Enhancing Dental Diagnostics: Advanced Image Segmentation Models for Teeth Identification and Enumeration

Mohsin Ali[1] [0000-0002-5409-7368], Moin Hassan[1], Esra Konsa[3], John Q Gan[1], Akhilanand Chaurasia[2], and Haider Raza[1]

[1] School of Computer Science and Electronics Engineering, University of Essex, Colchester, UK
ma22159@essex.ac.uk
[2] Faculty of Dental Sciences, King George Medical University, Lucknow, India
[3] Department of Periodontology, Oral Medicine and Oral Surgery, Berlin, Germany

Abstract. With recent advancements in Artificial Intelligence (AI) influencing various medical fields, dentistry faces several challenges. Among these challenges, accurate tooth counting and identification are essential for effective treatment and oral health monitoring. While several approaches exist for tooth identification and counting, they often entail drawbacks such as high costs or excessive manual labour. Panoramic X-ray imaging, a cost-effective and widely utilized method, is vital in dental healthcare, aiding in treatment planning and monitoring patient progress pre- and post-treatment. However, the complexity of panoramic X-rays, including non-uniform tooth shapes, misalignment, and overlapping teeth, pose challenges in tooth identification and counting. This study presents a novel approach to address these challenges by introducing a tooth identification and counting technique using advanced image segmentation models. We comprehensively evaluate multiple segmentation models, such as U-Net, Attention U-Net, Feedback U-Net, and Feedback U-Net with LSTM, specifically tailored to panoramic X-ray images, utilizing the open-source Tufts Dental Dataset. Our analysis demonstrates that the U-Net model surpasses other evaluated segmentation models for panoramic X-ray image segmentation because it can be effectively trained with limited datasets, which is crucial in dentistry where extensive labelled data is often unavailable. The primary goal of this research is to develop a technique that assists dental professionals in accurately identifying and counting teeth, thereby enhancing treatment planning and patient diagnosis. Code available on https://github.com/game-sys/Dental-Segementation-and-Enumeration

Keywords: Image Segmentation · Tooth Identification · Teeth Counting

1 Introduction

The integration of AI models into various fields has seen remarkable advancements, with the dental healthcare sector being no exception [1]. In recent years, the adoption of deep learning techniques has surged, fundamentally transforming practices within dentistry [2]. From analyzing dental radiographs to assessing patient risk, deep learning models have become invaluable tools in optimizing dental treatment. However, some challenges persist in dentistry that require AI-driven solutions, and one such challenge is tooth identification and counting [3]. Accurate identification and counting of teeth are crucial for treatment planning and monitoring treatment progress. While researchers have proposed techniques for teeth identification and counting, many rely on expensive [4] and uncommon methods [5], limiting their adoption in the dental field. Dental radiographs, particularly panoramic X-rays, are commonly used by dentists for tooth identification. However, interpreting panoramic X-rays can be challenging due to the complex structure of the jaw, as well as misaligned and overlapping teeth. Perceptual errors in interpreting panoramic X-rays contribute to a significant percentage of misdiagnoses in dentistry, estimated at 60-80% [6]. This can lead to incorrect tooth extractions, with a reported ratio of oversight as high as 21% [6], especially prevalent in underdeveloped countries where dental resources may be limited. These complexities highlight the urgent need for AI tools capable of automatically counting teeth and identifying those requiring treatment. Such tools could reduce the risk of performing therapy on the wrong tooth and improve treatment accuracy. Moreover, integrating deep learning models with dental radiology could empower dentists to conduct more precise dental treatments and streamline dental procedures.

The advancement in computation power has led to many computer vision (CV) solutions that led to the development of multiple tools in dental applications such as tooth classification [7] and cavity detection [8]. Among these advancements, image segmentation, a component of CV applications, has gained significant attention in recent years. Image segmentation involves annotating pixels for different parts of images, enabling more precise analysis and understanding of visual data. Fully convolutional networks (FCNs) have emerged as popular architectures for image segmentation tasks [9] [10]. In particular, U-Net [11], an advanced deep-learning model, has gained importance in this domain [12]. U-Net is favoured for its capability of being effectively trained with small datasets, a challenge often encountered with other types of FCNs. This capability is particularly relevant in dental healthcare, where datasets tend to be limited in size but rich in complexity. Thus, U-Net's adaptability to such datasets makes it a valuable tool for image segmentation in dental applications.

