Prabha Materials Science Letters
Vol. 3, No. 1, 94-104, 2024
https://doi.org/10.33889/PMSL.2024.3.1.007

Ram Arti Publishers

Evaluating the Deep Learning Models Performance for Segmentation of Oral Epithelial Dysplasia: A Histological Data-Driven Approach

Taibur Rahman
Mathematical and Computational Sciences Division,
Institute of Advanced Study in Science and Technology, Guwahati, 781035, Assam, India.
&
Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, Uttar Pradesh, India.
Corresponding author: taiburat8@gmail.com

Lipi B. Mahanta
Mathematical and Computational Sciences Division,
Institute of Advanced Study in Science and Technology, Guwahati, 781035, Assam, India.
&
Academy of Scientific and Innovative Research (AcSIR), Ghaziabad, 201002, Uttar Pradesh, India.
E-mail: lbmahanta@iasst.gov.in

(Received on February 7, 2024; Revised on March 11, 2024, Accepted on March 13, 2024)

Abstract
Oral epithelial dysplasia (OED) poses a significant precancerous risk, potentially progressing to oral squamous cell carcinoma (OSCC). Precise segmentation of OED within histopathological images is pivotal for early diagnosis and treatment planning. This study evaluates Deep Learning (DL) models for precise Oral Epithelial Dysplasia (OED) segmentation in biopsy slide images. The Vanilla UNET model is explored with the standard UNET and other transfer learning models (VGG16, VGG19, MobileNet, and DeepLabV3+) as the backbone of the model. For our application, U-Net demonstrated superior performance (IoU: 93.73%, precision: 97.96%, recall: 97.78%, F1-score: 96.76%). Visual examples highlight model strengths and limitations, providing insights beyond traditional metrics. This research advances computer-aided histopathological analysis, emphasizing DL models' crucial role in enhancing diagnostic accuracy and patient care.

Keywords- Oral epithelial dysplasia, Histopathological images, Deep learning, Segmentation, Automated diagnosis.

1. Introduction

Oral cavity health is not just about a bright smile; it's the gateway to overall well-being. Oral cavity health is crucial for overall well-being, yet despite efforts to combat bad habits, global cancer rates continue to rise. Oral cancers pose a significant threat, affecting approximately 3.5 billion people worldwide over their lifetime (The global status report on oral health 2022). In 2018, there were 355,000 global cases of oral cancer, resulting in 177,000 deaths, a number that increased to 377,713 new cases and 177,757 deaths by 2020. By 2023, the United States is projected to see nearly 1.96 million new cancer cases, leading to about 609,820 deaths (Siegel et al., 2023). In India, oral cancer is a significant public health concern, with approximately 20 cases per 100,000 people and accounting for 30% of all cancer types (Borse et al., 2020; Varshitha, n.d.). The Northeast region of India particularly faces a high burden of oral cancer, with an estimated increase in new cases to 57,131 by 2025 (North Eastern States of India Have a High Burden of Cancer According to a Report Released By-NCDIR, n.d.). Over 200,000 cases of head and neck cancers, with oral cancers comprising 40%, occur annually in India, predominantly in regions with higher tobacco consumption rates and lower awareness levels (Head & Neck Cancers in India, n.d.).

Histological diagnosis remains the gold standard for identifying oral cavity precancerous lesions, categorized based on architectural and cytological changes observed in histological images (Neville & Day, 2002; Brennan et al., 2007). These lesions, found in the lips, cheeks, tongue, gums, and hard palate, exhibit diverse characteristics potentially stemming from various factors (Neville & Day, 2002). Common types include leukoplakia, erythroplakia, submucous fibrosis, and oral lichen planus. Oral potentially malignant disorders (OPMD) increase the risk of oral cavity cancer compared to healthy mucosa, with diagnosis requiring biopsy and histopathological confirmation. Oral Epithelial Dysplasia (OED) represents the early stages of field cancerization, marked by cellular and architectural alterations (Müller, 2017). OED severity is significant, as it predisposes the lesion to malignant progression, resembling Squamous Cell Carcinoma (SCC) histopathologically (Kademani, 2020).

Artificial intelligence (AI) systems utilize machine learning algorithms to analyze features extracted from medical images such as oral photographs, radiographs, and histopathological slides, enabling the detection of subtle changes indicative of precancerous lesions or early-stage cancers (Bhinder et al., 2021; Cheung & Rubin, 2021). AI can also analyze patient data, medical records, and imaging results to suggest treatment plans and predict outcomes (Abdul et al., 2024). For instance, a study applied six deep convolutional neural network (DCNN) models using transfer learning for identifying precancerous tongue lesions directly from clinically annotated photographic images, aiding in the diagnosis of early signs of Oral Cavity Cancer (OCC) (Shamim et al., 2022). Another study demonstrated that automated pre-screening of oral tongue lesions using DCNNs achieved ‘near-human’ level classification performance for diagnosing early signs of OCC, aiding in prevention and early diagnosis (Nath et al., 2022). Furthermore, AI has been utilized in the segmentation, classification, and detection of oral precancerous lesions (OPMDs) and oral squamous cell carcinoma (OSCC) in histopathology images (Khanagar et al., 2023). It has shown promise in diagnosing malignant tumors of the oral cavity (Abdul et al., 2024). Overall, AI in medical imaging holds immense potential for earlier diagnoses, improved treatment, and personalized care, ultimately leading to better patient outcomes (Esteva et al., 2017; Litjens et al., 2017).

