome of LCDC-AOADL method on 80:20 of TR set/TS set.

FIGURE 6. $Accu_y$ curve of LCDC-AOADL method on 80:20 of TR set/TS set.

TS set, the LCDC-AOADL method accomplishes an average $accu_y$ of 96.02%, $prec_n$ of 92.10%, $reca_l$ of 91.87%, F_{score} of 91.86%, and AUC_{score} of 94.61%.

Fig. 6 shows the training accuracy TR_accu_y and VL_accu_y of the LCDC-AOADL system at 80:20 of the TR set/TS set. The TL_accu_y is defined by the estimate of the LCDC-AOADL method on TR datasets while the VL_accu_y is calculated by estimating the performance on individual testing datasets. The outcomes exhibit that TR_accu_y and VL_accu_y increased with an upsurge in epochs. Therefore, the performance of the LCDC-AOADL approach was acquired to enhance the TR and TS dataset with an increased number of epochs.

In Fig. 7, the *TR_loss* and *VR_loss*analysis of the LCDC-AOADL algorithm on 80:20 of the TR set/TS set is revealed. The *TR_loss* determines the error between the computed and original values on the TR dataset. The *VR_loss* signifies the measure of the performance of the LCDC-AOADL method on individual validation data. The results indicate that the *TR_loss* and *VR_loss* tend to reduce with increasing epochs. It depicted the improved performance of the LCDC-AOADL algorithm and its ability to generate accurate

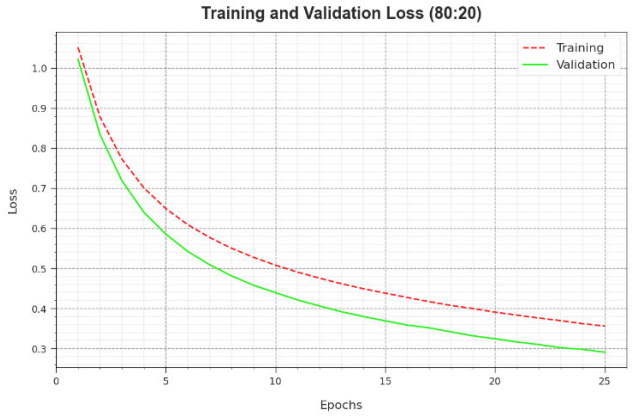

FIGURE 7. Loss curve of LCDC-AOADL method on 80:20 of TR set/TS set.

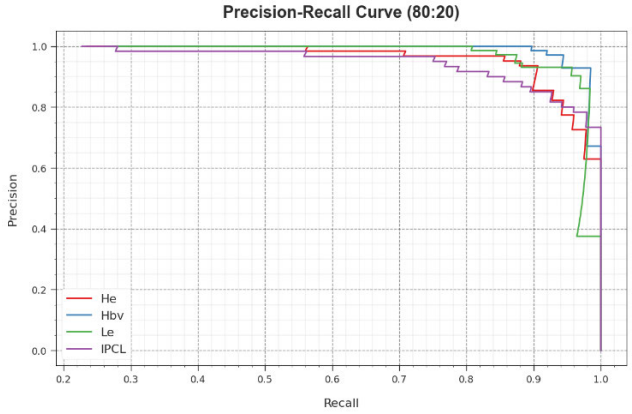

FIGURE 8. PR curve of LCDC-AOADL method on 80:20 of TR set/TS set.

classification. The decreased value of *TR_loss* and *VR_loss* proves the superior performance of the LCDC-AOADL system in capturing patterns and relationships.

A detailed precision-recall (PR) analysis of the LCDC-AOADL approach is shown on 80:20 of the TR set/TS set in Fig. 8. The outcomes stated that the LCDC-AOADL method results in raising values of PR. Moreover, the LCDC-AOADL algorithm can achieve greater PR values on all class labels.

