A Multi-task Learning Network for Automated Detection of Oral Epithelial Dysplasia

Abeer Aljuaid
Department of Computer Science
Umm Al-Qura University
Makkah, Saudi Arabia
aajuaid@uqu.edu.sa

Mai Almohaya
Department of Histopathology
Iman General Hospital
Riyadh, Saudi Arabia
almohayamai@gmail.com

Ashwaq Aljuaid
Department of Software Engineering
Umm Al-Qura University
Makkah, Saudi Arabia
aamjuaid@uqu.edu.sa

Mohd Anwar
Department of computer science
North Carolina A&T State University
Greensboro, NC, USA
manwar@ncat.edu

Abstract*—Oral cancer ranks as the sixth most common type of cancer worldwide, with 90% of cases in the form of oral squamous cell carcinoma (OSCC). OSCC has a high mortality rate, and early diagnosis can significantly increase the survival chances for patients. About 80% of OSCC is developed from Oral Epithelial Dysplasia (OED); thus, OED detection is vital in diagnosing and treating OSCC at its early stages. The conventional definition of OED encompasses sixteen criteria, including architectural and cytological features, meticulously evaluated through microscopic examination by expert oral pathologists. This manual detection is time-consuming and prone to human error, highlighting the need for precise automated detection of OED. However, the task of designing Computer-Aided Diagnosis (CAD) for OED detection is known to be challenging because each criterion in OED requires a particular medical image processing task to be performed. To address this issue, we proposed a novel multi-task learning network to combine semantic segmentation and classification to detect OED based on architectural and cytological characteristics. Our approach is the first study that jointly trained semantic segmentation and classification on a single network for automated OED detection. We developed two new frameworks called VGG16-Unet and InceptionV3-Unet, which were designed based on classic U-Net with ImageNet pre-trained VGG16 and InceptionV3 encoder components, as well as traditional classifier models attached at the end of the encoder branch of U-Net. The efficacy of the proposed models was rigorously verified through quantitative metrics and visualization. The experimental results demonstrate that our novel models perform superiorly on both classification and segmentation tasks for detecting OED by achieving a high accuracy of 90.6%, a sensitivity of 85%, a specificity of 94.6%, a precision of 92%, an F1-score of 88.4% in classification and 84.4% of Jaccard in segmentation tasks.

***Keywords*—Oral Epithelial Dysplasia (OED), U-Net, Classification, Semantic Segmentation, Computer-Aided Diagnosis (CAD), and histopathological images.**

I. INTRODUCTION

Oral cancer has been reported to result in high mortality rates [1], [2]. Oral Squamous Cell Carcinoma (OSCC) constitutes around 90% of all oral cancer cases [3], [2], [4]. OSCC's high mortality rate can be attributed to its diagnosis at advanced stages, and early detection has shown to reduce the rate by up to 90% [5], [2], [6]. About 80% of OSCC is developed from a potentially malignant disorder called Oral Epithelial Dysplasia (OED) [4], [7], [6]. Therefore, detecting OED is essential for the early diagnosis of OSCC.

The conventional method of diagnosing and grading the OED involves pathologists examining and analyzing sixteen criteria under the microscope [7], with half representing architectural disturbance of epithelium layers, and the other half cytological abnormalities. However, the manual method of OED detection is time-consuming and requires expertise [8], [9]. Therefore, a computer-aided diagnosis system (CAD) for the OED is necessitated.

Automated OED detection can reduce the workload on oral pathologists while delivering accurate diagnoses and speeding up the diagnostic procedure [10]. Machine learning (ML) and Deep Learning (DL) technologies have been widely adopted for CAD with accuracy and effectiveness [1], [11], [12]. Although classical ML techniques have proved accurate in detecting oral carcinomas, they are limited by their inability to analyze vast amounts of data [13], [14]. Therefore, this study explored the efficacy of deep learning-based techniques in detecting, analyzing, and segmenting OED.