2. Problem Statement and Objective of Research

Segmentation of Oral Epithelial Dysplasia (OED) is crucial for several reasons, as detailed below, particularly when employing a histological data-driven approach.

(a) Precision Diagnosis: OED encompasses a spectrum of morphological changes in oral mucosa, ranging from benign lesions to potentially malignant ones. Proper segmentation allows for precise identification and characterization of dysplastic areas within tissues, aiding in accurate diagnosis and subsequent treatment planning.

(b) Risk Assessment: By segmenting OED lesions based on histological data, clinicians can assess the degree of dysplasia and stratify patients based on their risk of malignant transformation. This information is invaluable for determining appropriate management strategies, including surveillance or intervention.

(c) Research and Development: A data-driven approach to segmentation facilitates the analysis of large datasets, enabling researchers to identify patterns, correlations, and predictive markers associated with OED progression. This information can inform the development of novel diagnostic tools, prognostic models, and therapeutic interventions.

With this motivation, we aim to develop a DL-based segmentation methodology for accurate oral epithelial layer segmentation at an architectural level. The novelty of this study lies in its comprehensive evaluation of DL models for the segmentation of OED, employing diverse architectures and rigorous performance analysis, thereby advancing computer-aided histopathological analysis and paving the way for enhanced oral health diagnostics.

3. Methodology

3.1 Dataset Used

For this study, we collected biopsy slides from Arya Wellness Center (AWC) and North East Cancer Hospital and Research Institute (NECHRI) in Guwahati, which are renowned for cancer diagnosis and treatment in the region. Image captures were conducted using a Leica DM 750 microscope equipped with the Canon ICC50 HD camera, magnified at 100x (10x objective lens and 10x eyepiece lens) with a size of 2048 × 1536 pixels. The study comprised 900 image strips, depicting normal, hyperplastic, mild, moderate, and severe dysplasia. We gathered the pathology reports to label the images. Certified pathologists delineated the region of interest (RoI), i.e., the epithelium, in certain images using Adobe Photoshop CS5 to prepare ground truth masks. The rest were completed by us under the supervision of the pathologists.

We grouped the collection of oral cavity histology biopsy images into two sets: the initial set (a) comprises original OED images, while the second set (b) contains corresponding image masks. A subset of the dataset's oral histology images and masks are illustrated in Figure 1.

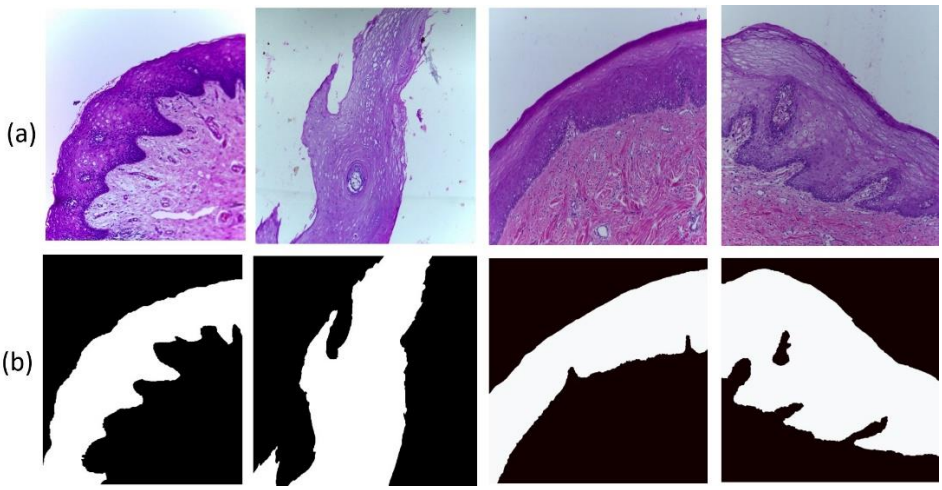

Figure 1. (a) Original image (b) Manual mask.

3.2 Data Preparation

Data preparation included first pre-processing, library importation, image loading, and resizing to 256×256×3, for consistency and effective model training. Masks were resized accordingly to ensure feature association. Next, augmentation techniques, such as horizontal and vertical flipping, expanded the dataset to mitigate overfitting risk due to limited data availability. The final dataset comprised of total 3600 images.

