Identifying Oral Carcinoma from Histopathological Image using Unsupervised Nuclear Segmentation

Rahul Shukla
BTech ECE
IIIT Naya Raipur
Raipur, India
rahuls20101@iiitnr.edu.in

Bhavesh Ajwani
BTech Data Scienece and AI
IIIT Naya Raipur
Raipur, India
bhavesh20102@iiitnr.edu.in

Shubham Sharma
BTech ECE
IIIT Naya Raipur
Raipur, India
shubhams20101@iiitnr.edu.in

Debanjan Das
Senior Member IEEE
IIIT Naya Raipur
Raipur, India
debanjan@iiitnr.edu.in

***Abstract*—Oral Cancer, a worldwide health concern, highlights the urgent need for accurate and swift detection and cure. Current diagnosing strategies primarily involve pathologists for analyzing tissue biopsy samples, a method that is time-consuming and heavily driven by pathologists' experience. To address these drawbacks, this study proposes a novel technique that incorporates machine vision for cancer detection, aiming to enhance diagnostic accuracy. Given the intricate nature of histopathological images, we adopt an unsupervised approach for cancer detection, in contrast to traditional deep learning or supervised approaches. The nucleus in a cancerous tissue biopsy image is identified as the region of interest (ROI), due to its key characteristics and form. We extract the ROI using K-means clustering augmented with a thresholding technique and apply a novel classification method for the final stage of cancer detection. Our proposed model achieved an accuracy of approximately 97.28%, with a closely following validation accuracy of roughly 96.34% making it more efficient and reliable at cancer detection. These results underscore the effectiveness of our two-stage process starting with image segmentation followed by CNN-based binary classification for accurately detecting cancer cells. They reveal improved speed and precision in identifying cancerous tissues, thus offering a promising pathway for enhancing the efficacy and efficiency of oral cancer detection.**

***Index Terms*—Oral Cancer, Histopathological, Segmentation, CNN**

I. Introduction

According to World Health Organization, every year, there are nearly six lakhs new cases of oral cancer and over three lahks reported deaths [1]. There are many different types of cancer, but oral squamous cell carcinomas, commonly known as OSCCs (Oral Squamous Cell Carcinomas), account for 90 percent of all cancer cases in the medical community [2]. Early detection and medical consultation are crucial for timely diagnosis and treatment of oral cancer. The traditional process of Cancer detection involves the classification of the biopsy image of the infected tissue by an experienced pathologist but not without its limitations, accurate diagnosis is highly dependent on experience and intelligence of pathologist [4] and its time-consuming and tiresome job.

Thus in this era of machine learning where it is playing a major role in the medical field by automating the work of the various medical-related diagnoses. The same can be incorporated into this field of oral cancer detection. This acts as a motivation for building an end-to-end solution for classifying biopsy images as Malign or Benign with the use of machine learning.

This paper presents a novel methodology for detection of the cancer using biopsy images with the help of machine learning approach. The majority of the behavior of the characteristics of a cancerous biopsy cell can be identified by observing the features of the nuclei including shape, size, and eccentricity [5], the approach for solution involves ROI extraction followed by classification. Machine learning approach is being incorporated despite of the presence of more complex deep learning algorithms as they are the supervised approaches and it is not feasible to use the supervised learning approach as data set with proper annotations is not available, following the complexity of the data set. Therefore, a unique, unsupervised methodology that integrates k-means clustering with thresholding is utilized for segmenting the region of interest - nuclei. Subsequently, a bespoke CNN-based model is applied for classification purposes.

The major contributions of this research are:

- Leveraging a sizable histopathological image dataset comprising a varied set of normal and cancerous oral cavity samples, offering a robust foundation for precise oral cancer detection.
- Utilizing an unsupervised method for nuclei segmentation through K-means clustering and thresholding technique, enabling effective extraction of the Region of Interest (ROI) from intricate histopathological images.
- Developing a compact yet potent Convolutional Neural Network (CNN) classifier, CanNet, for distinguishing between benign or malignant segmented nuclei. CanNet ensures computational efficiency while sustaining high detection accuracy, marking a considerable advancement in oral cancer detection techniques.