In Fig. 9, a ROC curve of the LCDC-AOADL system is shown on 80:20 of the TR set/TS set. The figure defined that the LCDC-AOADL method resulted in enhanced ROC values. Also, the LCDC-AOADL approach can extend higher ROC values on all class labels.

In Table 3 and Fig. 10, the cancer classification analysis of the LCDC-AOADL method is reported with 70:30 of the TR set/TS set. The outcomes represent that the LCDC-AOADL system attained efficient performance. On 70% of the TR set, the LCDC-AOADL algorithm achieved an average $accu_y$ of 95.56%, $prec_n$ of 91.15%, $reca_l$ of 91.20%, F_{score} of 91.08%, and AUC_{score} of 94.12%. Besides, on 30% of the TS set, the LCDC-AOADL approach accomplished the average $accu_y$ of 94.32%, $prec_n$ of 88.78%, $reca_l$ of 88.41%, F_{score} of 88.38%, and AUC_{score} of 92.31%.

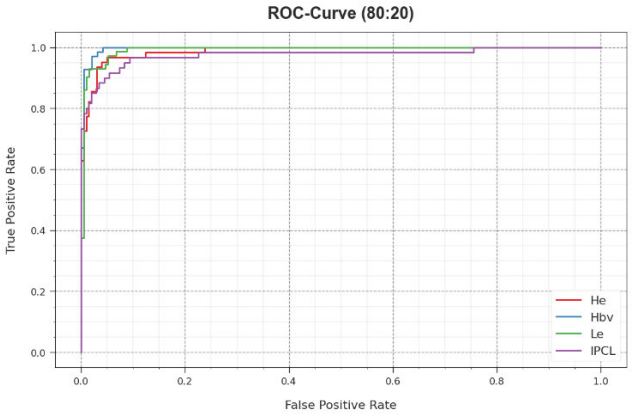

FIGURE 9. ROC curve of LCDC-AOADL method on 80:20 of TR set/TS set.

TABLE 3. Cancer classifier outcome of LCDC-AOADL method on 70:30 of TR set/TS set.

Class	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}	AUC_{Score}
Training Phase (70%)					
He	96.21	89.50	95.52	92.41	95.97
Hbv	95.56	92.00	90.00	90.99	93.70
Le	96.75	91.46	96.15	93.75	96.56
IPCL	93.72	91.63	83.12	87.17	90.25
Average	95.56	91.15	91.20	91.08	94.12
Testing Phase (30%)					
He	95.20	88.60	94.39	91.40	94.95
Hbv	93.43	87.76	86.00	86.87	90.97
Le	95.71	87.62	95.83	91.54	95.75
IPCL	92.93	91.14	77.42	83.72	87.55
Average	94.32	88.78	88.41	88.38	92.31

TABLE 4. Comparative outcome of LCDC-AOADL approach with other DL techniques.

Classifiers	$Accu_y$	$Prec_n$	$Reca_l$	F_{Score}
LCDC-AOADL	96.02	92.10	91.87	91.86
DCNN Model	84.00	89.23	85.92	86.93
Exception Model	90.10	87.56	86.84	86.15
ResNet50 Model	91.01	89.49	85.11	86.45
VGG19 Model	85.10	85.81	88.17	87.17
VGG16 Model	85.30	89.84	88.11	85.20
AlexNet Model	87.50	87.32	89.66	85.94

Fig. 11 shows the training accuracy TR_accu_y and VL_accu_y of the LCDC-AOADL method at 70:30 of the TR set/TS set. The TL_accu_y is defined by the estimation of the LCDC-AOADL system on TR datasets while the VL_accu_y is calculated by estimating the performance on individual testing datasets. The outcomes demonstrated that TR_accu_y and VL_accu_y rise with an upsurge in epochs. Therefore, the performance of the LCDC-AOADL algorithm was acquired to enhance the TR and TS datasets with an increased number of epochs.