3.3 Model Implemented

We utilized the Vanilla UNET (Ronneberger et al., 2015) model for oral epithelial layer segmentation during training, as depicted in Figure 2. The U-Net model, a convolutional neural network (CNN) design, comprises a contracting path for contextual information extraction and a symmetric growing path for precise localization. Tailored for biomedical image segmentation, it excels in identifying and segmenting Regions of Interest (RoI) in medical images. Its distinctive structure, blending context awareness and precise localization, renders the UNET model ideal for biomedical tasks, notably in the challenging realm of medical image analysis and segmentation.

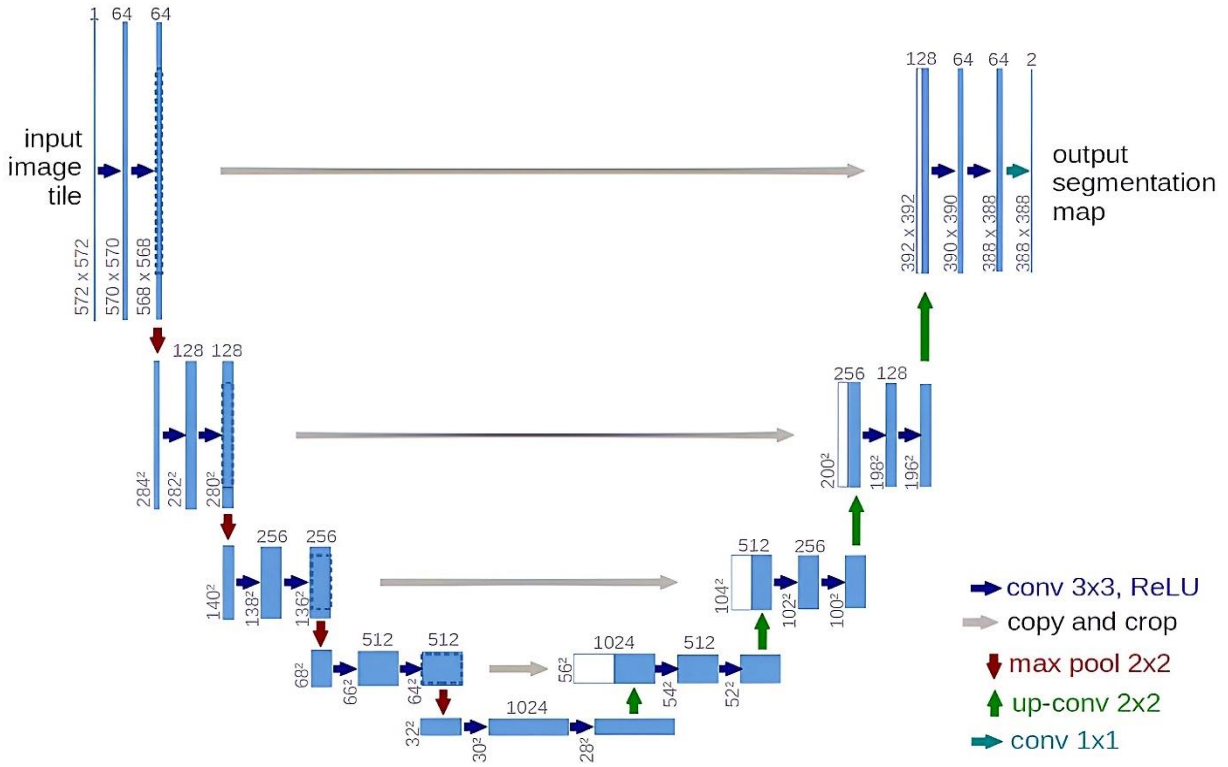

Figure 2. The vanilla UNET architecture.

Different models can serve as the encoder backbone, offering flexibility to adapt to diverse imaging characteristics and tasks while leveraging pre-trained weights to boost performance. In evaluating UNET models' performance for our application, we replaced various DL algorithms as backbones, including VGG16 (Pravitasari et al., 2020), VGG19 (Nguyen et al., 2022), MobileNet (Howard et al., 2017), and DeepLabV3+ (Chen et al., 2018). These models were incorporated into existing UNET architectures with adjusted hyperparameters. While UNET models are primarily designed for medical image segmentation, other architectures like VGG16, VGG19, MobileNet, and DeepLabV3+ have also proven effective in diverse segmentation tasks. We selected these models based on their superior oral epithelial layer segmentation (OED) performance. Each model offers distinct architectural characteristics, such as complexity, depth, and receptive fields. VGG16 and VGG19 were chosen for their strong feature extraction capabilities, MobileNet for its lightweight structure, and DeepLabV3+ for its efficient ASPP module in semantic segmentation. Table 1 shows the strengths and weaknesses of the Unet, VGG16, VGG19, MobileNet, and DeepLabv3+ models.