However, designing a CAD system for the OED diagnosis is challenging due to the requirement to detect the architectural and cytological features in histopathological images. Each feature is identified by a particular medical image processing task. To address this issue, we proposed a novel approach that combines the segmentation and classification tasks within a single network using both architectural and cytological features. We applied the U-Net [15]for the semantic segmentation of the epithelium layer, linking a classifier model at the end of the encoder of the U-Net to distinguish the dysplastic tissue from normal tissue. The connection between the classifier and the encoder branch allowed the sharing of high-level features between the two tasks, thus enhancing detection accuracy.

In addition, the dataset used for this study is small; thus, we adopted the U-Net with pre-trained ImageNet networks, particularly VGG16 [16] and Inception-v3 [17], as encoder backbones. Transfer learning techniques have been shown to improve the models' performance and generalization while reducing training time when training a complex model with a small dataset [18], [19], [14], [8], [9].

2835-5776/24/$31.00 ©2024 IEEE
DOI 10.1109/IRI62200.2024.00018

The contributions of the proposed study are as follows.

- We collected, annotated, and extracted Regions of Interest (ROIs) in collaboration with Mount Sinai Health System.
- A novel end-to-end multi-task learning framework was developed for the segmentation and classification of OED using cytological and architectural features.
- We developed a U-Net-based model architecture using pre-trained VGG16 and inception-v3 as backbones, respectively.

II. Related Work

This section summarizes literature that used ML and DL algorithms to diagnose OED, highlighting the limited number of studies focusing on both cytological and architectural features.

Automated OED Detection based on OED's Cytological Features. Cytological features represent nuclear disturbance, including nuclear-cytoplasmic ratio, shape, and size.
A custom CNN framework is proposed in [13] for classifying dysplastic cell images from the oral squamous epithelium layer, achieving 90% for precision, recall, and F1-score.

Other studies applied deep learning algorithms for nuclei semantic segmentation of OED in mice tongue histopathological images [20], [21] using Mask R-CNN to segment nuclei of oral dysplasia, and they had an accuracy of 89.5% and 88%, respectively.

Automated OED Detection based on OED's Architectural Features. Architectural characteristics used to diagnose OED are about the changes in the structure of the stratified epithelium. A study by [22] used a DeepLabV3 approach to classify OED based on the morphometric features, achieving an accuracy of 57%.

Automated OED Detection based on OED's cytological and architectural Features. In a recent study, researchers suggested segmenting the epithelium layers and the nuclei would detect OED's cytological and architectural features. A study by [23] proposed HoVer-Net+ for simultaneous nuclear instance segmentation and intra-epithelial layer semantic segmentation using H&E WSI for OED. Consequently, they achieved an F1-score for nuclear and layer segmentation of 0.73 and 0.82, respectively. Likewise, [24] proposed a model to segment the epithelium layer and nuclear. As a result, they achieved an F1-score of 93.5% and 94.5% for the entire epithelium and nuclear segmentation, respectively.

III. Methodology

We proposed a novel multi-task learning network that segments the oral tissues using the architectural features and classifies them using the cytological features into dysplastic tissue and normal. Therefore, the proposed network combines classification and segmentation in an end-to-end Deep Convolutional neural network model (DCNN) that takes histopathological images as input and provides two outputs, including the segmentation map for the epithelium architecture and the classification probability. Therefore, our proposed network involves an encoder, a decoder for segmentation, and a classifier network for classification. Our methodology involves five steps: 1) dataset acquisition and

