2023 10th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON) | 979-8-3503-8247-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/UPCON59197.2023.10434514

Histopathological Image based Oral Squamous Cell Carcinoma Classification Using Deep Network Fusion

Kumar Ankit
Department of Computer Science and Engineering
Graphic Era deemed to be university
Dehradun, India.
ankitmarvel22102004@gmail.com

Vidit Kumar
Department of Computer Science and Engineering
Graphic Era deemed to be university
Dehradun, India.
viditkumaruit@gmail.com
https://orcid.org/0000-0002-6157-7562

***Abstract*— Oral cancer represents a pressing public health concern, posing a substantial threat to individuals due to its impact on critical anatomical structures, including the oral cavity, tongue, and lips. The continual evolution of technology has played a pivotal role in advancing disease diagnostics, and this holds true for oral cancer detection. Utilizing histopathology images as a diagnostic aid enhances the accuracy of disease identification. When coupled with deep learning methodologies, the potential for significantly improved disease prediction becomes evident, promising practical benefits in real-world scenarios. Our research dives deep into these cutting-edge techniques, systematically exploring a range of possibilities to enhance the predictive capacity for this disease. In this regard, we propose a deep fusion-based approach that combines three different models, namely Resnet50, Efficientnet-b0, and Convnext-tiny, exhibit noteworthy advancements in predictive accuracy. Our findings strongly indicate that the fused model outperforms the individual models and other methods in prediction accuracy, thus having substantial potential to advance the field of its analysis in the future.**

Keywords*— oral squamous cell carcinoma, convolutional neural network, deep fusion, fusion network, oral cancer

I. INTRODUCTION

Oral cancer continues to be an important global health issue. It is a subgroup of malignancies of the head and neck region, and it mostly affects the cells of the squamous membrane which encompass the mouth and its neighbouring parts. One of the most prevalent forms of oral cancer is Squamous cell carcinoma (OSCC) of the oral cavity. Based on information from the Global Cancer Observatory (GCO), there were an estimated 377,713 new cases of OSCC reported yearly around the globe in 2020, with Asia reporting the highest number of cases (248,360), subsequently followed by Europe (65,279) and North America (27,469)[1]. The total number of cases of OSCC throughout the course of five years was nearly one million (959,248) and followed the same developments [2]. The GCO anticipates that by two decades in future, the number of cases of OSCC will increase by up to 40%, increasing death in concordance. The major causes of oral cancer include Smoking, tobacco use, betel nut chewing, alcohol use, poor nutrition and even genetic factors. Most dangerous risk factors that harms a person from oral cancer which add it to factor by 35% are smoking and alcohol consumption that becomes lethal at heavy dose [3]. Also the taking of Areca nut strongly shows its ties with oral cancer specially OSCC in the Pacific and East Asian region. [2].

In spite of the importance of early detection, most patients are diagnosed at the very end of their illness, where they have a dismal prognosis. Oral cancer's appearance makes it difficult to diagnose. the abnormal cells in the tissue, and its analysis or limit, due to the stated reason, selective therapy of this illness is insufficient. Conventional anamnesis and clinical exam are used to diagnose oral cancer, and a more common method is coupled with visualisation and haematoxylin-eosin histopathological examination. [4, 5]. The most prevalent indication of a cancer is a ulcer that can sometimes causes immeasurable pain. Some of the most common indications are Unhealing white and red patches on the lips,an abnormal growth around tongue, gums, or cheeks; a lump in the mouth; issues ingesting; jaw enlargement; difficulty communicating; and long-term pain in the throat [6]. In the future of healthcare, having a deeper understanding of innovations in technology like artificial intelligence may be beneficial [7]. When detecting tumors lesions, the use of Artificial intelligence can minimize the work needed for monitoring and big data samples [17]. Oncology is employing artificial intelligence because of its capacity to increase the success rate for oral cancer monitoring. [8].