During model training, we set the parameters as follows: a learning rate of 0.001, batch size of 16, dropout rate of 0.2, and 200 epochs. The patience level of a UNet model during training refers to the number of epochs it waits before considering early stopping. This is crucial for preventing overfitting. The specific value for patience is determined through experimentation and validation on a separate dataset. In our case, a patience level of 10 was established by rigorous testing, and data shuffling was disabled. These parameters were carefully selected to optimize model performance and ensure robust oral health diagnostics and early disease detection through effective segmentation techniques.

Table 1. Strength and weakness of DL models.

Model Name	Pros	Cons	Example
Unet	• Excellent at capturing context and fine details due to its skip connections. • Handles small and irregular OED lesions well. • Efficient use of memory compared to deeper networks.	• May struggle with high-resolution images due to limited receptive field. • Prone to overfitting with limited data.	• Accurately segments a small, irregularly shaped OED lesion with clear boundaries.
VGG16 & VGG19	• Pre-trained on ImageNet, providing a good feature extractor for OED segmentation. • Relatively fast inference due to their shallower architecture.	• Prone to overfitting due to their complexity and limited data for OED. • May struggle with small and intricate OED lesions.	• Accurately segments a large, well-defined OED lesion with clear boundaries.
MobileNet	• Lightweight and efficient, ideal for resource-constrained environments. • Performs well on mobile devices due to its low computational cost.	• Lower accuracy compared to more complex models due to its limited capacity. • May struggle with complex OED patterns and high-resolution images.	• Accurately segments a simple OED lesion with clear boundaries on a low-resolution image.
DeepLabv3+	• State-of-the-art performance for medical image segmentation tasks. • Handles diverse tissue structures and complex OED patterns effectively. • Incorporates atrous convolution for capturing long-range dependencies.	• Requires a large amount of training data, potentially challenging for smaller datasets. • Computationally expensive compared to other models.	• Accurately segments a complex OED lesion with intricate boundaries and surrounding structures.

3.4 Platform Used

We implemented all methodologies using the Python language within the Anaconda software distribution. The coding procedures were performed using the PyCharm framework, and code execution was conducted via the Command Prompt terminal.

3.5 Performance Evaluation

The segmentation model's performance is evaluated by analyzing the following factors from the confusion matrix.

• TP (True Positive): Pixels correctly identified as oral epithelial layers.
• TN (True Negative): Pixels correctly identified as non-oral epithelial layers.
• FP (False Positive): Pixels incorrectly identified as oral epithelial layers.
• FN (False Negative): Pixels incorrectly identified as non-oral epithelial layers.

The performance metrics of the models, viz. Intersection over Union (IoU), Precision, Recall, and F1_Score, are calculated as per the mathematical formula below (Venkatachalam & Chandrabose, 2023).

$$IoU = TP / (TP + FP + FN)$$
$$Precision = TP / (TP + FP)$$
$$Recall(Sensitivity) = TP / (TP + FN)$$
$$F1\_Score = 2 * (Precision * Recall) / (Precision + Recall).$$

These metrics were computed throughout the model training process for OED segmentations. Higher values of IoU, Precision, Recall, and F1_Score signify precise segmentation of Oral Epithelial Dysplasia (OED). Precision ensures accurate identification of OED regions, reducing false positives during biopsies, while Recall reflects the model's ability to avoid missing 'true' OED regions, crucial for early detection. The F1-score, balancing precision and recall, comprehensively assesses the model's effectiveness, providing a fair

evaluation during training, akin to a discerning judge considering detection accuracy and thoroughness. These metrics are pivotal for assessing segmentation algorithms.

Additionally, we computed the model's loss using the "binary cross-entropy" loss function. This loss function is commonly used during the training phase of segmentation models, such as pixel-wise binary segmentation, where each pixel in the output is classified as either a target object (OED) or background. The binary cross entropy loss calculates the difference between the predicted probabilities and the actual target labels for each pixel, penalizing deviations from the ground truth.

The mathematical expression for binary cross-entropy is as follows:
Given:

y - the ground truth label (0 or 1)

$\hat{y}$ - the predicted probability that y =1.

The binary cross-entropy loss function $L_{\text{binary cross-entropy}}$ is defined as:

$$L_{\text{binary cross-entropy}}(y, \hat{y}) = -(y \cdot \log(\hat{y}) + (1-y) \cdot \log(1-\hat{y})).$$

where,

- If y = 1, the loss reduces to $-\log(\hat{y})$, penalizing the model more when it predicts a low probability for the positive class.
- If y = 0, the loss reduces to $-\log(1-\hat{y})$, penalizing the model more when it predicts a high probability for the negative class.
- The negative sign is applied to ensure that the loss is minimized during training.

