

## Deep Transfer Learning Driven Oral Cancer Detection and Classification Model

Radwa Marzouk<sup>1</sup>, Eatedal Alabdulkreem<sup>2</sup>, Sami Dhabbi<sup>3</sup>, Mohamed K. Nour<sup>4</sup>, Mesfer Al Duhayyim<sup>5</sup>, Mahmoud Othman<sup>6</sup>, Manar Ahmed Hamza<sup>7,\*</sup>, Abdelwahed Motwakel<sup>7</sup>, Ishfaq Yaseen<sup>7</sup> and Mohammed Rizwanullah<sup>7</sup>

<sup>1</sup>Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, 11671, Saudi Arabia

<sup>2</sup>Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, 11671, Saudi Arabia

<sup>3</sup>Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia

<sup>4</sup>Department of Computer Sciences, College of Computing and Information System, Umm Al-Qura University, Saudi Arabia

<sup>5</sup>Department of Natural and Applied Sciences, College of Community-Aflaj, Prince Sattam Bin Abdulaziz University, Saudi Arabia

<sup>6</sup>Department of Computer Science, Faculty of Computers and Information Technology, Future University in Egypt, New Cairo, 11835, Egypt

<sup>7</sup>Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam Bin Abdulaziz University, AlKharj, Saudi Arabia

\*Corresponding Author: Manar Ahmed Hamza. Email: ma.hamza@psau.edu.sa

Received: 01 March 2022; Accepted: 26 April 2022

**Abstract:** Oral cancer is the most commonly occurring ‘head and neck cancers’ across the globe. Most of the oral cancer cases are diagnosed at later stages due to absence of awareness among public. Since earlier identification of disease is essential for improved outcomes, Artificial Intelligence (AI) and Machine Learning (ML) models are used in this regard. In this background, the current study introduces Artificial Intelligence with Deep Transfer Learning driven Oral Cancer detection and Classification Model (AIDTL-OCCM). The primary goal of the proposed AIDTL-OCCM model is to diagnose oral cancer using AI and image processing techniques. The proposed AIDTL-OCCM model involves fuzzy-based contrast enhancement approach to perform data pre-processing. Followed by, the densely-connected networks (DenseNet-169) model is employed to produce a useful set of deep features. Moreover, Chimp Optimization Algorithm (COA) with Autoencoder (AE) model is applied for oral cancer detection and classification. Furthermore, COA is employed to determine optimal parameters involved in AE model. A wide range of experimental analyses was conducted on benchmark datasets and the results were investigated under several aspects. The extensive experimental analysis outcomes established the enhanced performance of AIDTL-OCCM model compared to other approaches with a maximum accuracy of 90.08%.



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

**Keywords:** Oral cancer detection; image classification; artificial intelligence; machine learning; deep learning

## 1 Introduction

Oral cancer is one of the leading cancers across the globe based on its morbidity rate, late diagnoses, and high mortality rate. Tobacco and excessive alcohol usage are major risk factors that cause oral lesions [1]. Betel quid consumption is the most important causal factor behind high oral cancer prevalence in South and Southeast Asia. Betel quid usually contains tobacco, slaked lime, areca nut, and betel leaf [2–4]. At present, betel quid is commercially available in sachets and is commonly consumed in public because of its dynamic marketing strategy. Oral lesion is widely related to delay in presentation i.e., over two-third present at later stage which consequently results in poor survival rate [5]. Lesion management, particularly at later stages, is expensive [6]. Lack of knowledge and public awareness concern the medical experts since oral lesion is a major cause for late diagnosis [7]. Late detection does not define the characteristics of the lesion, since oral lesion is frequently headed by visible oral lesion named as Oral Potentially Malignant Disorder (OPMD). This is diagnosed during routine screening through Clinical Oral Examination (COE) conducted by dental specialists. When an apprehensive cancer is recognized, the patient is recommended for professional examination for further management and confirmation of diagnosis.

Prior research in India reveals that screening and earlier diagnoses have reduced the mortality rates and down-staging of disease among individuals who use alcohol and tobacco [8]. The screening programs for oral lesions must be vigorous, cost-effective and efficient since there is an ongoing constraint exists in terms of health resources and healthcare professionals. In this background, technology, in the form of telemedicine, can be a feasible method for early diagnosis. Advancements in the fields of Deep Learning (DL) and Computer Vision (CV) offer effective methods to design adjunctive technology that can automatically screens the oral cavity of an individual and provide feedback to medical professionals during patient examination and to the individual for self-examination [9]. The study conducted earlier [10] on image-based automated diagnoses of oral cancer has mainly focused on the usage of imaging techniques such as hyperspectral imaging, auto fluorescence imaging, and optical coherence tomography. While one or two studies have implemented white-light photographic image, almost all of the studies conducted earlier focus on the recognition of specific kinds of oral lesion.

The current study introduces Artificial Intelligence (AI) with Deep Transfer Learning Driven Oral Cancer detection and Classification Model (AIDTL-OCCM). The primary goal of the proposed AIDTL-OCCM model is to recognize the oral cancer using AI and image processing techniques. Firstly, the proposed AIDTL-OCCM model performs fuzzy-based contrast enhancement approach to perform data pre-processing. In addition, densely-connected networks (DenseNet-169) model is employed to produce a useful set of features. Moreover, Chimp Optimization Algorithm (COA) with Autoencoder (AE) model is applied for detection and classification of oral cancer. Furthermore, COA is employed to determine the optimal parameters involved in AE model. A wide range of experimental analyses was conducted on benchmark datasets and the results were investigated under several aspects.

Rest of the paper is organized as follows. Section 2 offers the works related to the domain and Section 3 introduces the proposed model. Next, Section 4 provides experimental validation and Section 5 concludes the study.