Artificial intelligence, in the view of specialists from the University of Sheffield in the United Kingdom, aids in accurate, dependable, and realistically predicting oral cancer. Thus, many methods based on machine learning techniques were explored. For instance, researchers in [9] used ANNs, employed texture analysis to classify oral cancer lesions into six groups but to create a more reliable and predictive model which can enhance in this domain, additional studies with larger data is needed. In OC histopathology images, previous techniques are combined with hybrid features derived from ResNet-18 and AlexNet models for a more accurate diagnosis. One of its benefits is that it trains dataset quickly [10], but requires a large set of data and more advanced and more computational power for further work. Some researchers in [11] used Naïve Bayes in assessing lymph node status

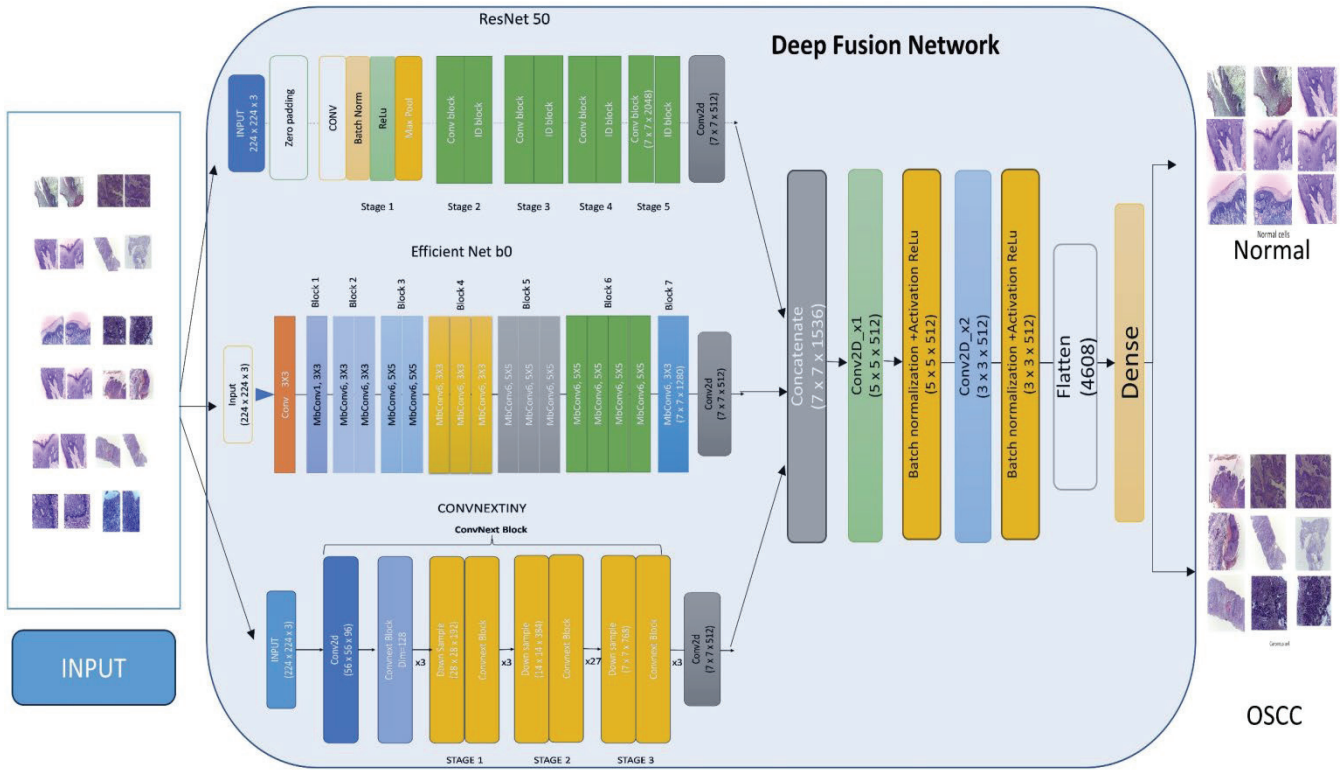

Fig. 1. Proposed work for oral cancer diagnosis

in OSCC using contrast-enhanced CT imaging. Although they achieved a high accuracy, Naïve Bayes relies on the assumption of feature independence, limiting its applicability in complex scenarios. Moreover, researchers in [12] conducted an experiment to accurately predict tumor behavior using machine learning algorithms in hope to improve clinical experiment to accurately predict tumor behavior using machine learning algorithms in hope to improve clinical outcome prediction in oncology research. They used four machine learning algorithms and trained them with 34 features and results were recorded and as a result Decision Tree gave most satisfactory result in identifying TP(true positive) progressive disease with total 70.59% accuracy over other models. According to researchers in [13], there were total 41 studies that have used these techniques in the diagnosis of OSCC. So, there is still possibility of using more advanced technologies such as deep learning models to further increase its accuracy and application to oral cancer.