II. Literature Survey and Background Work

Traditional methods to assess cellular changes in carcinogenesis typically rely on laboratory-based techniques. However, these approaches are often less accessible and affordable. Techniques like Multiphoton Autofluorescence Microscopy (MPAM) [7] have been utilized as an alternative, showing promise in noninvasive imaging of oral epithelial carcinogenesis. The Renishaw Raman

TABLE I
EXISTING WORKS ON DETECTING CANCER

Related Works	Targeted area	Methodology/Approach	Technique Used	classification model	Accuracy
Yang Liu et. al. [9]	standard	Determining nuclear morphology	To calculate radii, nucleus enclosed area, etc. as distinguishing factor	No	81.82%
S. -Y. Huang et. al. [12]	OSCC	Object Detection with Wavelet Filters	FCN model with Gabor Filters	FCN Model	95%
Pennisi A et. al., [11]	OSCC nuclei	Semantic Segmentation	Multi- Encoder U-Net	No	96%
V. Magoulianitis et. al., [14]	Cell nuclei	Unsupervised Segmentation	Proposed CBM technique	No	95.34%
Our Work	Cell Nuclie	Unsupervised segmentation with deep learning based classification	K-means clustering with CBM, classification	Yes	97.2%

microspectrometer, another non-laboratory method, has demonstrated potential in generating highly sensitive spectra from tissue samples, aiding in the understanding of Oral Squamous Cell Carcinoma (OSCC) [8]. Yet, challenges persist in effectively implementing these imaging technologies. Parallelly, research by Yang Liu, Shikhar Uttam, P. Dey, and others described in [9] have emphasized the importance of nuclear morphology in differentiating between malignant and benign tissues, thus playing a crucial role in detecting numerous cancers, like breast cancer, lung cancer and OSCC.

Previous literature demonstrates deep learning methodologies such as ResNet-101 for image classification and Faster R-CNN for object detection, both showing substantial potential in medical imaging [10]. However, like many other models for disease prediction, they demand high computing power and extensive datasets. The multi-encoder U-Net,A a revised version of the renowned U-Net architecture provides promising outcomes in segmenting oral squamous cell carcinoma through the use of whole slide images. [11]. Yet, these approaches often depend on cloud services and are not designed for edge computing or mobile platforms. Furthermore, the Fully Convolutional Network (FCN) provides another avenue for oral cancer detection, incorporating RGB channels and Gabor and wavelet filters to segment cancer and pre-cancer regions [12]. Nevertheless, it is imperative to balance high accuracy with computational demands to facilitate deployment across various platforms, including mobile devices.

In line with efficient computational methods and unsupervised algorithms, the literature encompasses key studies. A study focused on the utilization of an optimized band selection and Minimum Spanning Forest (MSF) algorithm for hyperspectral imaging of tissues [13]. Another research introduced an innovative unsupervised method, referred to as the CBM approach, for the segmentation of nuclei from histopathological images [14]. Additionally, a different study highlighted the pivotal role of decision tree classifiers in differentiating between malignant and normal cells in OSCC diagnosis [21].

These works contribute to the advancement of leveraging efficient computing power and sophisticated unsupervised techniques in disease diagnosis and classification.

In conclusion, current disease detection techniques either have high accuracy with substantial computational cost or save resources but capture fewer optimal histopathological cell features. To address this dichotomy, we propose an innovative approach: unsupervised segmentation followed by classification. This may offer a more balanced, efficient, and precise solution.

III. MATERIALS AND METHOD

This study outlines a methodology devised to distinguish benign and malignant cells in oral histopathological images using a streamlined process of segmentation and classification. As depicted in Fig 1, the process begins with the multi-staged preprocessing of histopathological images, followed by feeding them into a K-means clustering algorithm and further thresholding to extract the Region of Interest (ROI), which are cell nuclei, thereby creating a prediction mask. These preprocessed images are then processed through a miniature yet efficient network designed for binary classification, thereby generating a classification mask. The final step involves combining the two generated masks, one representing the ROI and the other for binary classification, resulting in the ultimate prediction of oral cancer presence.

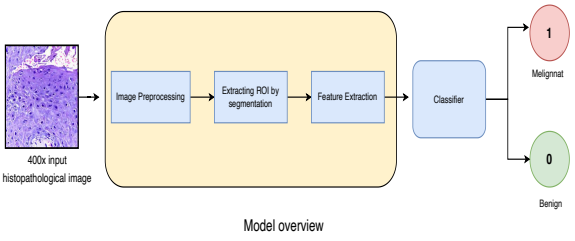

Fig. 1. Proposed methodology

A. Dataset Description

In this study, we employed a dataset [6] comprising 1224 histopathological images, which are segmented into two sets based on the magnification levels - 100× and 400×, as shown in Fig 2.