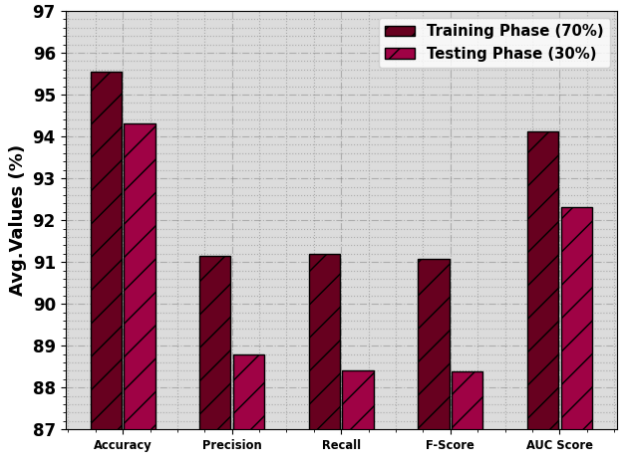

FIGURE 10. Average outcome of LCDC-AOADL method on 70:30 of TR set/TS set.

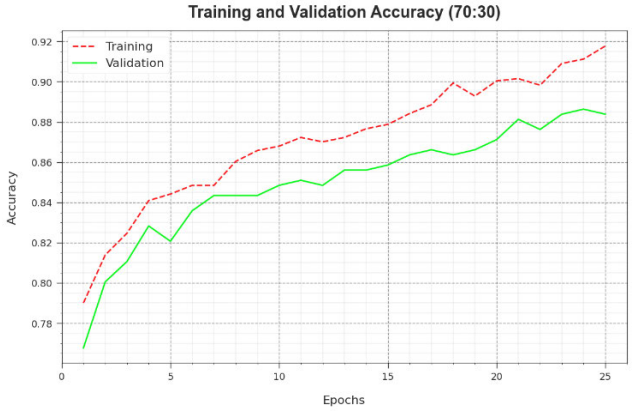

FIGURE 11. $Accu_y$ curve LCDC-AOADL method on 70:30 of TR set/TS set.

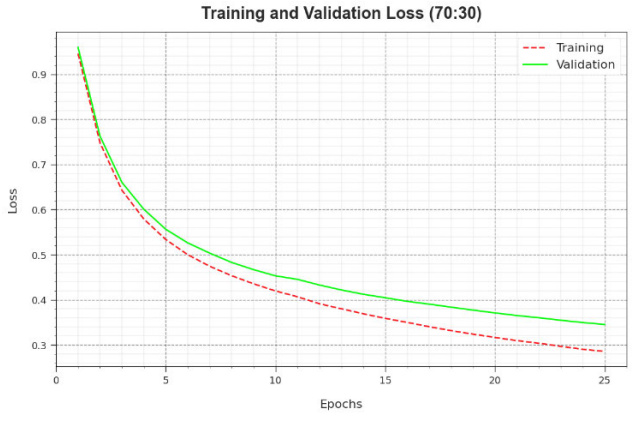

FIGURE 12. Loss curve of LCDC-AOADL method on 70:30 of TR set/TS set.

In Fig. 12, the *TR_loss* and *VR_loss*analysis of the LCDC-AOADL system at 70:30 of the TR set/TS set is demonstrated. The *TR_loss* defines the error amongst the predicted performance and original values on the TR datasets. The *VR_loss* signifies the measure of the performance of the LCDC-AOADL approach on separate validation data.

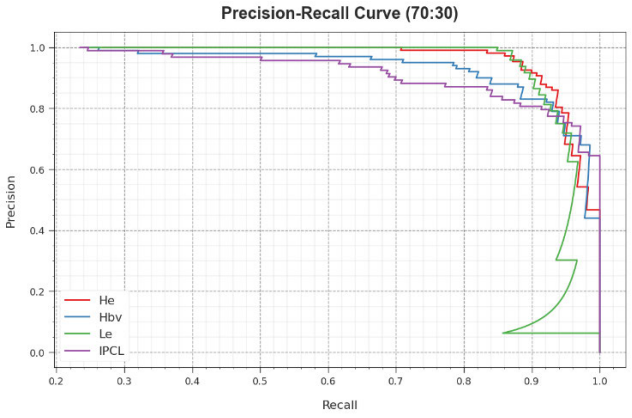

FIGURE 13. PR curve of LCDC-AOADL method on 70:30 of TR set/TS set.