With recent advancements in the field of artificial intelligence, deep learning is very important in the medical field [18][30-32]. When compared to older approaches, recent deep learning-driven study has resulted in promising outcomes.. Researchers in [14] acquired the images of oral cavities and applied a different model for diagnosis, and they find that it performs well with accuracy of 84.3% compared to Resnet50 and Densenet 169 models. Researchers in [15] employed R-CNN and DenseNet121 model to train a dataset consisting of 490 total images of OSCC. They then evaluated the performance of both systems on a set of 140 test images. Remarkably, the DenseNet121 model outperformed R-CNN. The researchers in [16] have conducted experiments involving hybrid combinations of Convolutional Neural Networks (CNNs), including GoogLeNet, ResNet101, and VGG16 models, which demonstrated varying performance depending on the specific conditions of the study. Furthermore, their efforts extended to image quality enhancement, which played a crucial role in significantly improving the overall results of their work.

In summary, numerous studies have conducted to improve the diagnosis of oral cancer, and there is considerable optimism for further advancements in this area. One potential avenue for improvement lies in the exploration of fusion of multiple CNN models, as we have investigated in our proposed research.

II. Methodology

The proposed work as shown in fig. 1 is to diagnose oral cancer with deep fusion technique. In this we are going to use different types of CNNs in incorporation of more layers and predict the outcome of inputted data with fusion of these different CNNs types. The different Convolution neural network used are explained below:

A. Resnet 50:

Resnet which was proposed by author in [19], is a system where the layers are reformulated as learning residual functions with regard to the inputs of the layers. This structure can produce results that are better than previous networks by increasing layers that enhances depth and improving object classification accuracy. Resnet 50, which is 50 neural network layer is used in this work.

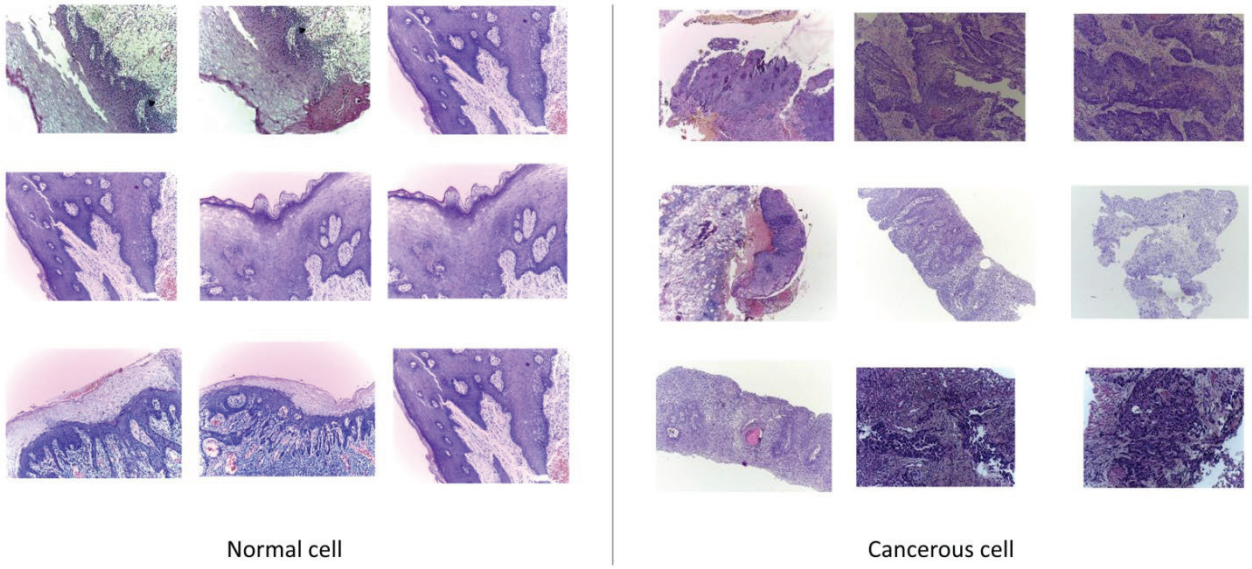

Fig. 2. Oral histopathological images of normal and cancerous cell

B. *Efficientnet b0*

Efficientnet which was proposed by authors in [20], is an architectural and scaling technique for convolutional neural networks that uses some user defined values to evenly map the dimensions correctly. In contrast to conventional practise, which varies the scaling of these variables, the scaling approach as mentioned earlier equally distribute the dimensions(width, depth) with some predefined constant value as given by user. We are using Efficientnet b0 architecture in our proposed system.

