Journal of Trends and Challenges in Artificial Intelligence

Vol. 01, No. 02 (2024) 45-50, doi: 10.61552/JAI.2024.02.001 - http://jai.aspur.rs

EMPOWERING ORAL SQUAMOUS CELL CARCINOMA DETECTION WITH DEEP LEARNING: INSIGHTS FROM CONVOLUTIONAL NEURAL NETWORK ANALYSIS OF HISTOPATHOLOGICAL IMAGES

Akshat Mishra
K. Srinivas[1]
A. Charan Kumari

Received 17.01.2024.
Revised : 27.02.2024.
Accepted 20.04.2024.

Keywords:

Convolutional Neural Networks, Deep learning, Early detection, Histopathological Images, Oral cancer, Oral Squamous Cell Carcinoma.

Original research

A B S T R A C T

The global health is heavily impacted by Oral Squamous Cell Carcinoma (OSCC), marked by high mortality rates often due to late diagnoses. This research delves into leveraging Convolutional Neural Networks for the early detection of OSCC. The model's architecture is intricately designed to discern critical features of OSCC, incorporating convolution, max-pooling, and dense layers, culminating in a sigmoid function for binary classification. This research findings reveal the model's high proficiency in OSCC identification, achieving 98.49% training accuracy, 86.89% validation accuracy, and 89.37% testing accuracy. Notably, for the Normal class, the model demonstrated a robust precision of 0.88, recall of 0.90, and the F1-score of 0.89, whereas for the OSCC class, precision is 0.90, recall is 0.88, and the F1-score is 0.89, underscoring its effectiveness in differentiating between normal and cancerous tissues. This study suggests CNNs as a viable and promising approach for OSCC early detection, potentially transforming screening and improving patient outcomes.

1. INTRODUCTION

Oral cancer is a serious global health concern caused by the malignant growth of tissues in the oral cavity, including the lips, gums, tongue, floor of the mouth, and palate. With 177,757 deaths and 377,713 new cases in 2020, it was the 13th most prevalent cancer worldwide. One health issue is the increasing patient-to-doctor ratio; in rural regions, patients with mouth cancer have a 15% five-year survival rate due to delayed discovery and treatment. However, the five-year survival rate is almost 65% in developed countries (Hertel *et al.*, 2022). Oral Squamous Cell Carcinoma (OSCC), the most aggressive kind of oral cancer, represents 90% of cases and is deadly if found too late (Ling *et al.*, 2020). Early identification of OSCC is critical for successful therapy, increased probability of survival, and low death and morbidity rates (Chakraborty *et al.*, 2019). OSCC frequencies are significantly higher in South Asian countries. Pakistan ranks first and second in terms of the frequency of cancer cases in men and women, respectively; however, India has the highest number of cases (one-third) (Anwar *et al.*, 2020).

OSCC remains a serious concern despite advancements in medical technology and therapeutic approaches. A poor prognosis and limited possibilities for treatment are present when it is identified at an advanced stage. Understanding the cause of oral cancer is crucial for

[1] Corresponding author: K. Srinivas
Email: ksri12@gmail.com

effective prevention and timely identification efforts. This disease develops due to a complicated interaction between genetic predisposition, environmental variables, and lifestyle decisions. The most significant risk factor for oral cancer, particularly OSCC, is the use of tobacco products, including both smoking and smokeless categories. Approximately one-third of all occurrences of mouth cancer worldwide are caused by tobacco use. Furthermore, drinking too much alcohol raises the risk of oral cancer substantially, especially when paired with tobacco usage.

The first step in diagnosing OSCC using traditional procedures is a microscopy-based histopathological examination of tissue samples (Deif & Hammam, 2020; Kong *et al.*, 2009). However, innovative alternatives are required because standard diagnostic tools have limits, and early diagnosis is critical. In order to detect OSCC in histopathology images, this study looks into the application of Convolutional Neural Networks (CNNs). Examining these images can assist healthcare providers in detecting irregularities that may suggest a variety of diseases, including oral cancer. However, the procedure is time-consuming and requires expertise. Furthermore, inter-observer variability might result in inconsistencies in diagnoses since different pathologists may interpret the same image differently.