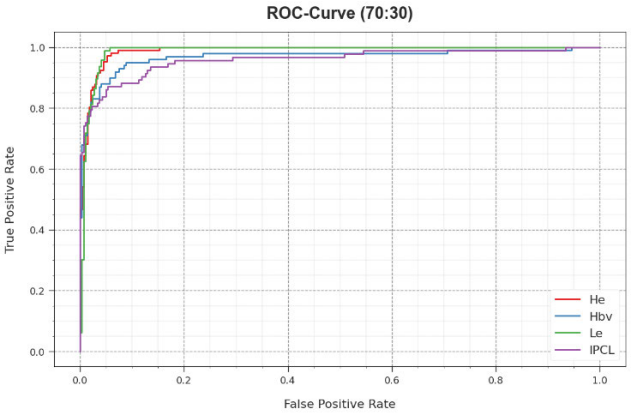

FIGURE 14. ROC curve of LCDC-AOADL method on 70:30 of TR set/TS set.

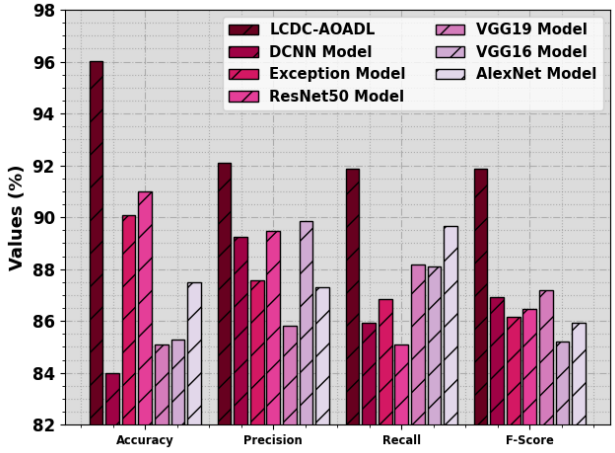

FIGURE 15. Comparative outcome of LCDC-AOADL approach with other DL techniques.

The outcomes indicate that the *TR_loss* and *VR_loss* tend to reduce with increasing epochs. It depicted the improved performance of the LCDC-AOADL algorithm and its ability to generate accurate classification. The decreased value of *TR_loss* and *VR_loss* shows the better performance of the LCDC-AOADL method in capturing relationships and patterns.

TABLE 5. CT outcome of LCDC-AOADL approach with other DL techniques.

Classifiers	Computational Time (sec)
LCDC-AOADL	1.85
DCNN Model	2.43
Exception Model	3.00
ResNet50 Model	4.83
VGG19 Model	4.30
VGG16 Model	3.85
AlexNet Model	5.12

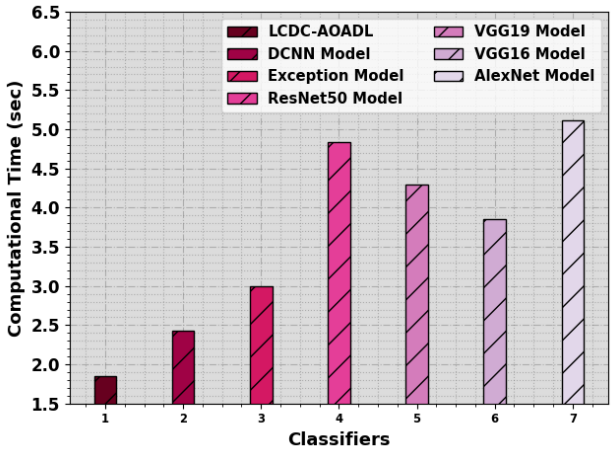

FIGURE 16. CT outcome of LCDC-AOADL approach with other DL methods.