## 2 Related works

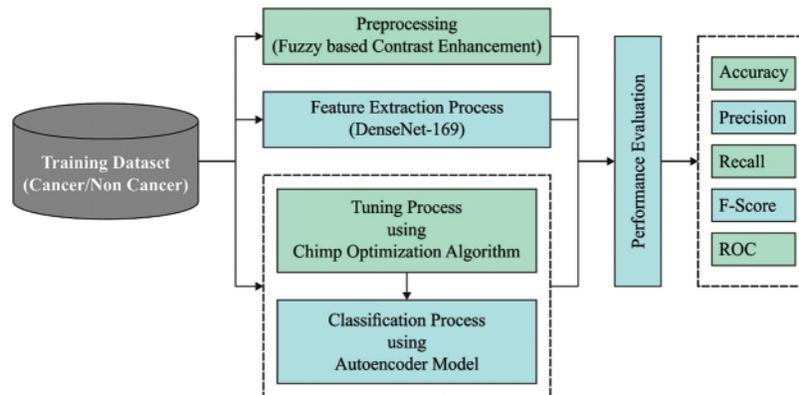
In literature [11], a new approach was proposed to combine the bounding box annotation from several clinicians. In addition, Deep Neural Network (DNN) was also utilized to create automated systems, where difficult patterns are created as a result to tackle this complex task. Since the primary data collected under this case is utilized, two DL-based CV methods were considered for automated recognition and classification of oral lesions. Jeyaraj et al. [12] established a DL technique for automated and computer-aided oral cancer identification method in which patient hyperspectral image is considered. In order to validate the presented regression-based partitioned DL technique, its performance was compared with other approaches in terms of classifier sensitivity, accuracy, and specificity. To achieve accurate medicinal image classification, a novel infrastructure named partitioned deep Convolution Neural Network (DCNN) was presented with two partitioned layers for labeling and classified by labeling the Region Of Interest (ROI) from multi-dimensional hyperspectral images.

Jubair et al. [13] established a lightweight DCNN for binary classification of oral lesions so as to segregate benign and malignant or potentially-malignant lesions using normal real-time medical images. The data set of 716 medical images was utilized for training and testing the presented method. Tanriver et al. [14] found that both CV and DL approaches have enormous potential in oral cancer detection through photographic images. They analyzed the diagnostic outcomes of automated schemes to identify the potential oral-malignant disorder with a 2-phase pipeline. The initial outcomes demonstrate the possibility of DL-based techniques for automated recognition and classification of oral lesions from real time.

Lin et al. [15] projected an effectual smartphone-based image analysis approach controlled by DL technique to address the automatic detection challenges of oral diseases. It can be conducted as a retrospective analysis. An easy yet effectual centered-rule image-capturing technique is presented in this study to gather oral cavity images. Next, according to this technique, a medium-sized oral data set with five types of diseases are generated, and a resampling approach is projected to alleviate the effect of images' variability in hand-held smartphone cameras. Eventually, a new DL network (HRNet) has been established to evaluate the performance of the proposed technique in terms of oral cancer recognition. Nanditha et al. [16] established an automated system to analyze oral lesions with the help of DL approaches. An ensemble DL technique that integrates the advantages of Resnet-50 and visual geometry group (VGG)-16 was established in this study. This method was trained with augmented data set of oral lesion images. This technique was compared and contrasted with other commonly-utilized DL approaches in terms of implementing the classifier of oral images.

## 3 The Proposed Model

The current study has developed a novel AIDTL-OCCM model for effectual recognition and classification of oral cancer. The presented AIDTL-OCCM technique undergoes a series of operations (as shown in Fig. 1) such as fuzzy-enabled pre-processing, DenseNet-169 feature extraction, AE-based classification, and COA-based parameter optimization. Furthermore, COA is employed to determine the optimal parameters involved in AE model.



**Figure 1:** Workflow of AIDTL-OCCM model

### 3.1 Fuzzy Based Contrast Enhancement

In order to enhance the contrast level of images, Histogram Equalization (HE) is employed. Fuzzy-enabled image pre-processing helps in generating highly-sophisticated image divergence compared to input images. High resolution medical images are difficult to understand due to which image pre-processing is carried out to improve the quality of the image. A major benefit of this model is the transformation of color images into grayscale ones. In order to visualize the image with high clarity, fuzzy-based Contrast Limited Adaptive Histogram Equalization (CLAHE) model is used.

### 3.2 Dense-169 Based Feature Extraction

Followed by, DenseNet-169 model processes the pre-processed images to create a set of feature vectors [17]. DCNN is a highly efficient framework in image recognition process due to the existence of particular kinds of pooling layer and convolution layer. Nonetheless, when the system goes in-depth, the input data or gradient that passes by the layer gets disappeared with the time, the previous layer of the network is obtained. DenseNet overcomes the challenge of vanishing gradients by interconnecting each layer with equivalent feature-size straightaway with one another. The main objective of utilizing DenseNet model, as a feature extractor is that deeper the network, more generic feature can be attained. DenseNet-169 model includes one pooling and convolutional layer at first followed by three transition layers and four dense blocks. After this layer, the last layer i.e., classification layer exists. The layered architecture is depicted in Fig. 2.

Layers	Output Size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112		7 × 7 conv, stride 2		
Pooling	56 × 56		3 × 3 max pool, stride 2		
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56 × 56		1 × 1 conv		
	28 × 28		2 × 2 average pool, stride 2		
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28 × 28		1 × 1 conv		
	14 × 14		2 × 2 average pool, stride 2		
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 64$
Transition Layer (3)	14 × 14		1 × 1 conv		
	7 × 7		2 × 2 average pool, stride 2		
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$
Classification Layer	1 × 1		7 × 7 global average pool		
			1000D fully-connected, softmax		

**Figure 2:** Layers in DenseNet model

The initial convolution layer executes  $7 \times 7$  convolution with stride 2 and  $3 \times 3$  max pooling utilized by stride 2. Next, the network comprises of a dense block and three sets of them contain a dense block and a transition layer. The dense connectivity in DenseNet is established by transporting a direct connection from one layer to another. The  $l$ th layer receives the feature-map of each previous layer, thereby, ameliorating the flow of gradient all over the network. This needs the concatenation of feature-map of the preceding layer that could not be completed except the case in which every feature-map is of the similar size. However, Convolution Neural Network mainly intends toward the downsampling of size of the feature-map. The DenseNet model is separated into several densely-connected dense blocks as mentioned above. The layer between these dense blocks is represented as a transition layer. All the transition layers contain  $2 \times 2$  average pooling layer, batch normalization layer, and  $1 \times 1$  convolution layer which employs a stride of 2. As mentioned above, there are four dense blocks and all of them comprise of two convolutional layers,  $1 \times 1$  and  $3 \times 3$  sizes. The size of each 4 dense block in DenseNet169 model pre-trained on ImageNet is 6, 12, 32, and 32. Then, the last layer is the classification layer that implements the global average pooling of  $7 \times 7$  whereas the last FC layer employs ‘softmax’ as the activation.