The main contributions of this paper are as follows:

- This study proposes a framework for highly accurate identification and counting of individual teeth, empowering dental experts to make informed treatment decisions, as illustrated in Fig. 7.
- Additionally, the study evaluates various image segmentation models, such as U-Net, Attention U-Net, Feedback U-Net, and Feedback U-Net with LSTM, on panoramic X-rays utilizing the Tufts Dental Database, a widely acknowledged benchmark dataset in the dental domain [13].

This paper is structured as follows: Section 2, presents a comprehensive overview of the historical evolution of panoramic X-ray and the challenges associated with its interpretation. Section 3, the methodology section, offers detailed insights into the dataset and provides information on the image segmentation models used. It further describes the experimental setup and outlines the functioning of the framework for tooth identification and counting. Section 4 presents the findings of comparing multiple image segmentation models, along with an explanation of the output generated by the proposed framework. Finally, Section 5 summarizes the study.

2 Related Work

Several deep-learning techniques have been proposed for teeth identification. However, many of these studies utilize expensive methods such as 3D imaging, despite demonstrating accurate tooth identification. Unfortunately, these approaches are not widely adopted in dentistry due to the high costs associated with specialized equipment requirements [14]. Furthermore, more commonly used manual methods utilize panoramic X-rays. However, these techniques require significant human resources and dental expertise for tooth identification [15]. Panoramic radiography, commonly referred to as panoramic X-ray holds prominence in the field of dentistry due to its cost-effectiveness, low radiation levels, and ability to provide a comprehensive view of the dental framework and surrounding bone structures [16]. Panoramic X-rays do present challenges as they summarise three-dimensional space into a two-dimensional image, potentially leading to overlapping structures and making interpretation difficult. Similarly, specific acquisition methods of panoramic X-ray can introduce distortions. Further, patient positioning may compromise the clarity and quality of panoramic X-ray [17]. Due to this reason, panoramic X-rays require dental experts to interpret and identify the individual tooth.

Image segmentation is a popular method for annotating multiple objects within an image. In recent years, image segmentation has gained popularity in dental healthcare, as it aids in outlining specific sections of images, providing critical visual insights essential for diagnosis. Multiple frameworks have been designed utilizing image segmentation techniques to segment different parts of the panoramic X-rays. Traditional manual techniques employed by radiologists are time-consuming and yield inconsistent results, necessitating automated solutions, especially given the rapid advancements in medical imaging technology. Numerous frameworks have been introduced to automate segmentation tasks in healthcare, such as "Image Engineering" [18], a comprehensive framework that emphasizes image analysis and positions segmentation as a crucial component. This automated approach is vital for improving the accuracy of subsequent analyses in medical contexts, spanning diagnostics to computer-assisted surgeries. The inherent complexity of human anatomy ensures that medical image segmentation remains a challenging field, requiring continual innovation and refinement of segmentation techniques to meet the evolving needs of healthcare professionals [19].

U-Net [11] stands out as a popular image segmentation architecture utilized in various applications, including healthcare, particularly dentistry, where it has been innovatively applied to segment dental panoramic X-ray images. Radiographic images such as panoramic X-rays play a pivotal role in dentistry for accurate diagnosis, yet inherent complexities and noise may compromise their clarity. Therefore, dental image segmentation becomes indispensable, assisting professionals in identifying impactions, pinpointing implant sites, and understanding dental alignments. The effectiveness of U-Net in dental image segmentation has been compared to other segmentation solutions, underscoring its potential in this domain. Image segmentation forms the foundation for many healthcare applications facilitating early detection of diseases and refining diagnostic precision, thereby enhancing patient outcomes and streamlining healthcare processes [20].

There is a growing focus on automating dental restoration segmentation to improve diagnostics and tailor treatment plans [21]. Traditional manual methods may occasionally miss restorations due to factors like dentist fatigue or varying levels of expertise. To address this challenge, numerous image segmentation techniques employing deep learning models have been proposed for tasks such as identifying dental fillings or detecting missing teeth [21]. Notably, one study [22] utilized the U-Net model to segment both amalgam and composite fillings. This approach presents a promising advanced diagnostic tool for dental practitioners, facilitating more accurate identification and analysis of dental restorations, which is crucial for effective treatment planning and patient care.