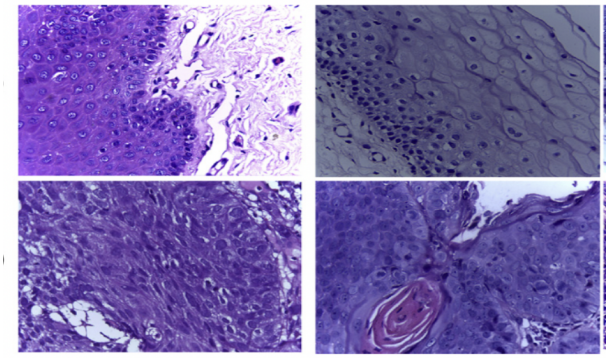

Fig. 2. Images from the dataset

The first set includes 89 images depicting regular oral tissue and 439 images displaying a particular form of oral cancer, all acquired at 100× magnification. Meanwhile, the second set consists of 201 images showing regular oral tissue and 495 images demonstrating the same oral cancer type, each taken at 400× magnification.

B. Nuclei Segmentation

For the purpose of extracting Region of interest (ROI) i.e nuclei from the dataset images the biopsy images with 400× magnification are being used. As the size of each image is large the specific section of the complete image is cropped and then further processed.
The images are initially preprocessed so to even further increase the efficiency, followed by segmentation and classification.

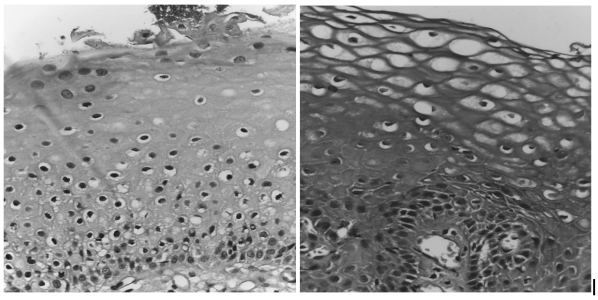

Fig. 3. Preprocessed Images Left(Benign) and Right(Malignant)

1) Image preprocessing: Biopsy slides are prepared for microscopic analysis through several steps, starting with staining of nearly transparent tissue samples to capture intricate cell structures. Stained slides are imaged at 400× magnification in RGB, transformed into HSV color space to counteract staining effects, and enhanced via Contrast Limited Adaptive Histogram Equalization (CLAHE). This process boosts the segmentation model's accuracy and overall model effectiveness. Images are eventually converted into grayscale, focusing on the *'V'* component of HSV color space, resulting in the preprocessed image seen in Figure 3.

2) Segmentation Methodology : The preprocessed images undergo segmentation through the application of the k-means clustering algorithm, setting k to 3. This algorithm operates by segregating the images into distinct clusters. In this context, K represents the predetermined quantity of clusters intended to be generated during the procedure. As such, if K equals 2, the outcome will include two clusters, if K equals 3, the result will consist of three clusters, and the pattern continues accordingly.

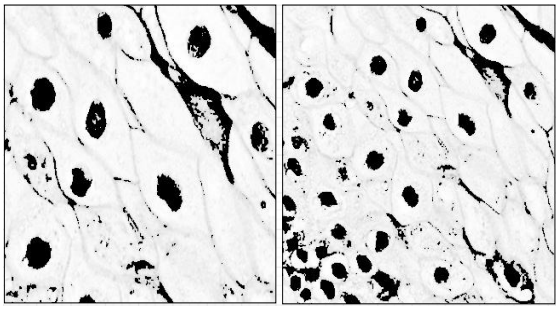

Fig. 4. K-means Segmentation Left(Benign) and Right(Malignant)

Algorithm 1 Enhanced Image Segmentation and Smoothing Algorithm

Input: Preprocessed image dataset, D
Inference: Segmented and smoothed image dataset

Initialize clusters, k, to 3.
Apply K-means to D with k clusters, yielding Img_{multi}.
for each data point $x_i \in D$ **do**
    Assign x_i to nearest centroid, creating Img_{bin}.
end for
Compute histogram H_{bin} of Img_{bin}.
Identify local minima θ_{min} between the first two peaks in H_{bin}.
for each pixel p_{ij} in Img_{bin} **do**
    if $I(p_{ij}) < \theta_{\text{min}}$ **then**
        Label p_{ij} as nucleus.
    else
        Label p_{ij} as background.
    end if
end for
Initialize $resultant_{\text{img}}$, $256 \times 256 \times 3$.
Apply smoothing to $resultant_{\text{img}}$.
return Segmented, smoothed $resultant_{\text{img}}$.