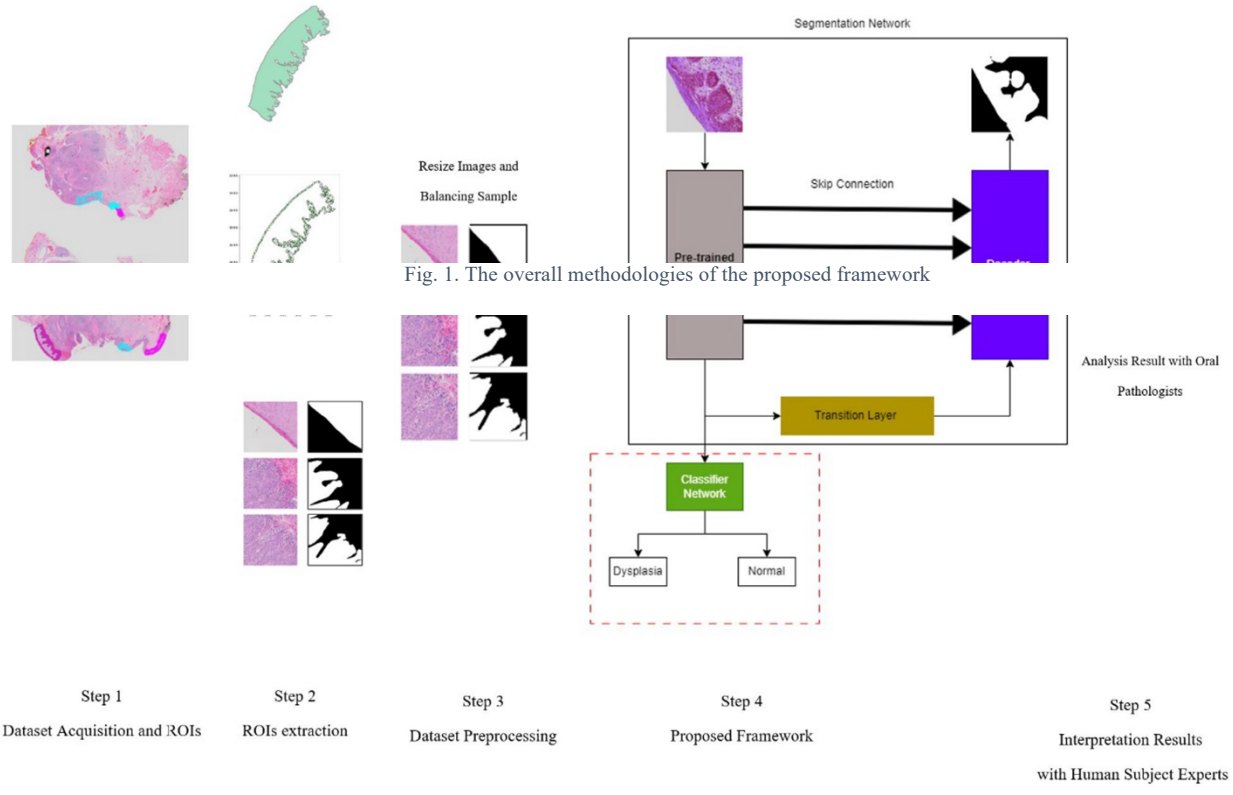

Fig. 1. The overall methodologies of the proposed framework

ROIs selection, 2) ROIs extraction, 3) data preprocessing, 4) proposed framework, and 5) interpretation of results with human subject experts. Fig. 1 illustrates the overall methodology. In this Section, we will discuss these steps in detail.

A. Data Acquisition and ROIs selection

A dataset of 151 whole slide images (WSI) of oral cancer cavity (OCC) was collected from the Erie County Medical Center. The dataset was scanned by an Olympus scanner at various resolutions of 40x, 20x, and 10x, and it was stained with hematoxylin and eosin. The dataset was stored in Digital Slide Archive (DAS) [25], which is a platform used for storing, managing, visualizing, and annotating whole-slide images.

Oral pathologists annotated the ROIs, consisting of polygons delineating the architecture of dysplasia and mucosa (normal tissue). We created a dataset consisting of 100 samples of oral dysplasia and 63 images of normal mucosa tissue. The dataset has a variety of challenges: 1. the dataset consists of varying image sizes due to differences in ROI shapes; 2. some samples contain tumor regions; 3. the dataset is small to train supervised learning networks; 4. the dataset is imbalanced.