It encourages the model to output high probabilities for the correct class and low probabilities for the incorrect class. During training, the model's parameters are updated iteratively using gradient descent to minimize the binary cross-entropy loss function, thereby improving the model's ability to make accurate predictions.

For our application we set the learning rates differently for various models, with 0.001 for vanilla UNET and 0.001 for other models (VGG16, VGG19, MobileNet, and DeepLabv3+), respectively, to optimize training performance.

4. Results and Discussion

During the training phase, we implemented a dropout rate of 0.20 across the entire model, randomly excluding 20% of neurons on average during each training iteration. This deliberate use of dropout aims to enhance the model's generalization capacity, effectively reducing overfitting and promoting the extraction of robust features. The dataset of 900 images was divided into an 80:20 ratio for training and validation, resulting in 2880 images for training and 720 images for validation after augmentation. The "Adam" optimizer was consistently applied to all models - VGG16, VGG19, MobileNet, and DeepLabv3+ throughout the training process. Table 2 below displays the outputs of all the models during the training of the OED images. A close look at the values in the table reveals that the Vanilla UNET model exhibits lower loss compared to other DL models used in the study.

Table 2. Performance metrics of the implemented models for comparison.

Models	Loss	IoU	Precision	Recall	F1-Score
UNet	**0.06**	**93.37**	**97.96**	**97.78**	**96.76**
Unet-VGG16	0.07	93.23	97.52	96.98	96.24
Unet-VGG19	0.07	93.26	97.35	96.60	95.69
Unet-MobileNetV2	0.07	93.17	97.21	97.49	96.40
DeepLabV3+	0.07	92.93	97.04	97.16	96.53

The subsequent figure (Figure 3**)** offers a comprehensive visualization of the behaviour of the performance metrics of the UNET model. Training and validation plots offer insight into many aspects of the performance of the model. It helps in the interpretation of the performance metrics.