The main contributions of this research work are:

1. The study focuses on the early identification of Oral Squamous Cell Carcinoma (OSCC) using Convolutional Neural Networks (CNNs) on histopathology images.
2. OSCC-CNN, a dedicated CNN model, was developed and trained on a carefully balanced and augmented dataset of histopathological images tailored explicitly for OSCC detection.
3. The model has demonstrated significant success in detecting OSCC accurately, as evidenced by its good training accuracy (98.49%), validation accuracy (86.89%), and testing accuracy (89.37%).
4. The model displayed high values of recall, precision, and F1-score metrics for all classes, indicating its effectiveness in distinguishing between the two tissue types.
5. This research suggests that using CNNs in clinical settings could significantly modify OSCC screening procedures, resulting in earlier detection and better patient outcomes.

The subsequent sections of this paper are outlined as follows: Section 2 presents a review of the relevant literature. Section 3 outlines the methodology employed, encompassing the dataset description, model description, parameter used, and the evaluation metrics. Section 4 delves into the results and discussions. The conclusions drawn from this research are summarized in Section 5, with Section 6 highlighting potential avenues for future scope.

2. LITERATURE REVIEW

This section provides a brief overview of the methods studied by various researchers in this field.

Aubreville *et al.* (2017) investigated a novel deep learning-based technique for identifying Oral Squamous Cell Carcinoma using Confocal Laser Endomicroscopy (CLE) images. Their methodology demonstrated 88.3% accuracy, exceeding prior methods and potentially leading to earlier and more accurate OSCC identification. In their research, Jubair *et al.* (2021) introduced a unique lightweight deep convolutional neural network (CNN) created specifically for the early detection of oral cancer. Their CNN model achieved an accuracy of 85%. Lin *et al.* (2021) used the HRNet model to diagnose oral cavities based on images they collected. The HRNet model demonstrated an accuracy of 84.3% and a sensitivity of 83%, outperforming the ResNet50 and DenseNet169 models.

Welikala *et al.* (2021) proposed an AI system using convolutional neural networks and trainable soft attention to detect high-risk lesions in oral cavity images. Their approach, which combined localisation loss with clinical assistance, produced an accuracy of 83.33%. Sharma *et al.* (2022) used pre-trained CNNs to analyse clinical images and identify between normal tissue, pre-cancer, and malignancy with a 76% accuracy using VGG19.

Yuan *et al.* (2022) built a unique Multi-Level Deep Residual Learning (MDRL) network for detecting malignant and benign (normal) tissues in Optical Coherence Tomography (OCT) images. The MDRL system gained an accuracy of 87.5%. Kavyashree *et al.* (2022) compared the CNN model with pretrained DenseNet models such as DenseNet121, DenseNet169, and DenseNet201 for Oral Cancer detection. Their study revealed that DenseNet201 achieved the best detection accuracy of 85% with a significant loss reduction.

Das *et al.* (2023) presented a deep learning-based approach for automatically detecting Oral Squamous Cell Carcinoma (OSCC) from histopathological images of oral mucosa. Achieving a high accuracy, their study showcases the potential of convolutional neural networks in enhancing oral cancer diagnosis.

3. METHODOLOGY

3.1 Dataset

The dataset of histopathological images comprises a total of 5192 images, with 2698 images categorized as OSCC and 2494 images as normal. It has been divided into three subsets to aid in the training and evaluation of the model. The training subset, representing 70% of the dataset, is utilized for training the model to identify features in both normal and OSCC tissues. The validation subset, representing 15% of the dataset, is employed to evaluate the model's performance during training and adjust parameters. Subsequently, the testing subset, accounting for 15% of the dataset, is utilized to assess the model's performance on unseen data post-training completion. The dataset utilized in this study is a publicly available dataset under the CC0 Public domain.

3.2 Model description

The OSCC-CNN model for detecting Oral Squamous Cell Carcinoma starts with a 32-filter convolutional layer with a three-kernel size. The activation function utilised is Rectified Linear Units (ReLUs). The model relies on images of dimensions (224, 224, 3) and employs max-pooling with a pool size of two to efficiently downsample spatial dimensions. More convolutional layers are added to the model, each with larger filters (64, 32, 32) and similar kernel configurations. A max-pooling layer follows each convolutional layer. The model's hierarchical structure enhances its capacity to detect complicated features in oral pathology images. The architecture concludes with a flatten layer that transforms the high-dimensional output. The subsequent layers are densely connected and include ReLU activation functions and filter sizes of 256, 128, and 64, respectively. A dropout layer is added after the first dense layer, with a dropout rate of 0.45, to help reduce overfitting during training. A sigmoid activation function is used for the last layer, which allows for binary classification between "Normal" and "Oral Squamous Cell Carcinoma (OSCC)". The model’s summary is shown in Figure 1.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 222, 222, 32)	896
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)	0
conv2d_1 (Conv2D)	(None, 109, 109, 64)	18496
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 64)	0
conv2d_2 (Conv2D)	(None, 52, 52, 32)	18464
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 32)	0
conv2d_3 (Conv2D)	(None, 24, 24, 32)	9248
max_pooling2d_3 (MaxPooling2D)	(None, 12, 12, 32)	0
flatten (Flatten)	(None, 4608)	0
dense (Dense)	(None, 256)	1179904
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
dense_2 (Dense)	(None, 64)	8256
dense_3 (Dense)	(None, 1)	65