### 3.3 AE-Based Classification

At the time of classification process, AE model is used to properly determine the class labels for test images [18]. AE consists of a hidden layer of  $h$  unit, one visible layer of  $d$  input, an activation function  $f$  and one reconstruction layer of  $d$  unit. In training process, initially it maps the input  $x \in R^d$  to an hidden state, thus producing the latent activity,  $y \in R^h$ . The network corresponds to the step as given in the boxed part and is named an ‘encoder’. Next,  $y$  is mapped using ‘decoder’ to output layer that has similar size of input layer and is named as ‘reconstruction’. The reconstructing value is represented by  $z \in R^d$ . They are mathematically expressed as follows.

$$y = f(W_y x + b_y) \quad (1)$$

$$z = f(W_z y + b_z) \quad (2)$$

whereas,  $b_y$  and  $b_z$  represent the bias of hidden and output units,  $f(\cdot)$  signifies the activation function,  $W_y$  and  $W_z$  denote the input-to-hidden and the hidden-to-output weights correspondingly. Generally,

non-linearity is given as  $f(\cdot)$ . There are many other possibilities for  $f(\cdot)$  namely, hyperbolic tangent, rectified linear function, and sigmoid function.

In this work, the subsequent constraints exist

$$W_y = W_z = W. \quad (3)$$

Consider that AE has tied weight that assists in halving the model parameter. Therefore, there are three classes of variables that remains to learn:  $W$ ,  $b_y$ ,  $b_z$ . The aim of the training is to reduce the “error” between reconstruction and input as given herewith.

$$\arg \min [c(x, z)] \quad (4)$$

In which  $z$  denotes the variables  $W$ ,  $b_y$ ,  $b_z$  whereas  $x$  is provided.  $c(x, z)$  denotes the “error,” which is determined through different ways. Therefore, weight update rule is determined as follows (whereas  $\eta$  indicates the learning rate)

$$W = W - \eta \frac{\partial \text{cost}(x, z)}{\partial W} \quad (5)$$

$$b_y = b_y - \eta \frac{\partial \text{cost}(x, z)}{\partial b_y} \quad (6)$$

$$b_z = b_z - \eta \frac{\partial \text{cost}(x, z)}{\partial b_z}. \quad (7)$$

After training the network, the reconstructed layer and the parameter are eliminated whereas the learned features in hidden node could be consequently utilized for classification or as input of high layers for the production of deep features.

### 3.4 COA-Based Parameter Optimization

At the time of parameter optimization process, COA is employed to identify the parameter values of AE model. In the event of non-uniform irradiance conditions (i.e., partial shading), COA utilizes the approach to drive and chase the neighborhood for maximal global functioning points with respect to voltage, as given herewith [19]:

$$\chi = |\rho V_{mp\_STC}(i) - \delta V_{pv}(i)| \quad (8)$$

$$V_{pv}(i+1) = V_{mp\_STC}(i) - \sigma \cdot \chi \quad (9)$$

At this point, ‘i’ refers to the iteration number, and  $\chi$  signifies the distance between actual voltage and the chosen functioning point voltage, but  $\sigma$ ,  $\delta$ , and  $\rho$  depict the co-efficient vectors.  $V_{mp\_STC}$  stands for maximal power point voltage at typical test situations, and  $V_{pv}$  represents the measured voltage value in non-uniform irradiance difference. The co-efficient vector is computed as follows.

$$\sigma = 2\beta n_1 - \beta \quad (10)$$

$$\rho = 2n_2 \quad (11)$$

$$\delta = \text{chaotic\_value} \quad (12)$$

At this point,  $\beta'$  decreases the non-linearly in 2.5 to 0 under iteration method. But  $n_1'$  and  $n_2'$  are arbitrary vectors in the range of zero to one. Finally,  $\delta'$  refers to chaotic vector that is defined as dependent upon several chaotic maps. In order to track the optimum functioning, COA technique utilizes their four chimp groups (for instance, attackers, drivers, chasers, and barriers). According to convergence conditions, all the chimps group upgrades their place based on the position of optimum chimp. It is demonstrated as follows.

$$\chi_A = \rho_1 \cdot V_A(i) - \delta_1 \cdot V_{pv}(i) |$$

$$\chi_B = \rho_2 \cdot V_B(i) - \delta_2 \cdot V_{pv}(i) | \quad (13)$$

$$\chi_C = \rho_3 \cdot V_C(i) - \delta_3 \cdot V_{pv}(i) |$$

$$\chi_D = \rho_4 \cdot V_D(i) - \delta_4 \cdot V_{pv}(i) |$$

and

$$V_1 = V_A(i) - \sigma_1 \cdot (\chi_A); V_2 = V_B(i) - \sigma_2 \cdot (\chi_B) \quad (14)$$

$$V_3 = V_C(i) - \sigma_3 \cdot (\chi_C); V_4 = V_D(i) - \sigma_4 \cdot (\chi_D)$$

At this point after upgrading the place of all the chimps, the optimum voltage is as follows.

$$V_{pv}(i+1) = \frac{V_1 + V_2 + V_3 + V_4}{4} \quad (15)$$

where  $\rho'$  vector is an arbitrary allocated variable. It avoids the traps from local optimum and increases the stochastic performance of COA approach. The threshold  $\mu < 0.5$  and  $|\sigma| < 1$  make sure to achieve convergence near the Global Maximal Power Point (GMPP); but,  $|\sigma| > 1$  force chimps diverge and globally explore further spaces. Furthermore, the flexible tuning of ' $\rho$ ' and ' $\delta$ ' vector permits for local optimum avoidance and faster convergence concurrently.