The approach is capable of splitting the data into multiple clusters, offering a pragmatic means for automatically identifying these groups within a dataset that lacks labels. Each cluster is associated with a centroid, given the centroid-based nature of the technique. The primary objective of this algorithm is to reduce or minimize the overall distances from each data point to its associated cluster. Starting with an unlabeled dataset, dividing it into *'k'* clusters, and continues until all clusters are utilized.

The algorithm requires the value of *'k'* to be defined in advance. The k-means clustering methodology primarily performs two key operations:

- Employs an iterative method to determine the ideal value for *K* central points or centroids.
- Allocates each data point to its nearest *k*-center. Data points in proximity to a specific *k*-center form a cluster.

Histogram was plotted for the above-observed results after employing k-means clustering. Hills and valleys are observed and the minima between the first two hills is considered to be threshold. If the pixel value is less than the threshold then pixel is designated as nuclei otherwise it is designated as the background. After the complete iteration of the received image, the segmented image is smoothed using the technique of contour smoothing.

C. Nuclei Classification

After extracting the ROI nuclei from the images, our task was to classify them as either benign or malignant. Given the complexity of histopathological images, a robust and computationally intensive model would typically be necessary. However, since we had already segmented the ROI section of the image, we mitigated this requirement. Once the segmentation output was classified, we successfully categorized the OSCC cells as either malignant or benign. Needing an efficient network, we implemented a deep convolutional neural network specifically designed for our task. Our network is structured with six convolutional layers followed by two fully connected layers. The first and second convolutional layers each employ 64 kernels of size 7×7, with each succeeded by a max-pooling layer. The third and fourth layers use 128 kernels of size 5×5 without max-pooling, while the fifth and sixth layers utilize 256 kernels of size 3×3, also without max-pooling.

Subsequent to the convolutional layers, there are two fully connected layers. The first contains 2048 neurons, and the final layer corresponds to the number of classes for classification - in this case, 2 classes: malignant or benign. A dropout layer with a rate of 0.25 is applied after the first fully connected layer to control overfitting.

In our model, the sigmoid activation function was applied, defined as follows:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

For the loss function, we used Hinge Loss. This loss function seeks to maximize the margin between the two classes, which leads to improved generalization performance. It is defined as:

$$L(y, \hat{y}) = \max(0, 1 - y \cdot \hat{y}), \tag{2}$$

By pushing the decision boundary further away from the data points, Hinge Loss significantly improved our model's performance and helped in preventing overfitting.

We used the Adam optimizer for optimization. The update rule for Adam optimizer is given as:

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\frac{\beta_2 v_t + (1-\beta_2)(\nabla L_t)^2}{1-\beta_2^{t+1}}} + \epsilon} \frac{\beta_1 m_t + (1-\beta_1)\nabla L_t}{1-\beta_1^{t+1}}, \tag{3}$$

In these equations, θ represents the parameters of the model, η is the learning rate, ∇L denotes the gradient of the loss function with respect to the model parameters. *m* and *v* are the first and second-moment estimates, $\beta 1$ and $\beta 2$ are the exponential decay rates for these moment estimates. *t* stands for the iteration count, and ϵ is a small constant to ensure numerical stability.

IV. Results and Discussion

A. Experimental Setup

Our experiments demonstrate that the two-stage approach, consisting of image segmentation followed by a CNN-based binary classifier, is effective in detecting malignant and benign cancer cells. The introduced segmentation algorithm and CNN classifier were both built on Kaggle cloud Nvidia GPU P100.