Total params: 1268225 (4.84 MB)
Trainable params: 1268225 (4.84 MB)
Non-trainable params: 0 (0.00 Byte)

Figure 1. OSCC-CNN Model Summary for OSCC Detection

3.3 Parameters

This research aims to construct OSCC-CNN model that can correctly diagnose OSCC. The parameters are tuned accordingly to achieve the objective. The proposed model uses Rectified Linear Unit (ReLU) as an activation function to recognise complex patterns in data. In addition, the model incorporates a 0.45 rate dropout layer to help reduce overfitting and encourage generalization. During the training stage, this dropout mechanism randomly deactivates a section of neurons, which encourages the model to extract a wider range of robust features from the input data, thus improving its generalizability. Binary crossentropy is a loss function that is appropriate for binary classification problem, such as detecting oral cancer, hence it was used. The Adam optimizer is used for optimization because of its adjustable learning rate properties and efficient convergence, with 0.0001 as the learning rate. A batch size of 32 has been selected. Furthermore, the model architecture has 1,268,225 trainable parameters, which allow it to dynamically adjust its weights and biases to enhance prediction performance by lowering the loss function.

3.4 Evaluation metrics

To evaluate the performance of OSCC-CNN, the following evaluation metrics are used:

- **Accuracy:** The accuracy of a model is obtained by dividing the number of accurately predicted cases by the total number of cases. Its high value indicates a model's ability to make correct class predictions.

$$\text{Accuracy} = \frac{TP+TN}{TP+FN+TN+FP} \quad (1)$$

- **Precision:** Precision evaluates the accuracy of the model's positive predictions, with a focus on reducing false positives. Its high value indicates the low rate of false positives within the model, suggesting that it is accurate at predicting positive outcomes.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

- **Recall (Sensitivity):** It measures the model's capacity to accurately identify all relevant events. Its high value indicates that the majority of positive cases are correctly identified by the model while reducing false negatives.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

- **F1-Score:** It is a performance metric that gives a fair evaluation of a model's correctness. It is calculated using the harmonic mean of precision and recall, which accounts for both false positives and false negatives. The F1-Score is particularly effective for attaining a balance between precision and recall.

$$\text{F1- Score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision}+\text{recall}} \quad (4)$$

The terminology used in the evaluation metrics are defined below:

True Positives (TP) are cases in which the model correctly predicts the positive class, hence detecting examples of OSCC. For example, the model properly identifies and classifies an OSCC-positive patient.

True Negatives (TN) are those occurrences where the model accurately predicts the negative class, hence identifying occurrences when no OSCC is present. For example, the model accurately detects and classifies a patient without OSCC as negative.

False Positives (FP) are cases where the model inaccurately identifies a condition as positive, even though they do not have OSCC. For example, the model incorrectly detects and classifies a patient without OSCC as positive.

False Negatives (FN) occur when the model incorrectly predicts the negative class, such as ignoring OSCC cases and classifying them as negative when they are actually positive. For example, the model fails to identify and classify a patient with OSCC as positive.

4. RESULTS AND DISCUSSION

The effectiveness of the OSCC-CNN in detecting OSCC is evaluated based on its accuracy at various stages of training and testing. The training accuracy, indicating the model's proficiency on the training data, stands notably high at 98.49%. This showcases the model's successful learning process in accurately identifying OSCC patterns during training. Additionally, the model achieves a validation accuracy of 86.89%, testing its performance on new, untested data and facilitating parameter adjustments. Lastly, the testing accuracy, gauging the model's effectiveness on entirely new data not included in training or validation, is recorded at 89.37%. This demonstrates the model's consistent and reliable performance when applied to real-world conditions, giving confidence in its potential to detect OSCC in clinical settings accurately.