#### 4 Experimental Validation

The experimental validation was conducted for the proposed AIDTL-OCCM model using the benchmark dataset from Kaggle repository as illustrated in Fig. 3 (available at <https://www.kaggle.com/shivam17299/oral-cancer-lips-and-tongue-images>). The dataset holds lips and tongue images under two classes namely, cancer (87 images) and non-cancer (44 images).

Fig. 4 illustrates the set of confusion matrices generated by the proposed AIDTL-OCCM model under five runs. The figure points out that AIDTL-OCCM model recognized the oral cancer effectually under every run. For instance, with run-1, the proposed AIDTL-OCCM model classified 81 images under cancer and 37 images under non-cancer. Similarly, with run-2, the proposed AIDTL-OCCM model classified 84 images under cancer and 28 images under non-cancer. Likewise, with run-5, the proposed AIDTL-OCCM model classified 81 images under cancer and 35 images under non-cancer.



Figure 3: Sample images

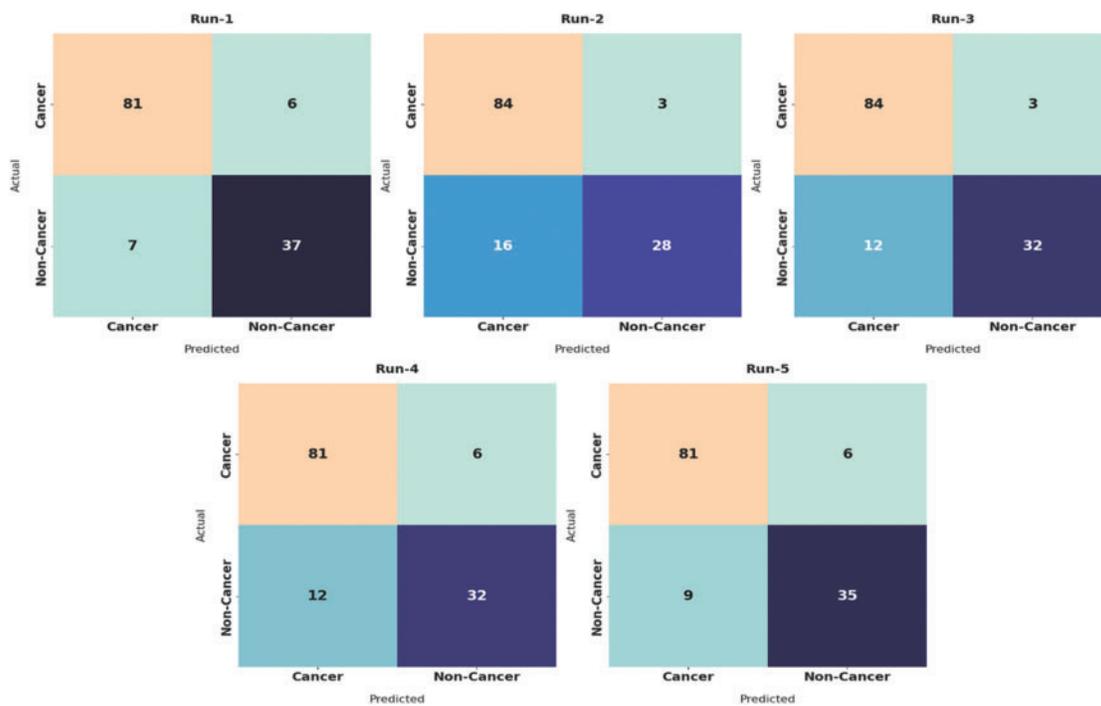


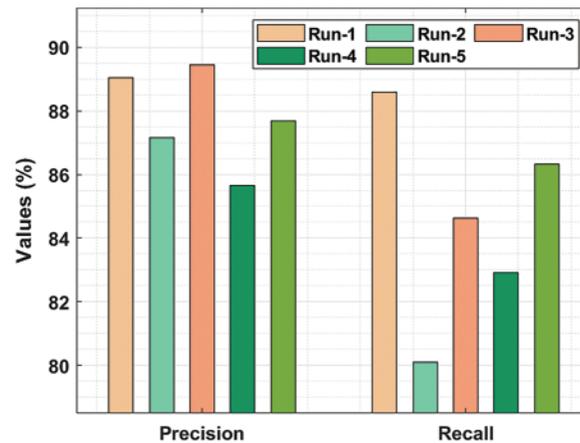
Figure 4: Confusion matrices of AIDTL-OCCM model

Tab. 1 showcases the overall oral cancer classification outcomes of the proposed AIDTL-OCCM model after five runs of execution. The experimental values indicate that the proposed AIDTL-OCCM model has accomplished superior outcomes under all the runs.

**Table 1:** Classification results of AIDTL-OCCM model

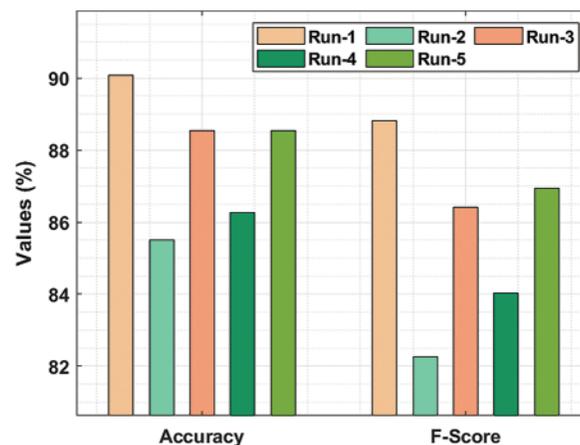
Class labels	Accuracy	Precision	Recall	F-Score
Run-1				
Cancer	90.08	92.05	93.10	92.57
Non-cancer	90.08	86.05	84.09	85.06
Average	90.08	89.05	88.60	88.81
Run-2				
Cancer	85.50	84.00	96.55	89.84
Non-cancer	85.50	90.32	63.64	74.67
Average	85.50	87.16	80.09	82.25
Run-3				
Cancer	88.55	87.50	96.55	91.80
Non-cancer	88.55	91.43	72.73	81.01
Average	88.55	89.46	84.64	86.41
Run-4				
Cancer	86.26	87.10	93.10	90.00
Non-cancer	86.26	84.21	72.73	78.05
Average	86.26	85.65	82.92	84.02
Run-5				
Cancer	88.55	90.00	93.10	91.53
Non-cancer	88.55	85.37	79.55	82.35
Average	88.55	87.68	86.32	86.94