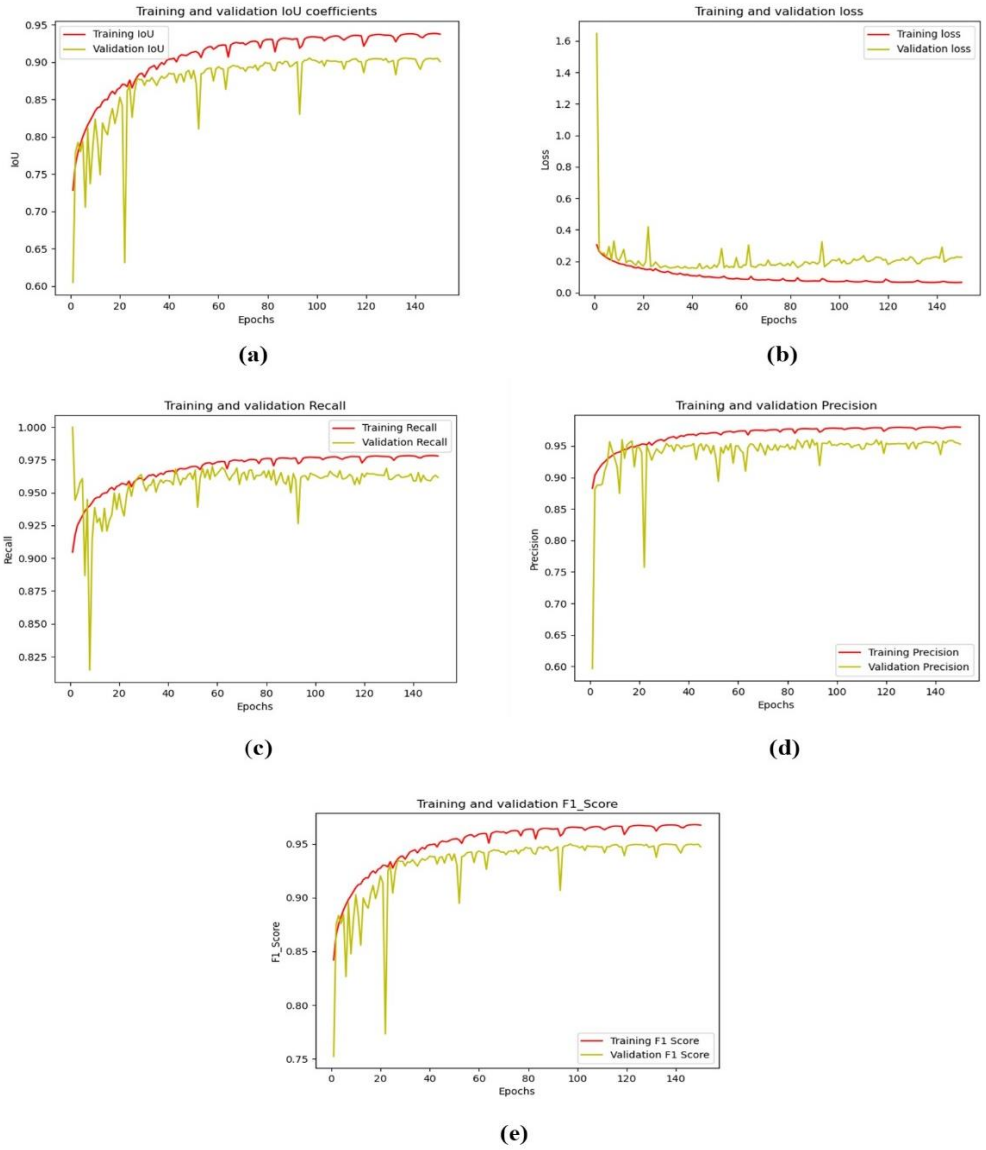

Figure 3. UNET model's training and validation curves of (a) IoU, (b) Loss (c) Precision, (d) Recall (e) F1-Score.

(a) IoU: The X-axis of the IoU plot typically represents different IoU thresholds, ranging from 0 to 1. Here, we observe that as the model progresses, there are higher and more stable matches. Our model displays consistent performance across a wide range of IoU thresholds, which is generally considered more robust.

(b) Loss: The training curve shape reveals insights into overfitting or underfitting. Plateaus or fluctuations may indicate convergence issues or unstable optimization, suggesting the need for hyperparameter adjustments like learning rate tuning. A curve converging to a low and stable loss signifies effective learning and readiness for evaluation on unseen data., as in our case.

(c) Recall: The shape of the Recall plot and the distribution of points along the X-axis in our case provide insights into the overall good performance of the segmentation model.

(d) Precision: The Precision plot demonstrates high values across thresholds, reflecting the model's ability to identify positive instances while minimizing false positives.

(e) F1_Score: A higher F1 Score signifies better overall model performance in balancing precision and recall, indicating effective segmentation by our model.

In the final phase, we extensively evaluated all the trained models on 18 new oral cavity OED histopathology datasets to gauge the segmentation performance of the images. It was found that the UNET model segmented the oral epithelial layers in histopathology images of OED most successfully. Figure 4 illustrates the segmented data as a visual representation of the testing phase's results. The visualization of the segmentation process unfolds across three detailed panels. Firstly, the unaltered, raw image—serving as the foundational input data for the segmentation task—is prominently featured in the initial panel. This unprocessed image provides the contextual backdrop against which the subsequent segmentation output will be assessed.

Moving to the second panel, we note the predicted mask generated by the meticulously trained model. This mask, meticulously crafted through the model's learning process, serves as a critical guidepost for the segmentation task. It delineates areas of interest and outlines the boundaries of objects within the image, laying the groundwork for precise segmentation.

Moving to the third panel offers a culmination of the segmentation journey. Here, the predicted mask is superimposed onto the original image, resulting in a visual amalgamation. This final display showcases the intricate interplay between the predicted mask and the underlying image, providing a visualization of the segmentation outcome. Through this layered presentation, Figure 4 offers an exploration of the segmentation process, from raw input to refined output.

The segmented outputs provide compelling evidence for the effectiveness of UNET models in accurately identifying the intricate structures present in histopathology images of the oral cavity. This milestone marks a significant advancement with the successful segmentation of Oral Epithelial Layers, suggesting that top-performing models could aid pathologists and clinicians in precisely identifying and analyzing oral diseases. This accomplishment underscores the potential of the UNET model to significantly contribute to computer-aided diagnostics by offering a valuable tool for the precise and efficient evaluation of pathological conditions in oral tissues.