B. ROIs Extraction

The second step involved using shapely polygons' bound to extract the annotated ROIs by the oral pathologists from the previous step and create their corresponding masks. To overcome the dataset challenges explained in section 3.1, we randomly generated points at polygon boundaries for extraction of the annotated ROIs and their masks with a patch size of 2048×2048×3 based on the pathologist's recommendation.

To eliminate undesired tissue (tumor tissue) involved in our ROIs, we applied multi-pixel labels instead of binary labels with foreground/background distinction (Fig 2; the red arrows point to dysplasia tissue, whereas the blue arrow points to the tumor tissue.).

The patch methods increased the number of samples from 63 to 719 for mucosa and from 100 to 580 for dysplasia. All samples have a shape of 2048×2048×3 and do not include any tumor regions, and all patches involve ROIs and background parts.

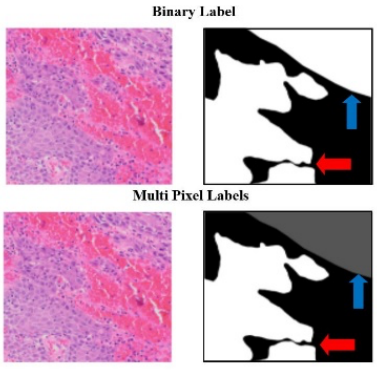

Fig. 2. The way to identify the tumor tissue.

In addition, all regions were validated by the oral pathologists.

C. Data Preprocessing

The third step of the framework involves data preprocessing in two phases: data balancing and resizing. Our dataset is imbalanced by having more normal samples than dysplastic samples, which can significantly affect deep learning networks' performance and generalization [26]. To solve this problem, we downsampled the number of normal samples to the number of dysplastic images.

Trained deep learning networks with large image shapes are computationally expensive, and large image resolution requires large Graphic Processing units (GPU) and longer training time [27]. Moreover, the input shape of the most pretrained CNN is smaller than our image shape. Therefore, we resized our extracted patches to the size of 512×512×3.

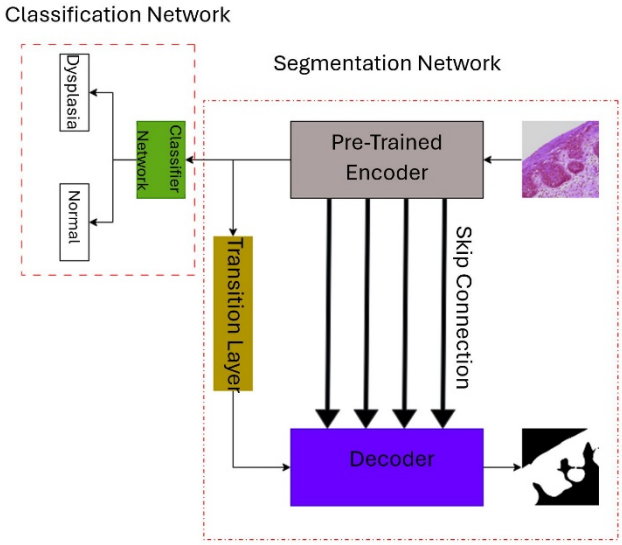

Fig. 3. The proposed framwork.

D. Proposed Framework

The proposed network jointly trains classification and segmentation tasks in a single network. The models' input is histopathological images and produces a segmentation map for epithelium layer architecture and a class label (i.e., dysplasia or normal). Therefore, our proposed network involves an encoder, a decoder for semantic segmentation, and a classifier network attached at the end of the encoder (see Fig. 3).

1) Semantic Segmentation Network: We utilize U-Net for our multi-task learning network due to its significant performance in medical image segmentation. The U-Net consists of three parts forging symmetrical architecture: (1) a contracting (encoding) path, (2) an expanding (decoding) path, and (3) skip connections between them. More detiles about U-Net can be found in [15]. However, the U-Net requires extensive data for robust performance and generalization, and our dataset is small. Therefore, we

applied the U-Net with the ImageNet pre-trained encoder, particularly pre-trained VGG16 and pre-trained InceptionV3, and we called them as the VGG16-UNet and InceptionV3-UNet, respectively.