Fig. 5 demonstrates the  $prec_n$  and  $reca_i$  inspection results accomplished by AIDTL-OCCM model under five runs. The figure reports that the proposed AIDTL-OCCM model obtained the maximum values of  $prec_n$  and  $reca_i$ . For instance, with run-1, the AIDTL-OCCM model identified the cancer images with  $prec_n$  and  $reca_i$  of 92.05% and 93.10% respectively. Along with that, with run-1, the proposed AIDTL-OCCM model identified non-cancer images with  $prec_n$  and  $reca_i$  of 86.05% and 84.09% respectively. Similarly, with run-2, the presented AIDTL-OCCM model recognized the cancer images with  $prec_n$  and  $reca_i$  of 84% and 96.55% respectively. In line with these, with run-2, the proposed AIDTL-OCCM model recognized the non-cancer images with  $prec_n$  and  $reca_i$  of 90.32% and 63.64% respectively.



**Figure 5:**  $Prec_n$  and  $Reca_l$  analysis results of AIDTL-OCCM model

Fig. 6 demonstrates the  $accu_y$  and  $F_{score}$  inspection results attained by the proposed AIDTL-OCCM model under five runs. The figure reports that AIDTL-OCCM model obtained maximum  $accu_y$  and  $F_{score}$  values. For instance, with run-1, the proposed AIDTL-OCCM model identified cancer images with  $accu_y$  and  $F_{score}$  of 90.08% and 92.57% respectively. Along with that, with run-1, AIDTL-OCCM model identified the non-cancer images with  $accu_y$  and  $F_{score}$  of 90.08% and 85.06% respectively. Similarly, with run-2, the proposed AIDTL-OCCM model recognized the cancer images with  $accu_y$  and  $F_{score}$  of 85.50% and 96.55% respectively. In line with these, with run-2, AIDTL-OCCM model recognized the non-cancer images with  $accu_y$  and  $F_{score}$  of 85.50% and 74.67% respectively.



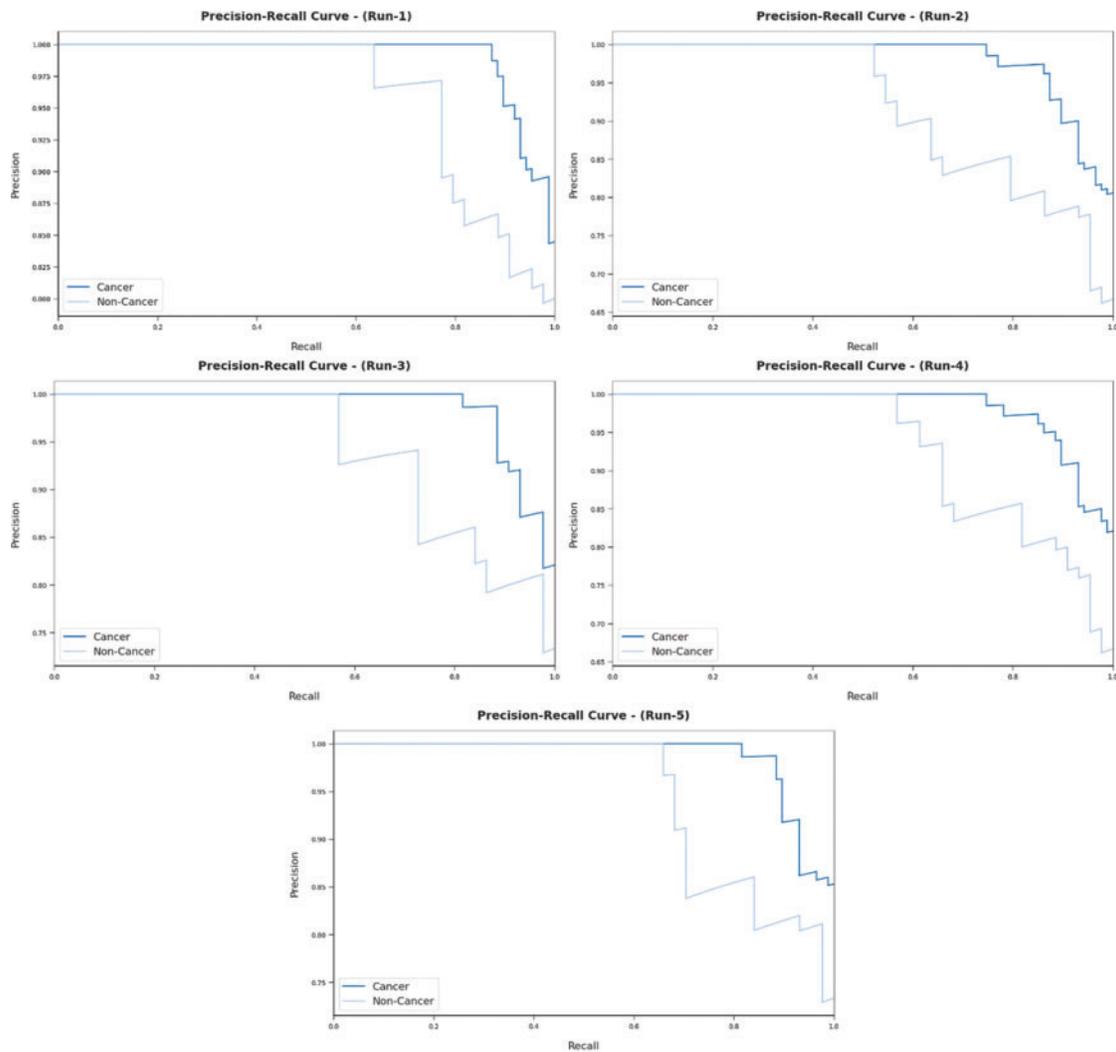
**Figure 6:**  $Accu_y$  and  $F_{score}$  analysis results of AIDTL-OCCM model