B. Unsupervised Segmentation Results

The efficacy of our unsupervised segmentation technique is demonstrated through a comparative analysis of the model's accuracy before and after segmentation. The model exhibited an approximate accuracy of 93.2% prior to segmentation, which dramatically increased to around 97.2% after applying our method. Figure 5 visually traces the process, starting with the initial preprocessed image. After undergoing thresholding, morphological operations, and contour smoothening, the image highlights the nuclei, aiding in accurate feature extraction. The method not only proves instrumental in oral cancer diagnosis but also has the potential for wider use in histopathological image analysis.

C. Classification Results

Our model's performance was assessed using accuracy metrics, showcasing its competency against other cutting-edge classifiers. Following a 50-epoch training period, the model attained an accuracy of about 97.28% and a validation accuracy close behind at approximately 96.34%, as illustrated in Figure 6. The almost convergent graph of loss and validation loss in Figure 7 after 50 epochs suggests a well-tuned model. These outcomes corroborate the effectiveness of the two-stage approach—image segmentation followed by CNN-based binary classification—in detecting cancer cells. The algorithm and CNN classifier, both developed on Kaggle cloud Nvidia GPU P100, underscore the effectiveness of this approach.

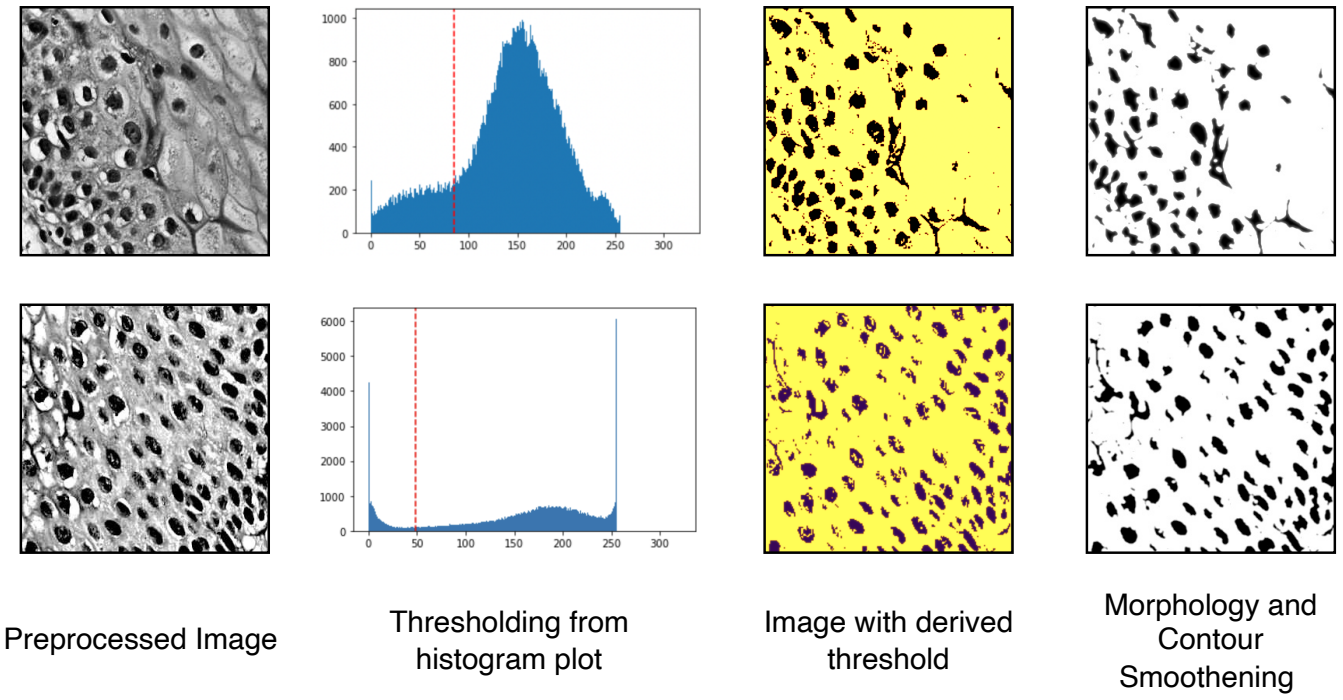

Fig. 5. Results of end-to-end Unsupervised Segmentation process

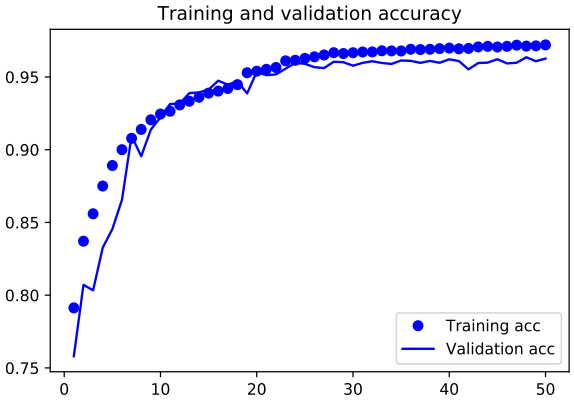

Fig. 6. Performance Trends: Training and Validation

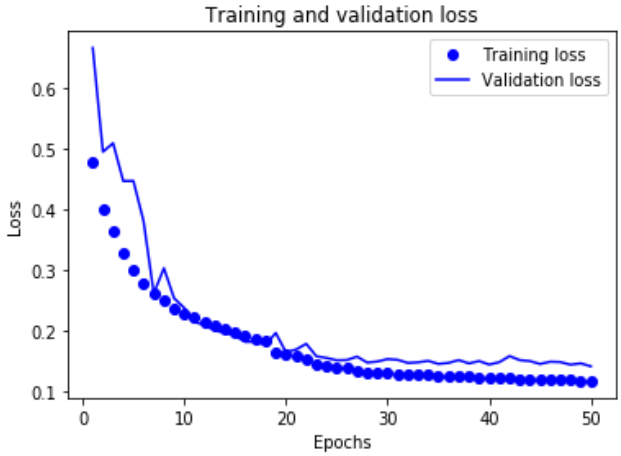

Fig. 7. Loss Metrics Progression Chart

D. Comparison with Prior Arts

In comparison with prior research, our proposed deep learning-based methodology for medical imaging distinguishes itself in its efficiency and accessibility. Traditional deep learning methodologies such as ResNet-101 for image classification and Faster R-CNN for object detection have been demonstrated to hold substantial potential in medical imaging [10]. However, these models frequently require substantial computing power and extensive datasets, potentially limiting their practical applications. Additionally, while alterations of existing architectures like the multi-encoder U-Net have shown promise, particularly in oral squamous cell carcinoma segmentation [11], these methods typically heavily rely on cloud services, making them unsuitable for edge computing or mobile platforms.