2) Classification Network: We exploited the encoder's ability to extract the input features by attaching a classifier model for the classification task at the end of the VGG16-UNet and InceptionV3-Unet encoders, thus allowing for sharing of those features between the two tasks. After removing the top layers of both networks, we added global average pooling followed by a dropout layer. The output is then passed to a fully connected layer with 1024 neurons. The last layer applied the sigmoid activation function for the binary classification. Fig 4 represents the outline of the traditional classifier.

E. Interpretation Results with Human Subject Experts

Our method employs deep learning for detecting oral precancer lesions (OED), however, it requires a pathologist's expertise for analysis.

IV. Results and Discussion

In this Section, we discussed the implementation and results of the proposed novel multi-task frameworks for semantic segmentation and classification of OED. We explored the environment used to implement our networks. We explained the dataset utilized for this experiment. We discussed the training protocols and the hyperparameters applied to set up the proposed networks. We investigated the results of our frameworks by using quantitative analysis and visualization matrices and by comparing with the proposed methods by [24].

A. Environment

We run our experiments on a personal computer (PC) with an Intel® Core™ i7-7500U CPU running at 2.70 GHz using 16 MB of RAM. The PC is running on Windows 10 Home Edition. We implemented our proposed models in Python using Keras and TensorFlow libraries.

B. Dataset

We evaluated our proposed networks using 151 H&E WSI of oral cancer cavity (OCC), containing both normal and dysplasia tissue in ROIs. After ROIs extraction, we generated a dataset containing 1299 histopathological images, where 719 images belong to the normal class, and 580 samples are dysplasia.

For the semantic segmentation, we classified each pixel in the image as the foreground or the background. We balanced the number of classes in the training dataset for the classification task; meanwhile, we exploited the entire dataset for the semantic segmentation task.

We split the dataset in the ratio of 80:10:10 for training, validation, and testing. Table 1 provides more details about the number of samples for the classification and segmentation tasks.

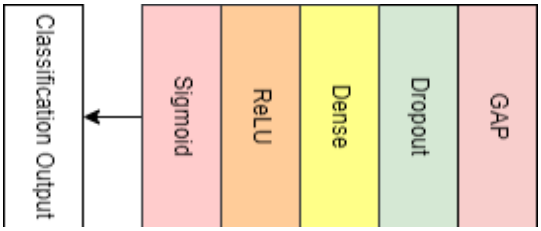

Fig. 4. The simple classifier architecture.

C. Training Protocols and Hyperparameters Setting

We performed the classification and segmentation tasks using transfer learning. First, we built a classification network and trained it on a balanced classification dataset. Then, we used the classification network as the pre-trained encoder and frozen it to maintain the classification performance. After that, we added a decoder branch and trained it for semantic segmentation using the entire dataset. This way, we have a single network for both tasks, and they share common high-level features.

We utilized an early stopping method described in [28] to avoid overfitting. We also exploited the learning rate schedules instead of the fixed learning rate. In the end, we trained the proposed model with 200 iterations, an Adam optimizer, and a batch size of 8.

D. Results

We evaluated the classification task performance using confusion matrix, ROC, accuracy, precision, recall (sensitivity), specificity, AUC, and F1-score. We used F1-score, precision, recall, and Jaccard for the semantic segmentation task validation. More information about these performance metrics can be found in [26].

1) Classification Task Result: The Inception-Unet model surpasses the VGG16- Unet model for differentiating mucosa from dysplasia by achieving a high accuracy of 90.6%, a sensitivity of 85%, a specificity of 94.6%, a precision of 92%, an F1-score of 88.4%. The Inception-Unet correctly classified 71 out of 75 mucosae and 46 out of 54 dysplasia. In contrast, the VGG16-Unet misclassified 12 and 10 mucosae and dysplasia, respectively.