Advancements in medical imaging and software technologies have brought significant improvements to diagnostic practices in dentistry. While panoramic X-rays play a crucial role in detecting various oral conditions, their manual analysis can be labour-intensive and susceptible to errors. Previous research has often focused on individual types of panoramic X-rays. However, the highlighted study stands out by employing a transformer-integrated U-Net model to segment both panoramic and bitewing images, achieving noteworthy results. Panoramic X-rays produce images based on tissue radiological density, providing invaluable insights into hidden dental issues. With various dental X-ray methods available, the transition towards automated tools utilizing deep learning models, particularly the transformer-augmented U-Net, represents a significant advancement in accuracy and efficiency [23]. This approach holds promise for streamlining diagnostic workflows and improving patient outcomes in dentistry.

3 Materials and Methodology

3.1 Dataset

The dataset used in this research is Tufts Dental Database (TDD) [13], which serves as a pivotal tool for dental radiograph studies, offering a collection of 1,000 panoramic radiographs that detail dental irregularities, teeth structures, and crucial maxillomandibular regions, as shown in Fig. 1 and 2. Unlike other types of X-rays, panoramic X-ray images have the unique advantage of showcasing the entire oral area, encompassing all teeth in both jaws and indicating the positions of fully and partially emerged teeth in one image. An integral aspect of the TDD is its ability to encapsulate radiologist expertise. By integrating eye-tracking and the think-aloud protocol, the database effectively captures the nuanced relationship between the radiologist's perception of abnormalities and their cognitive interpretation. This is further enriched by the inclusion of gaze maps and recorded oral observations, stored in a .json file, shedding light on the radiologist's analytical approach and thought processes during image assessments [13].

Fig. 1. panoramic X-ray

3.2 Image Segmentation

The implementation of AI models on dental panoramic X-rays requires a careful balance between image resolution and computational efficiency. This study acknowledges the trade-off between higher resolution (e.g., 512x512 pixels) and the computational demand it entails. Moreover, lower resolution (e.g., 256x256 pixels) is chosen for its resource efficiency but raises concerns about potential information loss, especially in capturing finer anatomical details essential for precise dental diagnostics. The selection of segmentation model architecture is crucial in this investigation. The U-Net [11], Attention U-Net [24], Feedback U-Net [25] and Feedback U-Net with LSTM [26] were selected as image segmentation models as suggested by the literature. By pairing specific resolutions with appropriate models, this study aims to identify the optimal combination that maximizes segmentation accuracy and to develop a methodology for maximizing segmentation accuracy, which includes accurately counting and visually identi-

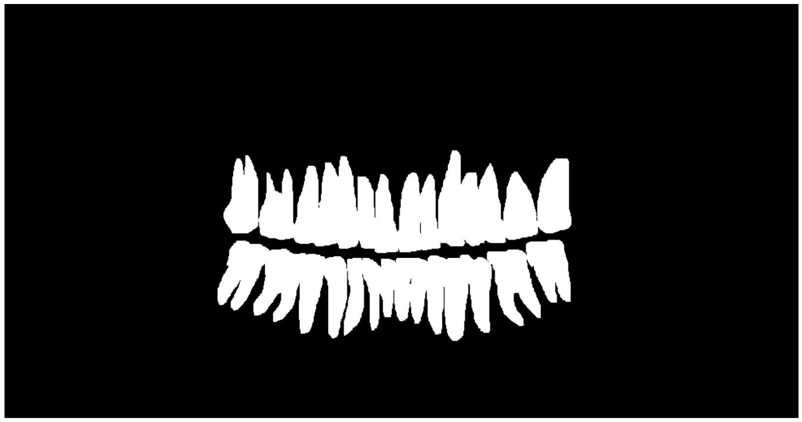

Fig. 2. Segmented mask

fying individual teeth from panoramic X-ray images, as demonstrated in Fig. 7.

U-Net Model: U-Net is a widely used biomedical image segmentation method distinguished by its unique architecture, tailored for limited biomedical datasets. Its contracting encoder efficiently reduces spatial dimensions while deepening feature channels, capturing intricate details from input images. The subsequent expansive decoder, featuring skip connections, facilitates precise feature localization. The overlap-tile strategy enables seamless image processing, overcoming GPU memory constraints and ensuring comprehensive coverage of input data. To address limited data, U-Net incorporates aggressive data augmentation techniques, enhancing its robustness in practical applications where large annotated datasets are scarce [11].