Another method, the Fully Convolutional Network (FCN), has been documented in the literature, which uses RGB channels along with Gabor and wavelet filters to detect cancerous and pre-cancerous regions [12]. However, balancing the high accuracy of these methods with their computational demands remains a persistent challenge to ensure their widespread deployment, including on mobile devices. Finally, although certain studies such as those on hyperspectral imaging of tissues using an optimal band

selection and the Minimum Spanning Forest (MSF) algorithm [13], the CBM approach for nuclei-segmentation from histopathological images [14], and the use of decision tree classifiers for distinguishing malignant cells in OSCC diagnosis [15], offer efficient and unsupervised techniques. This paper uniquely employs the Region of Interest (ROI) nuclei, a critical feature selection, which is often overlooked by other studies in the field of oral cancer detection. There is a consistent need for the development of more efficient, computationally light models for disease diagnosis and classification.

V. Conclusion and Future Works

In conclusion, the methodology employed in this research effectively utilizes a unique, unsupervised approach for oral cancer detection in histopathological images. Beginning with a meticulously curated dataset featuring images of Oral Squamous Cell Carcinoma (OSCC) and normal oral epithelium, proceeding with the implementation of K-means clustering algorithm to segment intricate histopathological images, the K-means clustering algorithm is implemented. This step allows for the identification and extraction of the region of interest (ROI), particularly the nuclei of the cells. Subsequently, the segmented images are classified by using an compact yet efficient convolutional neural network for classification, balancing the dichotomy between high accuracy and computational efficiency, traditionally seen in disease detection. Looking forward, this model holds potential for real-time medical imaging, offering a comprehensive solution for professionals. Future work involves expanding the model's capabilities through multi-class classification, exploring transfer learning, and optimizing for real-time performance, collaboration with clinical experts and extension to detect various cancers. These initiatives collectively underscore our commitment to advancing the model's versatility and effectiveness in the realm of medical diagnostics.