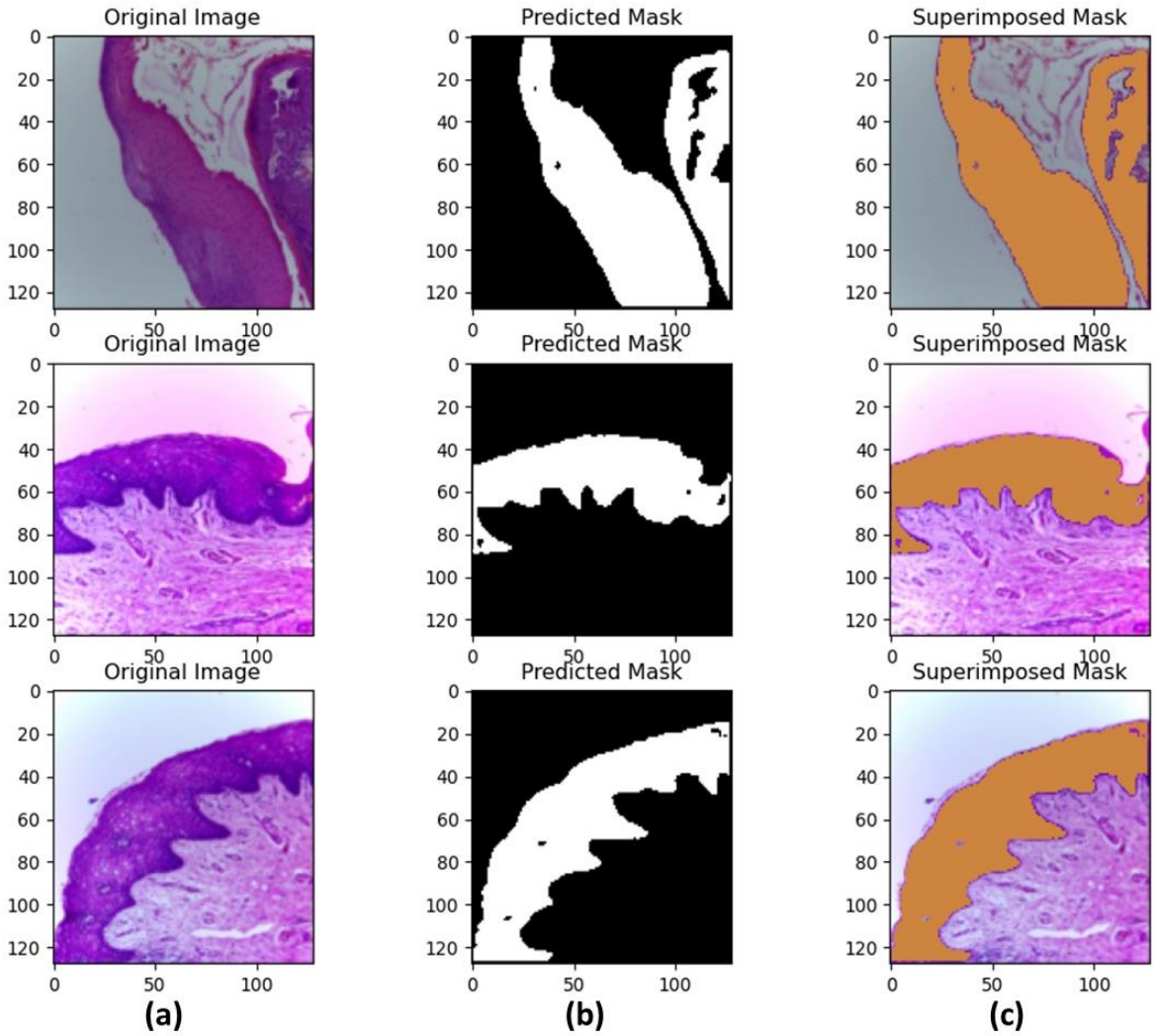

Figure 4. Testing output of the UNET Model; column (a) Original images, (b) Model predicted masks, (c) Predicted mask are superimposed over the original images.

4. Conclusion

In conclusion, our study emphasizes the critical role of segmentation in oral cavity health diagnostics, particularly for identifying precancerous lesions like OED. Leveraging DL models such as UNET, VGG16, VGG19, MobileNet, and DeepLabV3+, we achieved accurate segmentation of OED lesions in histopathological images. The UNET model emerged as the top performer, showcasing superior IoU: 93.73%, precision: 97.96%, recall: 97.78%, and F1_score: 96.76% metrics. Our findings underscore the potential of DL-based segmentation techniques in revolutionizing oral health diagnostics, offering clinicians valuable insights for early disease detection and treatment planning. Through meticulous segmentation and analysis of oral tissue structures, these models pave the way for enhanced patient care and improved outcomes in oral cancer management. Future research may investigate integrating multimodal data sources and exploring novel DL architectures to enhance OED segmentation models' precision and predictive power and improve oral health diagnostic tools.

Conflict of Interest

As far as this article is concerned, there is no conflict of interest among the authors. During this phase, all ethical protocols were upheld. This study has the appropriate institutional ethics approval (No. IEC(HS)/IASST/1082/2023/3 Date: 02/03/2023). For this investigation, no publicly available datasets were discovered. On an indigenous collection of datasets, segmentation work is done.

Acknowledgments

Mr. Taibur Rahman would like to thank the Mathematical and Computational Sciences lab members of the Institute of Advanced Study in Science and Technology (IASST) for their valuable suggestions and motivation for this manuscript. Mr. T Rahman also would like to thank Arya Wellness Centre (AWC) and North East Cancer Hospital and Research Institute (NECHRI) laboratory technicians and clinicians for their support.

References

Abdul, N.S., Shivakumar, G.C., Sangappa, S.B., Di Blasio, M., Crimi, S., Cicciù, M., & Minervini, G. (2024). Applications of artificial intelligence in the field of oral and maxillofacial pathology: A systematic review and meta-analysis. *BMC Oral Health*, *24*(1), 122. https://doi.org/10.1186/s12903-023-03533-7.