Feedback U-Net Model: The innovative Feedback U-Net establishes a symbiotic relationship between primary and feedback pathways, enhancing segmentation performance. The encoder crafts feature maps through meticulous downsampling, while the feedback convolution block enriches output with contextual understanding. In the decoder, up-sampled feature maps seamlessly integrate with feedback information through skip connections, ensuring refined details merge harmoniously with context. Short skip connections maintain semantic alignment, capturing local details while preserving global context. This synergy enhances segmentation accuracy and preserves crucial contextual cues within biomedical images [25].

Feedback U-Net Model with LSTM: The Feedback U-Net model with LSTM enhances the traditional U-Net architecture by introducing a feedback loop that redirects the model's output back to its input layer. This iterative process generates probability maps for each class, with subsequent iterations utilizing feedback from previous outputs. The model incorporates convolutional LSTMs for richer feature extraction, employing distinct batch normalizations for each iteration. The training utilizes two loss functions based on softmax cross entropy, one for the initial iteration and another for subsequent iterations. This architecture combines feedback loops with LSTM's memory capabilities, yielding more context-sensitive and accurate segmentation results, representing a promising advancement in segmentation methodologies [26].

Attention U-Net Model: The Attention U-Net Model merges the encoder-decoder architecture with the SE-Residual block, enhancing feature extraction depth and efficiency. It introduces the MixPool block, which iteratively refines segmentation maps using masks from previous cycles, initiated by a mask generated through the Otsu thresholding technique. FANet operates as a Fully Convolutional Neural Network with four encoder and decoder segments. The encoder comprises dual 3x3 convolutions in its SE Residual blocks, guided by hard attention from the MixPool block. The decoder expands feature maps via a 4x4 transpose convolution, merges them through skip connections, and integrates them with prior epoch's segmentation mask, facilitating iterative refinement [24].

3.3 Experimental Setup

In this study, the Tufts Dental Dataset was divided into three distinct subsets: training, validation, and testing. The division was carried out with 60% of the data allocated to training, 20% to testing, and the remaining 20% to validation. This partitioning strategy ensured an appropriate distribution of data for model training, validation, and subsequent evaluation. During training, we employed the Adam optimizer due to its efficient optimization capabilities and adaptive learning rate with an initial learning rate set to 1e-4. To maintain consistency with prior research efforts, we applied identical augmentation techniques to DENTECT [27] to enhance dataset diversity and improve model generalization. Each model training was done on 200 epochs, with early stopping mechanisms implemented to prevent overfitting and ensure optimal model performance. The PyTorch deep learning framework was utilized. Additionally, GPU acceleration was leveraged using two Nvidia 3080Ti GPUs, enabling accelerated computation. This carefully designed experimental setup aimed to ensure consistency, reproducibility, and comparability across various image segmentation models trained on the Tufts Dental Dataset.

3.4 Tooth Identification and Counting

When applied to panoramic dental X-ray images, segmentation requires high precision due to the intricate and closely situated structures within the oral cavity. While model outputs offer a foundational segmentation, they sometimes carry imperfections. These may arise from overlapping teeth, varying intensities, or

Fig. 3. Generating contours on the segmented mask of the panoramic X-ray image from the dataset

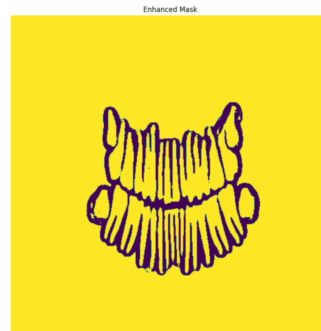

Fig. 4. Adaptive thresholding on predicted mask

artefacts inherent to radiographic procedures. Acknowledging these intricacies, a post-processing pipeline specifically tailored for dental X-ray images was devised to identify individual teeth. This pipeline draws inspiration from a similar method outlined by helli et al. [28]. Additionally, an algorithm was developed to accurately count the total number of teeth from the segmented image, ensuring refined and clinically meaningful segmentations.

Drawing Contours on Segmented Masks: Contours play a vital role, especially in dental X-rays, where the boundary between two adjacent teeth or between tooth and gum is paramount. By drawing contours on the segmented masks, we emphasize these boundaries, ensuring that each tooth and structure within the oral cavity stands out distinctly, aiding both visualization and subsequent analyses as shown in Fig. 3.

Adaptive Thresholding: The initial phase of our post-processing involves adaptive thresholding. Dental X-rays exhibit specific intensity variations, especially between the hard dental structures and the surrounding softer tissues. Adaptive thresholding tailors the threshold values region-wise, accommodating these intensity differences. Algorithmically, this is achieved by calculating the mean or median value of the intensities within a predefined window around each pixel and setting the threshold based on these local statistics. The outcome is an enhanced mask that accurately represents dental features as shown below in Fig.4.