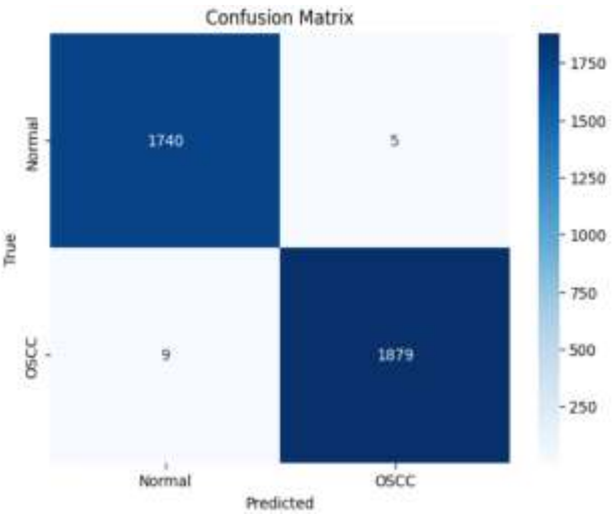

Figure 2. Confusion Matrix of Training Data

Overall, these data demonstrate the OSCC-CNN model's stability and efficacy in detecting OSCC, with excellent accuracies throughout multiple evaluation phases. These findings highlight its potential as a useful diagnostic tool for clinicians in real-world settings.

Figure 2 shows the proposed model's confusion matrix on Training data. Within the Normal class, which included 1745 cases, the model demonstrated good accuracy by correctly classifying 1740 cases with only 5 cases misclassified. Similarly, in the OSCC class, which contained 1888 cases, the model showed good results by correctly classifying 1879 cases, with only 9 cases misclassified.

Table 1. Classification Report of Training Data

Class	Precision	Recall	F1-score
Normal	0.99	1.00	1.00
OSCC	1.00	1.00	1.00

Table 1 presents the Classification Report generated from the Training data. In the Normal category, the precision is recorded as 0.99, and recall, along with the F1-score, comes out to be 1. Similarly, for the OSCC category, precision is recorded as 1, as is recall and the F1-score, demonstrating the model's precise classification within this category.

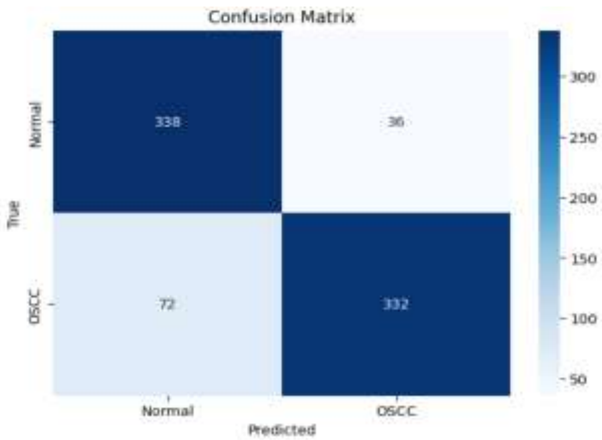

Figure 3. Confusion Matrix of Validation Data

Figure 3 shows the proposed model's confusion matrix on Validation data. Among the 374 cases in the Normal class, 338 were correctly classified, while 36 were misclassified. Similarly, in the OSCC class, which included 404 cases, the model properly classified 332 while 72 were misclassified.

Table 2. Classification Report of Validation data

Class	Precision	Recall	F1-score
Normal	0.82	0.90	0.86
OSCC	0.90	0.82	0.86

Table 2 presents the Classification Report generated from the Validation data. In the Normal category, precision is recorded as 0.82, recall as 0.90, and subsequently, the F1-score as 0.86, highlighting the model's overall performance. Similarly, the precision for the OSCC category is 0.90, followed by the recall of 0.82 and the F1-score of 0.86, highlighting the model's accuracy and effectiveness in classifying OSCC cases.

Figure 4. Confusion Matrix of Testing Data

Figure 4 shows the proposed model's confusion matrix on Testing data. Among the 375 cases in the Normal class, 337 were correctly classified, while 38 were misclassified. Similarly, in the OSCC class, which included 406 cases, the model properly classified 359 while 47 were misclassified.

Table 3. Classification Report of Testing Data

Class	Precision	Recall	F1-score
Normal	0.88	0.90	0.89
OSCC	0.90	0.88	0.89

Table 3 presents the Classification Report generated from the Testing data.