Tab. 2 highlights the average oral cancer classifier outcomes accomplished by the proposed AIDTL-OCCM model with five runs. With run-1, AIDTL-OCCM model achieved average  $accu_y$ ,  $prec_n$ ,  $reca_l$ , and  $F_{score}$  values such as 90.80%, 89.05%, 88.60%, and 88.81% respectively. Similarly, under run-3, AIDTL-OCCM model attained average  $accu_y$ ,  $prec_n$ ,  $reca_l$ , and  $F_{score}$  values such as 88.55%, 89.46%, 84.64%, and 86.41% respectively. Simultaneously, with run-5, AIDTL-OCCM model offered average  $accu_y$ ,  $prec_n$ ,  $reca_l$ , and  $F_{score}$  values such as 88.55%, 87.68%, 86.32%, and 86.94% respectively.

**Table 2:** Average classification results of AIDTL-OCCM model

No. of runs	Accuracy	Precision	Recall	F-score
Run-1	90.08	89.05	88.60	88.81
Run-2	85.50	87.16	80.09	82.25
Run-3	88.55	89.46	84.64	86.41
Run-4	86.26	85.65	82.92	84.02
Run-5	88.55	87.68	86.32	86.94
Average	87.79	87.80	84.51	85.69

Fig. 7 reports the results from precision-recall curve analysis achieved by AIDTL-OCCM model under five runs. The figure indicates that AIDTL-OCCM model produced effectual outcomes under all five runs.



**Figure 7:** Precision recall analysis results of AIDTL-OCCM model

Fig. 8 portrays ROC analysis results of AIDTL-OCCM model in classifying two classes under five runs. The figure indicates that the proposed AIDTL-OCCM model produced the maximum ROC values during the categorization of cancer and non-cancer images.

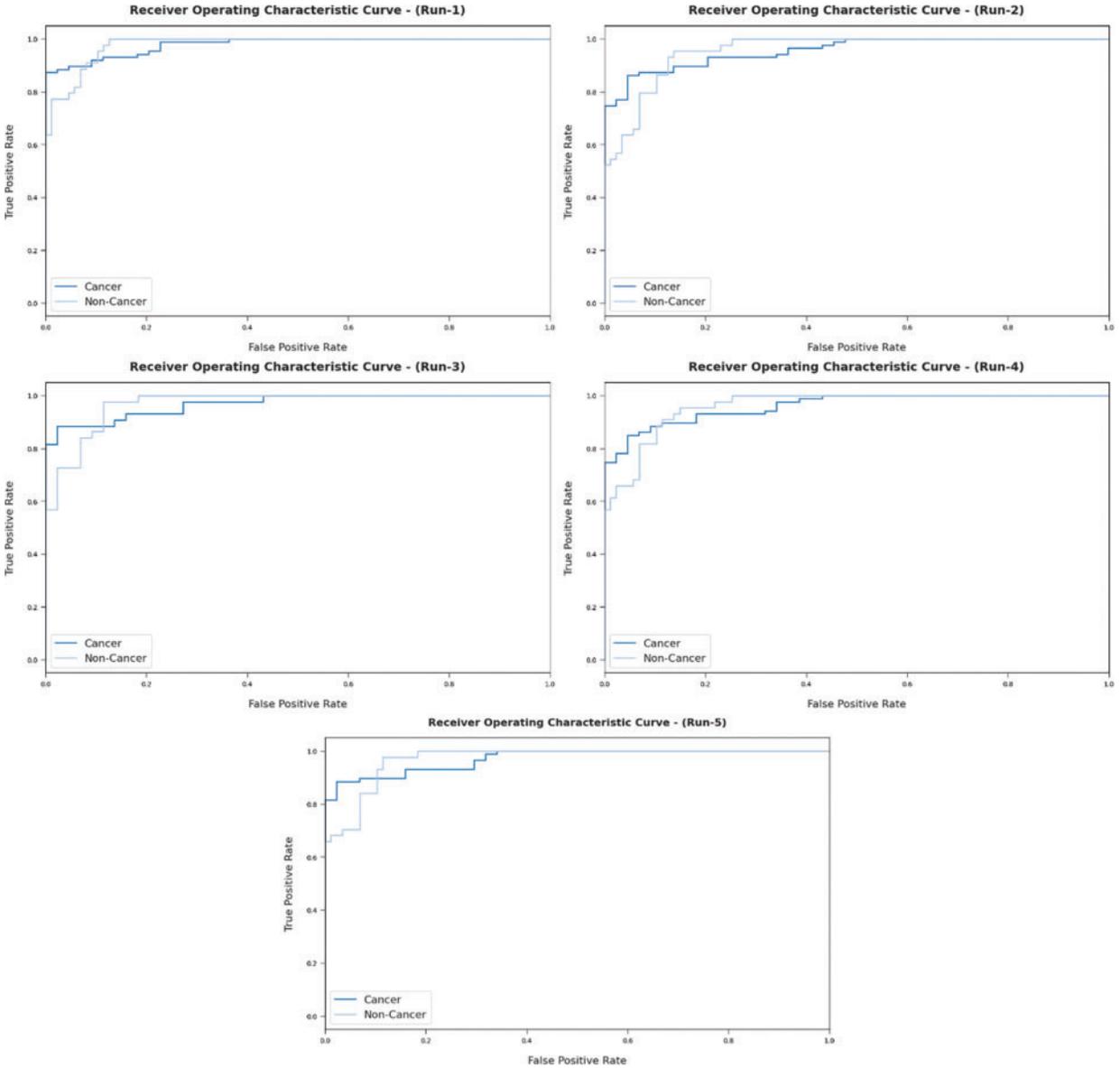


Figure 8: ROC analysis results of AIDTL-OCCM model

Fig. 9 illustrates a set of training and validation accuracy/loss graph plotted by AIDTL-OCCM model under five runs. In this figure, it can be observed that the validation accuracy is higher and validation loss is lower compared to training accuracy/loss.