Erosion: Subsequent to thresholding, we employ the erosion technique on the mask. As a morphological tool, erosion assists in minimizing noise and ridding

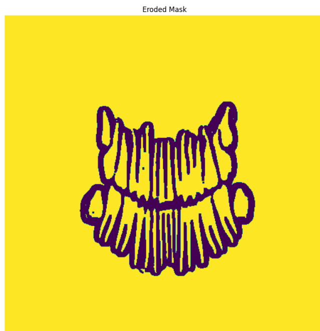

Fig. 5. Erosion on predicted mask

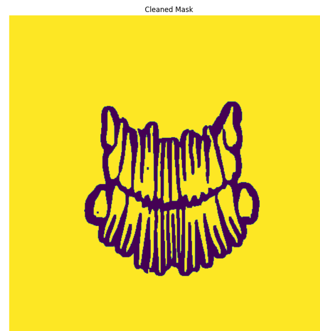

Fig. 6. Removing noise on mask

the mask of minor inconsistencies. This is achieved by convolving a structuring element with the mask and retaining only the regions where the template fits completely within the white areas of the mask. By strategically retracting the boundaries of the foreground (primary teeth in this context), we ensure unwanted specks or variations are curtailed, facilitating more distinct segmentation as shown in Fig. 5.

Addressing and Eliminating Fine Lines: Post-erosion, we tackle the challenge of subtle lines or unintended structures that may linger. Such streaks, possibly arising from imaging discrepancies or artefacts, could disrupt the true depiction of teeth. To eliminate these fine lines, we employ a combination of further morphological operations such as dilation and additional erosion, as well as image filters that target and remove high-frequency details. By addressing these, we are left with a mask that solely emphasises the salient dental structures as shown below in Fig. 6.

Teeth Counting: To count the teeth, we utilize a Connected Component Analysis (CCA) algorithm, which operates on the binary segmented image. CCA assigns a unique label to each connected region within the binary image. Two components are considered connected if they possess similar pixel values. This labelling process is performed pixel by pixel, and upon completion, each component receives a distinct label value from the CCA. Once CCA labelling is finalized, we calculate the number of unique components present in the image. Each unique component corresponds to a tooth, allowing us to determine the total count of teeth in the image as shown in Fig. 7.

Tooth Identification: The Watershed algorithm is a classic image segmentation method grounded in visualizing grayscale images as topographic relief,

Fig. 7. Final color-coded segmentation with teeth count

where high-intensity regions correspond to peaks and low-intensity regions represent valleys. When we metaphorically flood this topographic landscape from its valleys, the emerging watershed lines, or barriers, naturally demarcate regions of interest, or in our case, individual teeth [29]. Following the application of the Watershed algorithm, each tooth within the mask is then color-coded. This step goes beyond mere aesthetic appeal. By providing each tooth with a unique colour, we simplify subsequent analytical tasks such as examination, quantification, and detailed evaluation. Such an approach not only aids in the immediate visual differentiation of adjacent teeth but also enriches the explainability of our results, making them especially valuable for clinical assessments as shown in Fig. 7.

Table 1. Comparison of performance of image segmentation models.

Image size 256 x 256				
Model	**Attention U-Net**	**Feedback U-Net LSTM**	**U-Net** +	**Feedback U-Net**
Recall	**95.27**	94.59	93.59	92.87
Precision	78.91	85.53	**90.26**	82.97
F1 score	86.31	89.75	**91.86**	87.55
Dice Score	86.32	89.83	**91.89**	87.65
IoU	75.94	81.54	**85.01**	78.13
Accuracy	96.14	97.26	**97.89**	96.60
Image size 512 x 512				
Recall	93.62	95.34	**96.37**	95.14
Precision	79.48	88.21	**91.53**	85.21
F1 score	85.83	91.58	**94.00**	89.98
Dice Score	88.33	90.42	**92.19**	89.05
IoU	79.51	84.73	**86.42**	81.22
Accuracy	95.67	97.90	**98.28**	97.14