A detailed precision-recall (PR) curve of the LCDC-AOADL method is shown at 70:30 of the TR set/TS set in Fig. 13. The outcomes stated that the LCDC-AOADL method results in raising values of PR. Furthermore, it is noticeable that the LCDC-AOADL algorithm can attain better PR values on all class labels.

In Fig. 14, a ROC analysis of the LCDC-AOADL system is revealed at 70:30 of the TR set/TS set. The figure defined that the LCDC-AOADL approach resulted in enhanced ROC values. Also, the LCDC-AOADL method can extend superior ROC values on all class labels.

In Table 4 and Fig. 15, the LCDC-AOADL technique is compared with other DL models [15]. The results exhibit improvements in the LCDC-AOADL technique compared to recent models. It is noted that the DCNN, VGG19, VGG16, and AlexNet models accomplish lower performance. At the same time, the Exception and ResNet50 models obtain slightly improvised results. However, the LCDC-AOADL technique reaches effectual performance with maximum $accu_y$, $prec_n$, $reca_l$, and F_{score} of 96.02%, 92.10%, 91.87%, and 91.86% respectively.

In Table 5 and Fig. 16, the CT analysis of the LCDC-AOADL system with recent techniques is provided. The outcomes signified that the LCDC-AOADL method achieves the minimum CT of 1.85s.

Simultaneously, current DCNN, Exception, ResNet50, VGG19, VGG16, and AlexNet techniques obtained the higher CT of 2.43s, 3.00s, 4.83s, 4.30s, 3.85s, and 5.12s correspondingly. These results demonstrated the superior results of the LCDC-AOADL method over other approaches.

V. CONCLUSION

In this study, we have developed the LCDC-AOADL method. The main purpose of the LCDC-AOADL technique is to examine the histopathological images for the recognition and classification of LC. In the presented LCDC-AOADL technique, three stages of operations are involved Inception v3-based feature extraction, DBN classification, and AOA-based hyperparameter tuning. The utilization of Inceptionv3 for feature extraction, the application of DBN for classification, and the incorporation of AOA for hyperparameter tuning collectively aim to advance the accuracy and efficiency of laryngeal cancer diagnosis from histopathological images. The experimental results pointed out the enhanced detection results of the LCDC-AOADL method over other recent approaches with a maximum accuracy of 96.02%, precision of 92.10%, recall of 91.87%, and F-score of 91.86%. In future, the performance of the LCDC-AOADL technique can be improved by the ensemble voting classifier model. Future work will combine explainable AI approaches are improve the interpretability of the model predictions, fostering trust and understanding among medical staff. Furthermore, studies into real-time applications and the possible utilization of LCDC-AOADL in medical settings facilitate its practical utility. At last, the presented approach is extended to multi-modal image data namely integrating histopathological images with radiological data, which holds the ability for widespread cancer analysis.

ACKNOWLEDGMENT

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through large group Research Project under grant number (RGP2/112/44). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R77), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. This study is supported via funding from Prince Sattam bin Abdulaziz University project number (PSAU/2023/R/1444). This study is partially funded by the Future University in Egypt (FUE).

REFERENCES

[1] G. O. Young, "Synthetic structure of industrial plastics," in *Plastics*, vol. 3, J. Peters, Ed., 2nd ed. New York, NY, USA: McGraw-Hill, 1964, pp. 15–64.

[2] A. M. Bur, T. Zhang, X. Chen, H. Kavookjian, S. Kraft, O. Karadaghy, N. Farrokhian, C. Mussatto, J. Penn, and G. Wang, "Interpretable computer vision to detect and classify structural laryngeal lesions in digital flexible laryngoscopic images," *Otolaryngol.-Head Neck Surg.*, Jun. 2023.