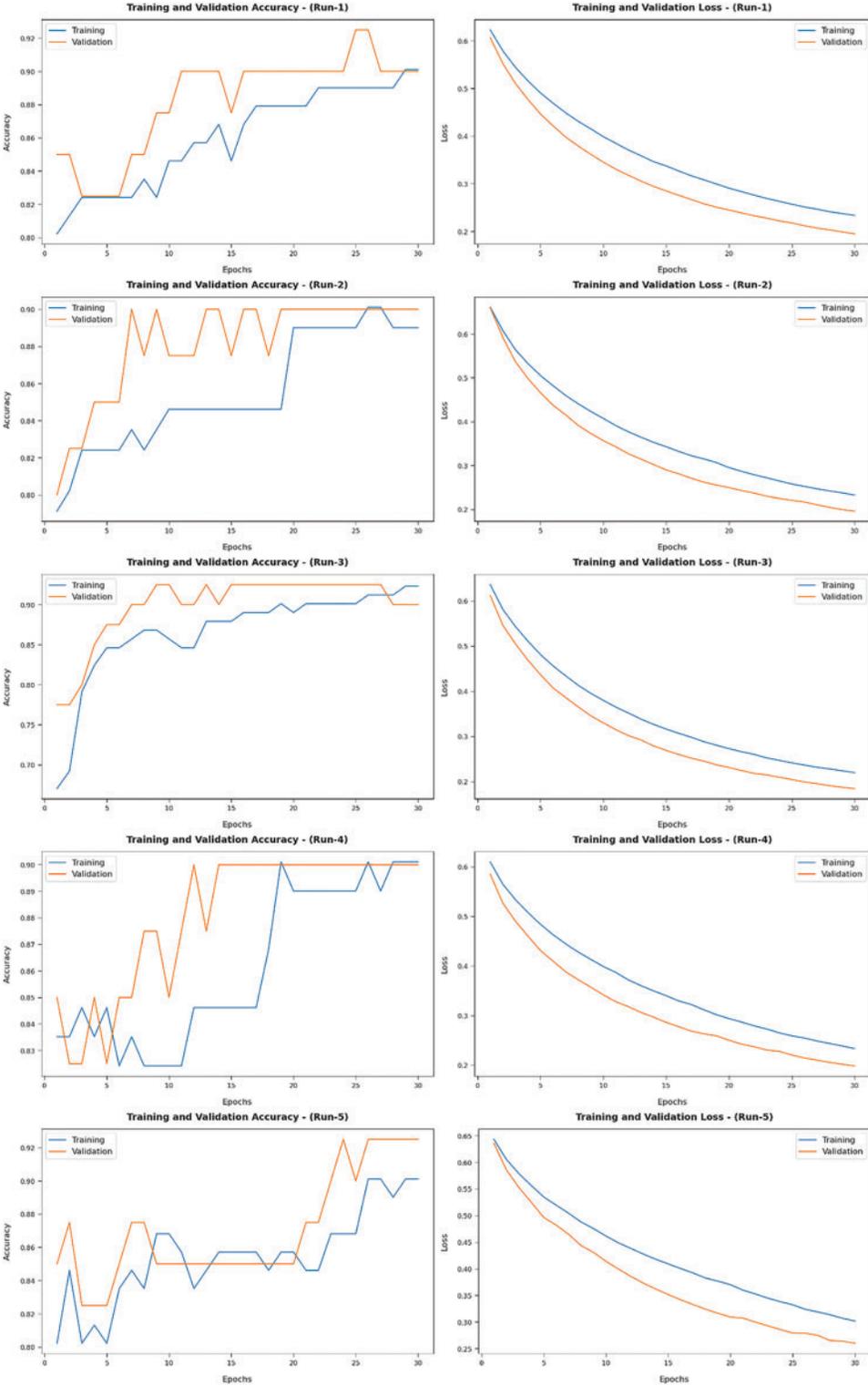
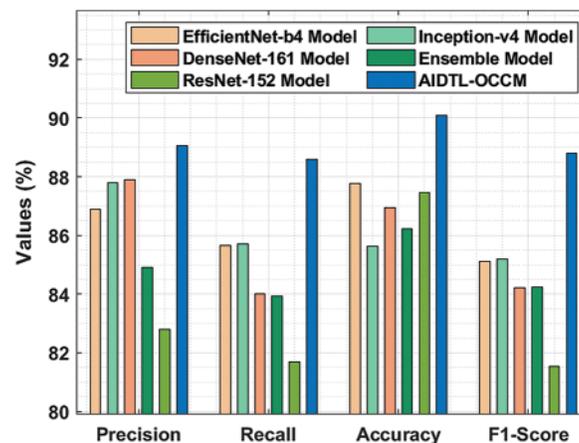


Figure 9: Training and validation accuracy/loss results of AIDTL-OCCM model

In order to demonstrate the superior outcomes of AIDTL-OCCM model, a brief comparison study was conducted and the results are shown in Tab. 3 and Fig. 10 [20]. The experimental values infer that Inception-v4 model produced the least classification outcomes. Followed by, DenseNet-161 and ensemble models resulted in slightly improved classification outcomes.

**Table 3:** Classification results of AIDTL-OCCM model and existing models

Methods	Precision	Recall	Accuracy	F1-score
EfficientNet-b4 model	86.90	85.65	87.76	85.12
Inception-v4 model	87.80	85.71	85.64	85.20
DenseNet-161 model	87.89	84.01	86.95	84.21
Ensemble model	84.92	83.93	86.23	84.24
ResNet-152 model	82.81	81.69	87.45	81.53
AIDTL-OCCM	89.05	88.60	90.08	88.81



**Figure 10:** Comparative oral cancer classification outcomes of AIDTL-OCCM model

At the same time, EfficientNet-b4 and ResNet-152 models accomplished reasonable performance. However, the proposed AIDTL-OCCM model produced superior results with maximum  $prec_n$ ,  $reca_l$ ,  $accu_y$ , and  $F1_{score}$  values such as 89.05%, 88.60%, 90.08%, and 88.81% respectively. After examining the above mentioned tables and figures, it is apparent that the AIDTL-OCCM model achieved maximum performance compared to other methods.