TABLE I. NUMBER OF SAMPLES FOR CLASSIFICATION TASK AND SEGMENTATION TASK.

	Classification Task			**Segmentation task**		
	Total	***Dysplasia***	***Normal***	***Total***	***Dysplasia***	***Normal***
Training	932	466	466	1041	467	574
Validation	129	59	70	129	59	70
Testing	129	54	75	129	54	75

Fig. 5. The ROC curve of the InceptionV3-Unet, and VGG16-Unet models.

TABLE II. PERFORMANCE COMPARISON FOR THE CLASSIFICATION TASKS

Models	Accuracy	Recall (Sensitivity)	Specificity	Precision	F1-score
Inception-Unet	**0.9069**	**0.8518**	**0.9466**	**0.92**	**0.8846**
VGG16-Unet	0.8294	0.8148	0.8400	0.7857	0.7999
Proposed model in [24]	-	0.793	-	0.786	0.786

The differences between the false positive and negative rates of the Inception-Unet and the VGG16-Unet affect the performance of the VGG16- Unet. Therefore, the VGG16-Unet performed less than the Inception-Unet by 7.7% in accuracy, 3.6% in sensitivity, 10.6% in specificity, 13.5% in precision, and 8.5% in F1-score. Table 2 summarizes these values for the Inception-Unet and the VGG16-Unet models. Moreover, the Inception-Unet outperforms the VGG16-Unet in terms of ROC and AUC. The Inception-Unet obtained 90% of AUC, whereas the VGG16-Unet achieved an AUC of 83%. Fig 5 shows the ROC curves of both networks.

2. Segmentation Task Result. The VGG16-Unet model achieved the highest semantic segmentation with 90.0% of F1-score, 84.4% of Jaccard, 92% of recall, and 90.9% of precision, whereas the inceptionV3-Unet model obtained the lowest performance of 89.7%, 82%, 89.6%, and 90.7% of F1-score, Jaccard, recall, and precision, respectively. Table 3 summarizes the semantic segmentation result for all models. Moreover, the VGG16-Unet model provides precisely predicted masks more than the Inception-Unet. The VGG16-Unet draws an accurate contour that detects the architecture criteria for the epithelium layers. Whereas the inceptionV3-Unet model has the worst mask detection and the highest rate of false negatives (blue circle, fig 6).

3. Performance Comparison with Other Methods. We compared our model with the proposed model in [24]. In [24], deep learning technique is utilized to detect the cytological and architectural features of Oral Epithelial Dysplasia (OED). However, their approach involved three phases: (1) semantic segmentation of the oral epithelium layer and nuclei, (2) feature extraction from the segmented nuclei, and (3) OED grade prediction based on these features. They trained and tested their framework using a private dataset of 32,384 H&E patches of OED.

For the classification task, the performance of our proposed models exceeded the proposed model in [24] by obtaining the highest recall, precision, and F1-score, as Table 2 shows. However, for the segmentation task, the model performance in [24] surpassed the VGG16-Unet by 2.6 % (table 3) in F1-score, perhaps because they had access to a dataset six times larger than the one we used. The significant performance of our proposed frameworks in the classification task and the comparable results in the segmentation task demonstrated that training both tasks jointly in a single framework improved the accuracy of OED detection.

.