C. *Convnextiny*

Convnextiny is a tiny sized convnext model which was proposed by authors in [22].This is light weight and efficient model with much better accuracy, with top accuracy of 81.3% ,it became popular and significant tool for image classification. It was significant improvement in speed, accuracy and other parameters which made it a good choice.

D. *Deep Fusion Network*

In our experiment we have made a slight change before concatenation of three models, we have added three layers, Conv layer of 1X1 sized kernel with 512 kernels, Batch normalization and activation relu just after the individual output of three models. We have also added these layers after the concatenation of these models, and Convolution layer of 3X3 sized kernel with 512 kernels. After the first Convolution layer the output of Batch normalization is (5x5x512) and that of relu layer is (5x5x512) and after the second convolution layer which is after these two layers, we again added these three layers which outputs to (3x3x512). Then a flatten layer followed by dense layer is added for classification.

III. Experiments

A. *Dataset*

The dataset which is used in this experiment is publicly made available by authors in [23]. The datasets contain oral histopathological images of normal and cancerous cell under 100x and 400x lens. There is total 1224 images which are further divided into three folds in which each fold comprises of training and testing sets in 70 to 30 ratio. Dataset distribution among folds are shown in Table I.

B. *Evaluation Parameters*

We have assessed our model using metrics like recall, accuracy, precision, and F1-score. of three folds and training each fold with single model and fusion model and testing it and recorded the result with different parameters.

$$Accuracy = \frac{Number\ of\ Correct\ predictions}{Total\ number\ of\ predictions\ made} \quad (1)$$

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive} \quad (2)$$

$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative} \quad (3)$$

$$F1 = 2 * \left(\frac{1}{\left(\frac{1}{precision}\right) + \left(\frac{1}{recall}\right)} \right) \quad (4)$$

Algorithm 1

Data size = *D_size*
batch size=b_s=32
max_epoch=60
mini_batch=m_b
learning_rate=l
$\max\_itr = \frac{(D\_size)}{b\_s}$

1. INPUT: *m_b*, Proposed model $f_{FCN}()$, b_s, *max_epoch*, *max_itr*, *D_size*.
2. OUTPUT: Trained Model.
3. for *epoch* 1 to *max_epoch:*
4. *m_b*[*epoch*]=extract image (*D_size*, *b_s*)
5. Apply Data augmentation to *m_b*[*epoch*]
6. for *itr* 1 to *max_itr:*
7. *feed m_b[epoch] to* $f_{FCN}()$
8. *Compute cross entropy loss L*
9. *optimize* f_{FCN} *with objective to minimize loss l*
10. *end for*
11. *end for*

TABLE I. DATASET FOLD SPLITING

Dataset Split	Traning (70%)	Testing (30%)
Fold1	First 70% from each class	Last 30% from each class
Fold2	Last 70% from each class	First 30% from each class
Fold3	Middle 70% from each class	Sideways 30% from each class

IV. RESULTS

In our proposed work, we have used single CNNs and fusion of different CNNs models and shown the result in Table II-IV for fold1-3. The scale we have taken for our experiment is from range (0-1).

A. Analysis of fusion on CNNs in differerent folds

As we can see from Table II, Resnet50 achieved accuracy of 0.67, Efficient b0 achieved accuracy of 0.72 and Convnextiny also achieved accuracy of 0.72, and when we combine two different models such as Resent 50 and Efficientnet b0, the accuracy is 0.69. The combination of Resnet50 and Convnextiny achieved 0.71 accuracy whereas Convnextiny and Efficientnet b0 marks 0.69 accuracy, but when we combine all three models, Resnet50, Efficientnet b0 and Convnextiny then we have achieved accuracy of 0.74 which is highest in all the cases.