## 5 Conclusion

In current study, the researchers developed a novel AIDTL-OCCM model for effectual recognition and classification of oral cancer. In the presented AIDTL-OCCM technique, various functions are involved such as fuzzy-enabled pre-processing, DenseNet-169 feature extraction, AE-based classification, and COA-based parameter optimization. Furthermore, COA is employed to determine the optimal parameters involved in AE model. A wide range of experimental analyses was conducted on benchmark datasets and the results were investigated under several aspects. The results from extensive experimental analysis established the enhanced performance of the proposed AIDTL-OCCM model

over other recent approaches with a maximum accuracy of 90.08%. Therefore, AIDTL-OCCM technique can be applied as an efficient tool in oral cancer classification. In future, advanced DL models can be used for the classification of medical images to diagnose oral cancer.

**Funding Statement:** The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant number (RGP 1/322/42). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R161), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Umm Al-Qura University for supporting this work by Grant Code: (22UQU4310373DSR06).

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

## References

- [1] R. K. Gupta, M. Kaur and J. Manhas, "Tissue level based deep learning framework for early detection of dysplasia in oral squamous epithelium," *Journal of Multimedia Information System*, vol. 6, no. 2, pp. 81–86, 2019.
- [2] J. Rajan, S. Rajan, R. Martis and B. Panigrahi, "Fog computing employed computer aided cancer classification system using deep neural network in internet of things based healthcare system," *Journal of Medical Systems*, vol. 44, no. 2, pp. 1–10, 2019.
- [3] P. Yang, G. Liu, X. Li, L. Qin and X. Liu, "An intelligent tumors coding method based on drools," *Journal of New Media*, vol. 2, no. 3, pp. 111–119, 2020.
- [4] X. R. Zhang, J. Zhou, W. Sun and S. K. Jha, "A lightweight CNN Based on transfer learning for COVID-19 diagnosis," *Computers, Materials & Continua*, vol. 72, no. 1, pp. 1123–1137, 2022.
- [5] N. Das, E. Hussain and L. B. Mahanta, "Automated classification of cells into multiple classes in epithelial tissue of oral squamous cell carcinoma using transfer learning and convolutional neural network," *Neural Networks*, vol. 128, no. 2, pp. 47–60, 2020.
- [6] X. Liu, L. Faes, A. UKale, S. K. Wagner, D. J. Fu *et al.*, "A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: A systematic review and meta-analysis," *The Lancet Digital Health*, vol. 1, no. 6, pp. e271–e297, 2019.
- [7] H. Yang, E. Jo, H. J. Kim, I. Cha, Y. S. Jung *et al.*, "Deep learning for automated detection of cyst and tumors of the jaw in panoramic radiographs," *Journal of Clinical Medicine*, vol. 9, no. 6, pp. 1839, 2020.
- [8] B. Ilhan, P. Guneri and P. Wilder-Smith, "The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer," *Oral Oncology*, vol. 116, no. 7, pp. 105254, 2021.
- [9] S. Corbella, S. Srinivas and F. Cabitza, "Applications of deep learning in dentistry," *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, vol. 132, no. 2, pp. 225–238, 2021.
- [10] S. B. Khanagar, S. Naik, A. A. A. Kheraif, S. Vishwanathaiah, P. C. Maganur *et al.*, "Application and performance of artificial intelligence technology in oral cancer diagnosis and prediction of prognosis: A systematic review," *Diagnostics*, vol. 11, no. 6, pp. 1004, 2021.
- [11] R. A. Welikala, P. Remagnino, J. H. Lim, C. S. Chan, S. Rajendran *et al.*, "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer," *IEEE Access*, vol. 8, pp. 132677–132693, 2020.
- [12] P. Jeyaraj and E. S. Nadar, "Computer-assisted medical image classification for early diagnosis of oral cancer employing deep learning algorithm," *Journal of Cancer Research and Clinical Oncology*, vol. 145, no. 4, pp. 829–837, 2019.
- [13] F. Jubair, O. Al-karadsheh, D. Malamos, S. Al Mahdi, Y. Saad *et al.*, "A novel lightweight deep convolutional neural network for early detection of oral cancer," *Oral Disease*, 2021.

- [14] G. Tanriver, M. S. Tekkesin and O. Ergen, "Automated detection and classification of oral lesions using deep learning to detect oral potentially malignant disorders," *Cancers*, vol. 13, no. 11, pp. 2766, 2021.
- [15] H. Lin, H. Chen, L. Weng, J. Shao and J. Lin, "Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis," *Journal of Biomedical Science*, vol. 26, no. 08, 2021.
- [16] B. R. Nanditha, A. Geetha, H. S. Chandrashekar, M. S. Dinesh and S. Murali, "An ensemble deep neural network approach for oral cancer screening," *International Association of Online Engineering*, vol. 17, no. 02, pp. 121, 2021.
- [17] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan and A. Mittal, "Pneumonia Detection using cnn based feature extraction," in *2019 IEEE Int. Conf. on Electrical, Computer and Communication Technologies (ICECCT)*, Coimbatore, India, pp. 1–7, 2019.
- [18] M. U. Salur and I. Aydin, "A novel hybrid deep learning model for sentiment classification," *IEEE Access*, vol. 8, pp. 58080–58093, 2020.
- [19] M. Elahi, H. Ashraf and C. Kim, "An improved partial shading detection strategy based on chimp optimization algorithm to find global maximum power point of solar array system," *Energies*, vol. 15, no. 4, pp. 1549, 2022.
- [20] R. A. Welikala, P. Remagnino, J. H. Lim, C. S. Chan, S. Rajendran *et al.*, "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer," *IEEE Access*, vol. 8, pp. 132677–132693, 2020.