[3] D. J. Wellenstein, J. Woodburn, H. A. M. Marres, and G. B. van den Broek, "Detection of laryngeal carcinoma during endoscopy using artificial intelligence," *Head Neck*, vol. 45, no. 9, pp. 2217–2226, Sep. 2023.

[4] F. Meyer-Veit, R. Rayyes, A. O. H. Gerstner, and J. Steil, "Hyperspectral wavelength analysis with U-Net for larynx cancer detection," in *Proc. Eur. Symp. Artif. Neural Netw. (ESANN), Comput. Intell. Mach. Learn.*, Bruges, Belgium, Oct. 2022.

[5] D. Bhattacharya, F. Behrendt, A. Felicio-Briegel, V. Volgger, D. Eggert, C. Betz, and A. Schlaefer, "Learning robust representation for laryngeal cancer classification in vocal folds from narrow band images," *Med. Imag. Deep Learn.*, Apr. 2022.

[6] S. S. Raoof, M. A. Jabbar, and S. A. Fathima, "Lung cancer prediction using machine learning: A comprehensive approach," in *Proc. 2nd Int. Conf. Innov. Mech. Ind. Appl. (ICIMIA)*, Mar. 2020, pp. 108–115.

[7] S. S. Raoof, M. A. Jabbar, and S. A. Fathima, "Lung cancer prediction using feature selection and recurrent residual convolutional neural network (RRCNN)," in *Machine Learning Methods for Signal, Image and Speech Processing*. Denmark, U.K.: River Publishers, 2022, pp. 23–46.

[8] M. A. Jabbar, "Breast cancer data classification using ensemble machine learning," *Eng. Appl. Sci. Res.*, vol. 48, no. 1, pp. 65–72, 2021.

[9] R. Gurumoorthy and M. Kamarasan, "Computer aided breast cancer detection and classification using optimal deep learning," in *Proc. Int. Conf. Sustain. Comput. Data Commun. Syst. (ICSCDS)*, Mar. 2023, pp. 143–150.

[10] P. Huang, P. He, S. Tian, M. Ma, P. Feng, H. Xiao, F. Mercaldo, A. Santone, and J. Qin, "A ViT-AMC network with adaptive model fusion and multiobjective optimization for interpretable laryngeal tumor grading from histopathological images," *IEEE Trans. Med. Imag.*, vol. 42, no. 1, pp. 15–28, Jan. 2023.

[11] Z. You, Y. Yan, Z. Shi, M. Zhao, J. Yan, H. Liu, X. Hei, and X. Ren, "Laryngeal leukoplakia classification via dense multiscale feature extraction in white light endoscopy images," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Jun. 2023, pp. 1–5.

[12] H. Xiong, P. Lin, J.-G. Yu, J. Ye, L. Xiao, Y. Tao, Z. Jiang, W. Lin, M. Liu, J. Xu, W. Hu, Y. Lu, H. Liu, Y. Li, Y. Zheng, and H. Yang, "Computer-aided diagnosis of laryngeal cancer via deep learning based on laryngoscopic images," *EBioMedicine*, vol. 48, pp. 92–99, Oct. 2019.

[13] Y. He, Y. Cheng, Z. Huang, W. Xu, R. Hu, L. Cheng, S. He, C. Yue, G. Qin, Y. Wang, and Q. Zhong, "A deep convolutional neural network-based method for laryngeal squamous cell carcinoma diagnosis," *Ann. Transl. Med.*, vol. 9, no. 24, p. 1797, Dec. 2021.

[14] H. Kim, J. Jeon, Y. J. Han, Y. Joo, J. Lee, S. Lee, and S. Im, "Convolutional neural network classifies pathological voice change in laryngeal cancer with high accuracy," *J. Clin. Med.*, vol. 9, no. 11, p. 3415, Oct. 2020.