TABLE II. RESULTS ON FOLD1

Model	Class	Precision	Recall	F1 score	Accuracy
Resnet50	Normal	0.35	0.48	0.41	0.67
	OSCC	0.82	0.72	0.77	
Efficientnet-b0	Normal	0.42	0.41	0.42	0.72
	OSCC	0.82	0.82	0.82	
Convnextiny	Normal	0.42	0.44	0.42	0.72
	OSCC	0.82	0.81	0.82	
Resnet50 + Efficientnet-b0	Normal	0.38	0.48	0.42	0.69
	OSCC	0.82	0.75	0.79	
Resnet50 + Convnextiny	Normal	0.4	0.39	0.39	0.71
	OSCC	0.81	0.81	0.81	
Convnextiny + Efficientnet-b0	Normal	0.39	0.48	0.43	0.69
	OSCC	0.83	0.76	0.79	
Resnet50 + Efficientnet-b0 + Convnextiny	Normal	0.45	0.40	0.42	0.74
	OSCC	0.82	0.85	0.83	

TABLE III. RESULTS ON FOLD2

Model	Class	Precision	Recall	F1 score	Accuracy
Resnet50	Normal	1	0.05	0.09	0.77
	OSCC	0.77	1	0.87	
Efficientnet-b0	Normal	0	0	0	0.76
	OSCC	0.76	1	0.87	
Convnextiny	Normal	0.67	0.05	0.09	0.77
	OSCC	0.77	0.99	0.87	
Resnet50 + Efficientnet-b0	Normal	1	0.01	0.02	0.77
	OSCC	0.77	1	0.87	
Resnet50 + Convnextiny	Normal	0.89	0.09	0.17	0.78
	OSCC	0.78	1	0.87	
Convnextiny + Efficientnet-b0	Normal	0.9	0.1	0.19	0.78
	OSCC	0.78	1	0.88	
Resnet50 + Efficientnet-b0 + Convnextiny	Normal	0.75	0.07	0.13	0.77
	OSCC	0.77	0.99	0.87	

TABLE IV. RESULTS ON FOLD3

Model	Class	Precision	Recall	F1 score	Accuracy
Resnet50	Normal	0.29	0.26	0.28	0.68
	OSCC	0.78	0.8	0.79	
Efficientnet-b0	Normal	0.29	0.23	0.25	0.68
	OSCC	0.77	0.82	0.8	
Convnextiny	Normal	0.34	0.18	0.24	0.72
	OSCC	0.78	0.89	0.83	
Resnet50 + Efficientnet-b0	Normal	0.28	0.22	0.25	0.68
	OSCC	0.77	0.82	0.8	
Resnet50 + Convnextiny	Normal	0.35	0.23	0.28	0.72
	OSCC	0.78	0.87	0.82	
Convnextiny + Efficientnet-b0	Normal	0.28	0.25	0.27	0.67
	OSCC	0.77	0.8	0.79	
Resnet50 + Efficientnet-b0 + Convnextiny	Normal	0.30	0.23	0.26	0.69
	OSCC	0.78	0.83	0.80	

From Table III, Resnet50 and Convnextiny achieved accuracy of 0.77 respectively whereas Efficientnet b0 achieved accuracy of 0.76. After that combination of Resnet50 and Efficientnet b0 achieved 0.77 and Resnet50, Efficientnet b0 and Convnextiny achieved the same 0.77, but when we combine Resnet50 and Convnextiny we achieved accuracy of 0.78 and also the combination of Convnextiny and Efficientnet b0 achieved the same 0.78, which is highest in all the cases.

From Table IV, we concluded that Resnet50 achieved accuracy of 0.68 and Efficientnet b0 also achieved the same 0.68 accuracy, Convnextiny on other hand achieved 0.72 accuracy, whereas when we combine Resnet50 and Efficientnet b0, we got 0.68 accuracy, the combination of Convnextiny and Efficientnet b0 yields 0.67 accuracy whereas the combination of Resnet50, Efficientnet b0 and Convnextiny gives 0.69 accuracy and also, Resnet50 and Convnextiny achieved 0.72 accuracy which is highest in all the cases.