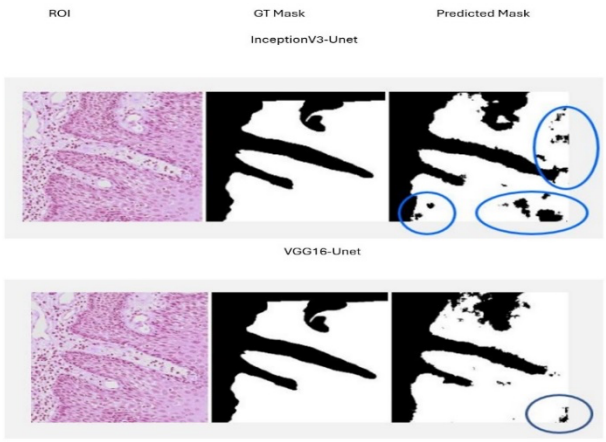

FIG. 6. EXAMPLES OF A PREDICATED MASK; GT IS A GROUND TRUTH MASK; BLUE CIRCLES SHOW FALSE NEGATIVES

TABLE III. PERFORMANCE COMPARISON FOR THE SEMANTIC SEGMENTATION TASK

Model	F1-score	Jaccard	Recall	Precision
Inception-Unet	0.89728	0.82271	0.89691	0.90744
VGG16-Unet	0.90996	**0.84346**	0.92277	0.90921
Proposed model in [24]	**0.935**	-	**0.933**	**0.962**

V. Conclusion

Oral cancer, particularly Oral Squamous Cell Carcinoma (OSCC), has a high mortality rate and requires early detection to improve survival rates. Oral Epithelial Dysplasia (OED) is the first stage of OSCC; thus, the diagnosis of OED is necessary. The traditional manual detection methods are slow and prone to errors, highlighting the need for automated systems. Our study presents an innovative multi-task learning network that integrates semantic segmentation and classification to identify OED accurately using architectural and cytological characteristics. We developed VGG16-Unet and InceptionV3-Unet models, which showed high performance with 90.6% accuracy in classification and an 84.4% Jaccard index in segmentation. These results demonstrate the potential of our models to enhance early detection and treatment of OSCC.

Despite significant results, the primary limitation of our study is the small dataset size, which is insufficient for the robust performance and generalization required by complex deep learning networks.

References

[1] A. B. Krishna, A. Tanveer, P. V. Bhagirath, and A. Gannepalli, "Role of artificial intelligence in diagnostic oral pathology-A modern approach," *Journal of Oral and Maxillofacial Pathology: JOMFP,* vol. 24, no. 1, p. 152, 2020.

[2] T. Radhika, N. Jeddy, S. Nithya, and R. Muthumeenakshi, "Salivary biomarkers in oral squamous cell carcinoma–An insight," *Journal of oral biology and craniofacial research,* vol. 6, pp. S51-S54, 2016.

[3] K. Awan, "Oral Cancer: Early Detection is Crucial," (in eng), *J Int Oral Health,* vol. 6, no. 5, pp. i-ii, Sep 2014.

[4] F. R. Pires, M. E. Barreto, J. G. Nunes, N. S. Carneiro, A. B. Azevedo, and T. C. Dos Santos, "Oral potentially malignant disorders: clinical-pathological study of 684 cases diagnosed in a Brazilian population," (in eng), *Med Oral Patol Oral Cir Bucal,* vol. 25, no. 1, pp. e84-e88, 01 01 2020, doi: 10.4317/medoral.23197.

[5] M. W. Ho *et al.*, "Outcomes of oral squamous cell carcinoma arising from oral epithelial dysplasia: rationale for monitoring premalignant oral lesions in a multidisciplinary clinic," (in eng), *Br J Oral Maxillofac Surg,* vol. 51, no. 7, pp. 594-9, Oct 2013, doi: 10.1016/j.bjoms.2013.03.014.

[6] A. Villa and A. Gohel, "Oral potentially malignant disorders in a large dental population," *Journal of Applied Oral Science,* vol. 22, pp. 473-476, 2014.

[7] K. Ranganathan and L. Kavitha, "Oral epithelial dysplasia: Classifications and clinical relevance in risk assessment of oral potentially malignant disorders," *Journal of oral and maxillofacial pathology: JOMFP,* vol. 23, no. 1, p. 19, 2019.

[8] C. Mazo, J. Bernal, M. Trujillo, and E. Alegre, "Transfer learning for classification of cardiovascular tissues in histological images," *Computer methods and programs in biomedicine,* vol. 165, pp. 69-76, 2018.