[15] P. K. Sahoo, S. Mishra, R. Panigrahi, A. K. Bhoi, and P. Barsocchi, "An improvised deep-learning-based mask R-CNN model for laryngeal cancer detection using CT images," *Sensors*, vol. 22, no. 22, p. 8834, Nov. 2022.

[16] M. Bengs, S. Westermann, N. Gessert, D. Eggert, A. O. H. Gerstner, N. A. Mueller, C. Betz, W. Laffers, and A. Schlaefer, "Spatio-spectral deep learning methods for in-vivo hyperspectral laryngeal cancer detection," *Proc. SPIE*, vol. 11314, pp. 369–374, Feb. 2020.

[17] M. N. Sachane and S. A. Patil, "Adaptive spotted hyena optimizer-enabled deep QNN for laryngeal cancer classification," in *Proc. Int. Conf. Edge Comput. Appl. (ICECAA)*, Oct. 2022, pp. 1025–1032.

[18] H. M. Tayade, "Early detection of laryngeal cancer using multiple instance learning based neural network," Doctoral dissertation, National College of Ireland, Dublin, Ireland, 2020.

[19] X. Zhou, C. Tang, P. Huang, F. Mercaldo, A. Santone, and Y. Shao, "LPCANet: Classification of laryngeal cancer histopathological images using a CNN with position attention and channel attention mechanisms," *Interdiscipl. Sci., Comput. Life Sci.*, vol. 13, no. 4, pp. 666–682, Dec. 2021.

[20] N. Esmaeili, E. Sharaf, E. J. G. Ataide, A. Illanes, A. Boese, N. Davaris, C. Arens, N. Navab, and M. Friebe, "Deep convolution neural network for laryngeal cancer classification on contact endoscopy-narrow band imaging," *Sensors*, vol. 21, no. 23, p. 8157, Dec. 2021.

[21] R. Maskeliūnas, A. Kulikajevas, R. Damaševičius, K. Pribuišis, N. Ulozaite-Staniene, and V. Uloza, "Lightweight deep learning model for assessment of substitution voicing and speech after laryngeal carcinoma surgery," *Cancers*, vol. 14, no. 10, p. 2366, May 2022.

[22] I. Kwon, S.-G. Wang, S.-C. Shin, Y.-I. Cheon, B.-J. Lee, J.-C. Lee, D.-W. Lim, C. Jo, Y. Cho, and B.-J. Shin, "Diagnosis of early glottic cancer using laryngeal image and voice based on ensemble learning of convolutional neural network classifiers," *J. Voice*, Sep. 2022.

[23] P. Dutta, K. A. Sathi, M. A. Hossain, and M. A. A. Dewan, "Conv-ViT: A convolution and vision transformer-based hybrid feature extraction method for retinal disease detection," *J. Imag.*, vol. 9, no. 7, p. 140, Jul. 2023, doi: 10.3390/jimaging9070140.

[24] N. Alqahtani, S. Alam, I. Aqeel, M. Shuaib, I. M. Khormi, S. B. Khan, and A. A. Malibari, "Deep belief networks (DBN) with IoT-based Alzheimer's disease detection and classification," *Appl. Sci.*, vol. 13, no. 13, p. 7833, Jul. 2023.

[25] L. Liu and Y. Li, "Research on a photovoltaic power prediction model based on an IAO-LSTM optimization algorithm," *Processes*, vol. 11, no. 7, p. 1957, Jun. 2023.

[26] *Zenodo*. Accessed: Mar. 14, 2023. [Online]. Available: https://zenodo.org/record/1003200

[27] S. Moccia, E. De Momi, M. Guarnaschelli, M. Savazzi, A. Laborai, L. Guastini, and G. Peretti, "Confident texture-based laryngeal tissue classification for early stage diagnosis support," *J. Med. Imag.*, vol. 4, no. 3, pp. 1–10, 2017.