References

[1] B R, Nanditha Annegowda, Geetha. (2022). Oral Cancer Detection using Machine Learning and Deep Learning Techniques. International Journal of Current Research and Review. 14. 64-70. 10.31782/IJCRR.2021.14104.

[2] Montero PH, Patel SG. Cancer of the oral cavity. Surg Oncol Clin N Am. 2015 Jul;24(3):491-508. doi: 10.1016/j.soc.2015.03.006. Epub 2015 Apr 15. PMID: 25979396; PMCID: PMC5018209.

[3] Chang HY, Jung CK, Woo JI, Lee S, Cho J, Kim SW, Kwak TY. Artificial Intelligence in Pathology. J Pathol Transl Med. 2019 Jan;53(1):1-12. doi: 10.4132/jptm.2018.12.16. Epub 2018 Dec 28. PMID: 30599506; PMCID: PMC6344799.

[4] Wu, King Laronde, Denise. (2014). Oral cancer and biopsy protocol: A primer for the dental hygienist. Canadian Journal of Dental Hygiene. 48. 34-39.

[5] Kumar R, Srivastava R, Srivastava S. Detection and Classification of Cancer from Microscopic Biopsy Images Using Clinically Significant and Biologically Interpretable Features. J Med Eng. 2015;2015:457906. doi: 10.1155/2015/457906. Epub 2015 Aug 23. PMID: 27006938; PMCID: PMC4782618.

[6] Rahman, Tabassum Yesmin, et al. 'Histopathological Imaging Database for Oral Cancer Analysis'. Data in Brief, vol. 29, 2020, p. 105114, https://doi.org10.1016/j.dib.2020.105114.

[7] G. Vargas, T. Shilagard, K. -H. Ho and S. McCammon, "Multiphoton autofluorescence microscopy and second harmonic generation microscopy of oral epithelial neoplasms," 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 2009, pp. 6311-6313, doi: 10.1109/IEMBS.2009.5332783.

[8] Y. Hu, T. Jiang and Z. Zhao, "Discrimination of Squamous Cell Carcinoma of the Oral Cavity Using Raman Spectroscopy and Chemometric Analysis," 2008 First International Conference on Intelligent Networks and Intelligent Systems, Wuhan, China, 2008, pp. 633-636, doi: 10.1109/ICINIS.2008.61.

[9] Liu Y, Uttam S, Alexandrov S, Bista RK. Investigation of nanoscale structural alterations of cell nucleus as an early sign of cancer. BMC Biophys. 2014 Feb 10;7(1):1. doi: 10.1186/2046-1682-7-1. PMID: 24507508; PMCID: PMC3928095.

[10] R. A. Welikala et al., "Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer," in IEEE Access, vol. 8, pp. 132677-132693, 2020, doi: 10.1109/ACCESS.2020.3010180.

[11] A. Pennisi, D. D. Bloisi, D. Nardi, S. Varricchio and F. Merolla, "Multi-encoder U-Net for Oral Squamous Cell Carcinoma Image Segmentation," 2022 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Messina, Italy, 2022, pp. 1-6, doi: 10.1109/MeMeA54994.2022.9856482.

[12] S. -Y. Huang, C. -Y. Chiou, Y. -S. Tan, C. -Y. Chen and P. -C. Chung, "Deep Oral Cancer Lesion Segmentation with Heterogeneous Features," 2022 IEEE International Conference on Recent Advances in Systems Science and Engineering (RASSE), Tainan, Taiwan, 2022, pp. 1-8, doi: 10.1109/RASSE54974.2022.9989871.

[13] A. Gopi and C. S. Reshmi, "A noninvasive cancer detection using hyperspectral images," 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), Chennai, India, 2017, pp. 2051-2055, doi: 10.1109/WiSPNET.2017.8300122.

[14] Magoulianitis, Vasileios, et al. 'Unsupervised Data-Driven Nuclei Segmentation For Histology Images'. arXiv [Eess.IV], 2021, http://arxiv.org/abs/2110.07147. arXiv.

[15] Rahman TY, Mahanta LB, Choudhury H, Das AK, Sarma JD. Study of morphological and textural features for classification of oral squamous cell carcinoma by traditional machine learning techniques. Cancer Rep (Hoboken). 2020 Dec;3(6):e1293. doi: 10.1002/cnr2.1293. Epub 2020 Oct 7. PMID: 33026718; PMCID: PMC7941561.