[9] M. Talo, "Automated classification of histopathology images using transfer learning," *Artificial Intelligence in Medicine,* vol. 101, p. 101743, 2019.

[10] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE reviews in biomedical engineering,* vol. 2, pp. 147-171, 2009.

[11] Y. Ariji *et al.*, "Contrast-enhanced computed tomography image assessment of cervical lymph node metastasis in patients with oral cancer by using a deep learning system of artificial intelligence," *Oral surgery, oral medicine, oral pathology and oral radiology,* vol. 127, no. 5, pp. 458-463, 2019.

[12] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *Ieee Access,* vol. 6, pp. 9375-9389, 2017.

[13] R. K. Gupta, M. Kaur, and J. Manhas, "Cellular Level Based Deep Learning Framework for Early Detection of Dysplasia in Oral Squamous Epithelium," in *Proceedings of ICRIC 2019*: Springer, 2020, pp. 137-149.

[14] S. Lu, Z. Lu, and Y.-D. Zhang, "Pathological brain detection based on AlexNet and transfer learning," *Journal of computational science,* vol. 30, pp. 41-47, 2019.

[15] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015: Springer, pp. 234-241.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556,* 2014.

[17] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.

[18] K. S. Beevi, M. S. Nair, and G. Bindu, "Automatic mitosis detection in breast histopathology images using Convolutional Neural Network based deep transfer learning," *Biocybernetics and Biomedical Engineering,* vol. 39, no. 1, pp. 214-223, 2019.

[19] S. Xu *et al.*, "An early diagnosis of oral cancer based on three-dimensional convolutional neural networks," *IEEE Access,* vol. 7, pp. 158603-158611, 2019.

[20] A. B. Silva, A. S. Martins, L. A. Neves, P. R. Faria, T. A. Tosta, and M. Z. do Nascimento, "Automated nuclei segmentation in dysplastic histopathological oral tissues using deep neural networks," in *Iberoamerican Congress on Pattern Recognition*, 2019: Springer, pp. 365-374.

[21] A. B. Silva *et al.*, "Segmentation of Oral Epithelial Dysplasias Employing Mask R-CNN and Color Normalization," in *2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020: IEEE, pp. 2818-2824.

[22] R. S. Bashir *et al.*, "Automated grade classification of oral epithelial dysplasia using morphometric analysis of histology images," in *Medical Imaging 2020: Digital Pathology*, 2020, vol. 11320: SPIE, pp. 245-250.

[23] A. J. Shephard *et al.*, "Simultaneous nuclear instance and layer segmentation in oral epithelial dysplasia," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 552-561.

[24] N. Azarmehr, A. Shephard, H. Mahmood, N. Rajpoot, and S. A. Khurram, "A Neural Architecture Search Based Framework for Segmentation of Epithelium, Nuclei and Oral Epithelial Dysplasia Grading," in *Annual Conference on Medical Image Understanding and Analysis*, 2022: Springer, pp. 357-370.

[25] D. A. Gutman *et al.*, "The Digital Slide Archive: A Software Platform for Management, Integration, and Analysis of Histology for Cancer Research," (in eng), *Cancer Res,* vol. 77, no. 21, pp. e75-e78, 11 01 2017, doi: 10.1158/0008-5472.CAN-17-0629.

[26] A. Aljuaid and M. Anwar, "Survey of supervised learning for medical image processing," *SN Computer Science,* vol. 3, no. 4, p. 292, 2022.

[27] K. da Rocha, J. C. Bermudez, E. R. Rivero, and M. H. Costa, "A pathology-based machine learning method to assist in epithelial dysplasia diagnosis," *Research on Biomedical Engineering,* vol. 38, no. 3, pp. 989-1002, 2022.

[28] L. Prechelt, "Early stopping—but when?," *Neural networks: tricks of the trade: second edition,* pp. 53-67, 2012.