Bhinder, B., Gilvary, C., Madhukar, N.S., & Elemento, O. (2021). Artificial intelligence in cancer research and precision medicine. *Cancer Discovery*, *11*(4), 900-915. https://doi.org/10.1158/2159-8290.CD-21-0090.

Borse, V., Konwar, A.N., & Buragohain, P. (2020). Oral cancer diagnosis and perspectives in India. *Sensors International*, *1*, 100046. https://doi.org/10.1016/j.sintl.2020.100046.

Brennan, M., Migliorati, C.A., Lockhart, P.B., Wray, D., Al-Hashimi, I., Axéll, T., Bruce, A.J., Carpenter, W., Eisenberg, E., Epstein, J.B., Holmstrup, P., Jontell, M., Nair, R., Sasser, H., Schifter, M., Silverman, B., Thongprasom, K., Thornhill, M., Warnakulasuriya, S., & van der Waal, I. (2007). Management of oral epithelial dysplasia: A review. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, *103*, S19-e1- S19.e12. https://doi.org/10.1016/j.tripleo.2006.10.015.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the European Conference on Computer Vision* (pp. 801-818). http://arxiv.org/abs/1802.02611.

Cheung, H.M.C., & Rubin, D. (2021). Challenges and opportunities for artificial intelligence in oncological imaging. *Clinical Radiology*, *76*(10), 728-736.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, *542*(7639), 115-118. https://doi.org/10.1038/nature21056.

Head & neck cancers in India. (n.d.). Retrieved February 2, 2024, from https://fhno.org/blog/Head-And-Neck-Cancers-in-India.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

Kademani, D. (2020). *Improving outcomes in oral cancer*. Springer Nature, Switerzerland. https://doi.org/10.1007/978-3-030-30094-4.

Khanagar, S.B., Alkadi, L., Alghilan, M.A., Kalagi, S., Awawdeh, M., Bijai, L.K., Vishwanathaiah, S., Aldhebaib, A., & Singh, O.G. (2023). Application and performance of artificial intelligence (ai) in oral cancer diagnosis and prediction using histopathological images: A systematic review. *Biomedicines*, *11*(6), 1-17. https://doi.org/10.3390/biomedicines11061612.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., & Sánchez, C.I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60-88. https://doi.org/10.1016/j.media.2017.07.005.

Müller, S. (2017). Update from the 4th edition of the world health organization of head and neck tumours : tumours of the oral cavity and mobile tongue. *Head and Neck Pathology*, *11*(1), 33-40. https://doi.org/10.1007/s12105-017-0792-3.

Nath, S., Raveendran, R., & Perumbure, S. (2022). Artificial intelligence and its application in the early detection of oral cancers. *Clinical Cancer Investigation Journal*, *11*(1), 5-9. https://doi.org/10.51847/h7wa0uhoif.

Neville, B.W., & Day, T.A. (2002). Oral cancer and precancerous lesions. *CA: A Cancer Journal For Clinicians*, *52*(4), 195-215.

Nguyen, T.H., Nguyen, T.N., & Ngo, B.V. (2022). A VGG-19 model with transfer learning and image segmentation for classification of tomato leaf disease. *AgriEngineering*, *4*(4), 871-887. https://doi.org/10.3390/agriengineering4040056.

North Eastern States of India have a high burden of cancer according to a report released byICMR-NCDIR. (n.d.).

Pravitasari, A.A., Iriawan, N., Almuhayar, M., Azmi, T., Irhamah, Fithriasari, K., Purnami, S.W., & Ferriastuti, W. (2020). UNet-VGG16 with transfer learning for MRI-based brain tumor segmentation. *Telkomnika (Telecommunication Computing Electronics and Control)*, *18*(3), 1310-1318. https://doi.org/10.12928/TELKOMNIKA.v18i3.14753

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention, 18th International Conference* (pp. 234-241), Springer International Publishing. Munich, Germany.

Shamim, M.Z.M., Syed, S., Shiblee, M., Usman, M., Ali, S.J., Hussein, H.S., & Farrag, M. (2022). Automated detection of oral precancerous tongue lesions using deep learning for early diagnosis of oral cavity cancer. *The Computer Journal*, *65*(1), 91-104. https://doi.org/10.1093/comjnl/bxaa136.

Siegel, R.L., Miller, K.D., Wagle, N.S., & Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, *73*(1), 17-48. https://doi.org/10.3322/caac.21763.

The global status report on oral health 2022. (n.d.). Retrieved February 2, 2024, from https://www.who.int/team/noncommunicable-diseases/global-status-report-on-oral-health-2022/

Varshitha, A. (n.d.). *Prevalence of oral cancer in India*.

Venkatachalam, J., & Chandrabose, S. (2023). Optimizing region detection in enhanced infrared images using deep learning. *Revue d'Intelligence Artificielle*, *37*(4), 1015-1021. https://doi.org/10.18280/ria.370423.