TABLE V. FUSION MODEL ACCURACY(BASED ON TRAINING AND TESTING 10 TIMES)

Training to Testing ratio	Accuracy ± std
70:30	84.3% **±4.2**
80:20	92.12% **±3.3**

B. Comparison with previous works

TABLE VI. COMPARISON TO OTHER METHODS

Models	Accuracy
XGBoost (Researchers in [24])	95.1%
VGG16 (Researchers in [25])	90.01%
Densenet201 (Researchers in [26])	85.55%
Densenet121 (Researchers in [28])	81.89%
Our Proposed Work	**92.12%**

We have summarized the comparison of our results with previous works which is in Table VI. Our approach achieves 92.12% is better than most methods.

V. CONCLUSION

In our paper, we conducted the classification of oral histopathological images into cancerous and non-cancerous categories using three distinct CNN models. Initially, we performed individual experiments with each model, and subsequently, we explored the fusion of various model combinations, which yielded improved results to some extent compared to the former approach. But it's crucial to remember that our research was constrained. by a relatively small dataset, which may not provide highly precise results and lacks comprehensive representation of real-world cancer diagnosis cases. In the future, our objective is to enhance our results by incorporating a larger dataset that encompasses a wider range of cases. Furthermore, we remain open to the possibility of integrating additional CNN models that could further enhance accuracy and facilitate practical applications in real-world cancer diagnosis scenarios.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel et al., Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021; 71: 209-249.

[2] K. Ali, "Oral cancer-the fight must go on against all odds." Evid.-Based Dent 23 (2022): 4-5.

[3] Y. Ghantous and I. Abu Elnaaj, Global incidence, and risk factors of oral cancer. Harefuah 2017; 156: 645-649

[4] A. Roi, C.I. Roi, N.I. Andreescu, M. Rivis, Badea, I.D. Badea, N. Meszaros, L.C. Rusu and S. lurciuc, "Oral cancer histopathological subtypes in association with risk factors: A 5 year retrospective study," Rom. J. Morphol. Embryol, vol. 61, pp. 1213-1220, 2020

[5] C.F. Leite, K.D.D. Silva, M.C.R. Horta and M.C.F Aguiar de, "Can morphological features evaluated in oral cancer biopsies influence in decision making? a preliminary study," Pathol. Res. Pract, no. 153138, p. 216, 2020.

[6] "Head and neck cancer - Wikipedia," Head and neck cancer - Wikipedia, Nov. 28, 2019. https://en.wikipedia.org/wiki/Head_and_neck_cancer

[7] V. Kearney, J.W. Chan, G. Valdes, T.D. Solberg and S.S Yom, " The application of artificial intelligence in the IMRT planning process," in Oral Oncol, 2018, pp. 111-116.

[8] N. Al-Rawi, A. Sultan, B. Rajai et al., "The effectiveness of artificial intellginece in detection of oral cancer," in Int Dent J. , S0020-6539(22)00064-00068., 2022.

[9] B. Thomas, V. Kumar and S. Saini, "Texture analysis-based segmentation and classification of oral cancer lesions in color images using ANN.," in IEEE International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 2013.

[10] S. Sunny, A. Baby, B.L. James, D. Balaji, P. Gurpur, A. Skandarajah, M. D' Ambrosio, R.D. Ramanjinappa, S.P. Mohan and M.H. Rana, "A smart tele-cytology point-of-care platform for oral cancer screening," PLoS ONE, no. 14, 2019.

[11] V. Romeo, R. Cuocolo, C. Ricciardi, L. Ugga, S. Coccoza, F. Verde, A. Stanzione, V. Napolitano, D. Russo and G. Improta, "Prediction of Tumor Grade and Nodal Status in Oropharyngeal and Oral Cavity Squamous-cell Carcinoma Using a Radiomic Approach.," Anticancer. Res., no. 40, pp. 271-280, 2019

[12] C. S. Chu, N. P. Lee, J. Adeoye, P. Thomson, and S. Choi, "Machine learning and treatment outcome prediction for oral cancer," Journal of Oral Pathology & Medicine, vol. 49, no. 10, pp. 977–985, Aug. 2020, doi: 10.1111/jop.13089.

[13] R. O. Alabi et al., "Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review," Artificial Intelligence in Medicine, vol. 115, p. 102060, May 2021, doi: 10.1016/j.artmed.2021.102060

[14] 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 Optics, vol. 26, no. 8, Article ID 086007, 2021.

[15] K. Warin, W. Limprasert, S. Suebnukarn, S. Jinaporntham, and P. Jantana, "Automatic classification and detection of oral cancer in photographic images using deep learning algorithms," Journal of Oral Pathology and Medicine, vol. 50, no. 9, pp. 911–918, 2021.