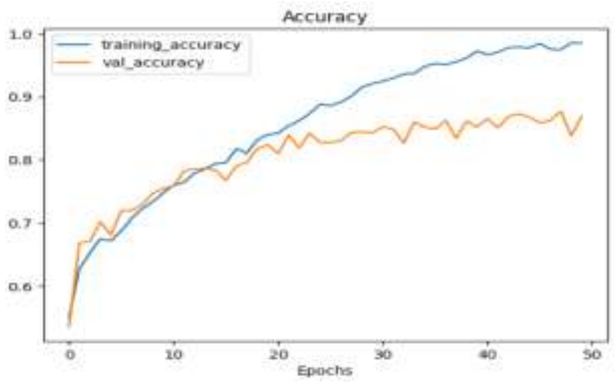

Figure 5. training accuracy vs validation accuracy

In the Normal category, precision is recorded as 0.88, recall as 0.90, and subsequently, the F1-score as 0.89, highlighting the model's overall performance. Similarly, the precision for the OSCC category is 0.90, followed by the recall of 0.88 and the F1-score of 0.89, confirming the model's correctness and usefulness in classifying OSCC cases.

Figure 5 depicts the graphical representation of Training accuracy versus Validation accuracy, providing a visual demonstration of the performance of OSCC-CNN. The graph demonstrates that over the course of 50 epochs, the suggested model attained a notable training accuracy of 98.49% and a corresponding validation accuracy of 86.89%.

5. CONCLUSION

It is evident from the research findings that the application of Convolutional Neural Networks (CNNs) heralds a significant leap forward in the early identification of Oral Squamous Cell Carcinoma. The diligent crafting and training of OSCC-CNN, a CNN model tuned explicitly to parse through a comprehensive dataset of histopathological images of OSCC patients, highlights the model's notable accuracy and adeptness at distinguishing between normal and cancerous tissues. The outcomes of this investigation illustrate a promising path toward overhauling current OSCC screening practices, potentially ushering in an era of earlier detection and markedly better patient survival rates. The model's good performance on accuracy metrics—98.49% during training, 86.89% in validation phases, and 89.37% in testing scenarios—reinforces the utility of CNNs in pinpointing OSCC with notable precision. Moreover, the model's good performance on various standard performance metrics, including recall, precision and F1-score particularly concerning OSCC, attests to its reliability and diagnostic ability. These achievements not only validate the model's performance but additionally hint at CNNs' vast potential in the realm of medical diagnostics. Incorporating CNNs into everyday clinical practices appears to hold immense promise for refining OSCC screening processes, creating more efficient screening programs, and enhancing the accuracy of diagnoses.

6. FUTURE WORK

This study lays a solid foundation for subsequent inquiries, suggesting that expanding the model's validation to encompass larger datasets and integrating clinical data could further sharpen diagnostic accuracy. Ultimately, this investigation accentuates the revolutionary impact of deep learning on medical diagnostics. By tapping into CNNs' capabilities, we stand on the cusp of pioneering advancements in the early detection and treatment approaches for OSCC and potentially other malignancies, marking a pivotal stride towards elevating patient care and improving outcomes in the field of oncology.

References:

Barakina E.Y., Popova A.V., Gorokhova S.S., Voskovskaya A.S. (2020). "Digital technologies and artificial intelligence technologies in education", *European Journal of Contemporary Education*, 10(2), 285-296.

Anwar, N., Pervez, S., Chundriger, Q., Awan, S., Moatter, T., & Ali, T. S. (2020). Oral cancer: Clinicopathological features and associated risk factors in a high risk population presenting to a major tertiary care center in Pakistan. *PLOS ONE*, *15*(8), e0236359. DOI: 10.1371/journal.pone.0236359

Aubreville, M., Knipfer, C., Oetter, N., Jaremenko, C., Rodner, E., Denzler, J., Bohr, C., Neumann, H., Stelzle, F., & Maier, A. (2017). Automatic classification of cancerous tissue in laserendomicroscopy images of the oral cavity using deep learning. *Scientific Reports*, *7*(1), 11979. DOI: 10.1038/s41598-017-12320-8

Chakraborty, D., Natarajan, C., & Mukherjee, A. (2019). Advances in oral cancer detection. In *Advances in Clinical Chemistry* (Vol. 91, pp. 181–200). Elsevier. DOI: /10.1016/bs.acc.2019.03.006

Das, M., Dash, R., & Mishra, S. K. (2023). Automatic detection of oral squamous cell carcinoma from histopathological images of oral mucosa using deep convolutional neural network. *International Journal of Environmental Research and Public Health*, *20*(3), 2131. DOI: 10.3390/ijerph20032131

Deif, M. A., & Hammam, R. E. (2020). Skin lesions classification based on deep learning approach. *Journal of Clinical Engineering*, *45*(3), 155–161. DOI: 10.1097/JCE.0000000000000405

Hertel, M., Hagedorn, L., Schmidt-Westhausen, A. M., Dommisch, H., Heiland, M., Preissner, R., & Preissner, S. (2022). Comparison of five-year survival rates among patients with oral squamous cell carcinoma with and without association with syphilis: A retrospective case-control study. *BMC Cancer*, *22*(1), 454. DOI: 10.1186/s12885-022-09583-4

Jubair, F., Al-karadsheh, O., Malamos, D., Al Mahdi, S., Saad, Y., & Hassona, Y. (2021). A novel lightweight deep convolutional neural network for early detection of oral cancer. *Oral Diseases*, *28*(4), 1123–1130. DOI: 10.1111/odi.13825

Kavyashree, C., Vimala, H. S., & Shreyas, J. (2022). Improving oral cancer detection using pretrained model. *2022 IEEE 6th Conference on Information and Communication Technology (CICT)*, 1–5. DOI: 10.1109/CICT56698.2022.9997897

Kong, J., Sertel, O., Shimada, H., Boyer, K. L., Saltz, J. H., & Gurcan, M. N. (2009). Computer-aided evaluation of neuroblastoma on whole-slide histology images: Classifying grade of neuroblastic differentiation. *Pattern Recognition*, *42*(6), 1080–1092. DOI: 10.1016/j.patcog.2008.10.035

Lin, H., Chen, H., Weng, L., Shao, J., & Lin, J. (2021). Automatic detection of oral cancer in smartphone-based images using deep learning for early diagnosis. *Journal of Biomedical Optics*, *26*(08). DOI: 10.1117/1.JBO.26.8.086007

Ling, Z., Cheng, B., & Tao, X. (2020). Epithelial-to-mesenchymal transition in oral squamous cell carcinoma: Challenges and opportunities. *International Journal of Cancer*, *148*(7), 1548–1561. DOI: 10.1002/ijc.33352

Sharma, D., Kudva, V., Patil, V., Kudva, A., & Bhat, R. S. (2022). A convolutional neural network based deep learning algorithm for identification of oral precancerous and cancerous lesion and differentiation from normal mucosa: A retrospective study. *Engineered Science*, *Volume 18 (June 2022)*(0), 278–287.

Welikala, R. A., Remagnino, P., Lim, J. H., Chan, C. S., Rajendran, S., Kallarakkal, T. G., Zain, R. B., Jayasinghe, R. D., Rimal, J., Kerr, A. R., Amtha, R., Patil, K., Tilakaratne, W. M., Cheong, S. C., & Barman, S. A. (2021). Clinically guided trainable soft attention for early detection of oral cancer. In N. Tsapatsoulis, A. Panayides, T. Theocharides, A. Lanitis, C. Pattichis, & M. Vento (Eds.), *Computer Analysis of Images and Patterns* (Vol. 13052, pp. 226–236). Springer International Publishing. DOI: 10.1007/978-3-030-89128-2_22

Yuan, W., Yang, J., Yin, B., Fan, X., Yang, J., Sun, H., Liu, Y., Su, M., Li, S., & Huang, X. (2022). Noninvasive diagnosis of oral squamous cell carcinoma by multi-level deep residual learning on optical coherence tomography images. *Oral Diseases*, *29*(8), 3223–3231. DOI: 10.1111/odi.14318

Akshat Mishra
Dayalbagh Educational Institute,
Dayalbagh, Agra, India
akshatsvmishra@gmail.com
ORCID: 0009-0001-3079-4740

K. Srinivas
Dayalbagh Educational Institute,
Dayalbagh, Agra, India
ksri12@gmail.com
ORCID: 0009-0002-3884-6282

A. Charan Kumari
Dayalbagh Educational Institute,
Dayalbagh, Agra, India
charankumari@dei.ac.in
ORCID: 0000-0002-3160-1912