[16] I. A. Ahmed, E. M. Senan and H. S. A. Shatnawi, "Analysis of Histopathological Images for Early Diagnosis of Oral Squamous Cell Carcinoma by Hybrid Systems Based on CNN Fusion Features", International Journal of Intelligent Systems, vol. 2023, Article ID 2662719, 33 pages, 2023. https://doi.org/10.1155/2023/2662719

[17] B. Ilhan, K. Lin, P. Guneri, P. Wilder-Smith, "Improving Oral Cancer Outcomes with Imaging and Artificial Intelligence," Journal of Dental Research, vol. 99, no. 3, p. 241, 2020.

[18] "Can AI help better predict mouth cancer risk? UK scientists think so | The Financial Express," Can AI help better predict mouth cancer risk? UK scientists think so | The Financial Express, Nov. 05, 2020. https://www.financialexpress.com/lifestyle/health/can-ai-help-better-predict-mouth-cancer-risk-uk-scientists-think-so/2121433/

[19] K. He, X. Zhang and S. Ren, J. Sun, "Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)," in IEEE, 2016.

[20] M. Tan, Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in International conference on machine learning. PMLR, 2019.

[21] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, K. Keutzer, "Densenet: Implementing efficient convnet descriptor pyramids.," arXiv preprint arXiv:1404.1869, 2014.

[22] Z. Liu, H. Mao, C-Y. Wu, C. Feichtenhofer, T. Darrell and S. Xie, "A ConvNet for the 2020s," vol. 1, 10 January 2022.

[23] Y. Rahman, "A histopathological image repository of normal epithelium of Oral Cavity and Oral Squamous Cell Carcinoma," Mendeley Data, 7 November 2019

[24] M. A. Deif, H. Attar, A. Amer, I. A. Elhaty, M. R. Khosravi, and A. A. A. Solyman, "Diagnosis of Oral Squamous Cell Carcinoma Using Deep Neural Networks and Binary Particle Swarm Optimization on Histopathological Images: An AIoMT Approach," Computational Intelligence and Neuroscience, vol. 2022, pp. 1–13, Sep. 2022, doi: 10.1155/2022/6364102.

[25] M. Das and R. Dash, "Data Augmentation and its Application in Histopathological Oral Cell Image Classification," 2022 International Conference on Machine Learning, Computer Systems and Security (MLCSS), Bhubaneswar, India, 2022, pp. 194-198, doi: 10.1109/MLCSS57186.2022.00043.

[26] C. Kavyashree, H. S. Vimala and J. Shreyas, "Improving Oral Cancer Detection Using Pretrained Model," 2022 IEEE 6th Conference on Information and Communication Technology (CICT), Gwalior, India, 2022, pp. 1-5, doi: 10.1109/CICT56698.2022.9997897.

[27] Z. Wang et al., "Deep-supervised adversarial learning-based classification for digital histologic images," Medical Imaging 2023: Digital and Computational Pathology, Apr. 2023, Published, doi: 10.1117/12.2654402.

[28] M. Li et al., "CIABNet: Category imbalance attention block network for the classification of multi-differentiated types of esophageal cancer," Medical Physics, vol. 50, no. 3, pp. 1507–1527, Nov. 2022, doi: 10.1002/mp.16067.

[29] D. K. Redie, S. Bilgaiyan, and S. Sagnika, "Oral cancer detection using transfer learning-based framework from histopathology images," Journal of Electronic Imaging, vol. 32, no. 05, Sep. 2023, doi: 10.1117/1.jei.32.5.053004.

[30] R. Gupta, V. Tripathi and A. Gupta, "An Efficient Model for Detection and Classification of Internal Eye Diseases using Deep Learning," 2021 International Conference on Computational Performance Evaluation (ComPE), Shillong, India, 2021, pp. 045-053

[31] M. Diwakar et al., "Low-dose COVID-19 CT Image Denoising Using CNN and its Method Noise Thresholding," Current Medical Imaging Reviews, vol. 19, no. 2. Bentham Science Publishers Ltd., pp. 182–193, Feb. 2023.

[32] P. Tiwari et al., "CNN based multiclass brain tumor detection using medical imaging," Comput. Intell. Neurosci., vol. 2022, p. 1830010, 2022.