4 Results and Discussion

Panoramic dental X-ray images are essential for capturing the intricate structures of the dental arch. However, due to their complexity and diverse densities, evaluating the effectiveness of models used for analysis requires a comprehensive approach. We employ a diverse set of metrics to assess the image segmentation model's performance: 1) Sensitivity, also known as recall, is a crucial metric that evaluates how well the model identifies genuine dental regions. It emphasizes the importance of detecting most dental structures to ensure comprehensive analysis. 2) Precision is another vital metric that measures the authenticity of the model's predictions. It ensures that the model accurately represents dental structures with minimal errors, thereby enhancing the reliability of the analysis. 3) Accuracy provides an overall perspective on the model's ability to differentiate dental structures from other elements present in the image. It assesses the correctness of the model's classifications, contributing to the overall effectiveness of the analysis. 4) F1 score, which combines recall and precision, offers a balanced assessment by considering both missed detections and false positives. This metric provides valuable insights into the model's performance, particularly in scenarios with imbalanced data distributions. 5) Dice Coefficient evaluates the alignment of the model's segmentation with the actual dental regions in the image. Scores approaching 1 indicate high accuracy and alignment between the predicted and actual regions, enhancing the reliability of the analysis. 6) Intersection over Union (IoU) measures the overlap between the model's predictions and the actual dental regions, further assessing the model's precision. This metric quantifies the agreement between predicted and ground truth regions, contributing to the overall validation of the model's capabilities. Collectively, these metrics validate the model's capability in interpreting the complexities of panoramic dental images and underscore its clinical relevance. To assess the performance of various neural network architectures, a comparative analysis was conducted on two image resolutions: 256×256 and 512×512, as shown in Table 1.

For the 256x256 images, the model with the highest precision is the U-Net model, achieving 90.26%, while the Attention U-Net model demonstrates the highest recall of 95.27%. However, the F1 score, which balances both recall and precision, is highest for the U-Net model at 91.86%. Similarly, the Dice Score and Intersection over Union (IoU) metrics, which evaluate segmentation accuracy, also show the U-Net model outperforming others. When considering images of size 512x512, the U-Net model maintains its dominance across various metrics, achieving the highest recall, precision, F1 score, Dice Score, IoU, and accuracy. This indicates its robustness and effectiveness in segmenting larger images. Overall, the results suggest that the U-Net model consistently outperforms other models across different image sizes, demonstrating its versatility and effectiveness in dental image segmentation tasks. The U-Net based method developed in this study not only finds and outlines each tooth in dental X-ray images but also colour-codes them, making it easier for dentists to see which teeth need attention, as shown in Fig. 7. For example, we might use different colours to highlight areas that need filling or treatment. Additionally, our method counts the total number of teeth, providing a quick assessment of the patient's dental health. This visual representation helps dental experts interpret the X-rays more effectively, leading to better treatment decisions. The integration of our U-Net-based segmentation framework with existing panoramic dental X-ray machines marks a significant technological advancement. This integration facilitates the automation of teeth counting and segmentation, streamlining the workflow in dental practices and enhancing the efficiency of diagnostic processes.

Furthermore, as demonstrated in the previous section, the importance of development of an automated framework for accurate identification of individual teeth and tooth counting is crucial for treatment planning. In this study, we not only analyzed multiple image segmentation models using comprehensive metrics but also developed a low-cost framework utilizing the U-Net model, for segmenting individual teeth and counting the total teeth in the panoramic X-ray as shown in the Fig 7.

5 Conclusions

In conclusion, this paper provides a comprehensive evaluation of image segmentation models for dental X-ray analysis, focusing on panoramic images of varying sizes (256x256 and 512x512). Through experimentation and analysis, we assessed the performance of several state-of-the-art segmentation models, including U-Net, Attention U-Net, Feedback U-Net, and Feedback U-Net+LSTM. Our findings consistently demonstrate that the U-Net model outperforms other models in dentistry, where data is often limited. This superiority is mainly due to U-Net's capability to be trained effectively with small datasets, unlike other models designed for larger datasets. Expanding the scope of our study, we developed a framework leveraging the capabilities of the U-Net model to identify individual teeth and return the total count of teeth. This framework can be integrated with panoramic dental X-ray machines, offering a low-cost solution for automated teeth segmentation using panoramic X-ray images. This integration assists dental experts in accurately planning patient treatments, ultimately improving patient care and outcomes. Future research could focus on integrating these segmentation models with real-time imaging systems to enhance diagnostic accuracy in dentistry. Further exploration into specialized areas such as endodontics and orthodontics could expand the applicability of these technologies.

Acknowledgement

High-performance computing resources were supported by the Economic and Social Research Council (ESRC) funded Business and Local Government Data Research Centre under Grant ES/S007156/1.

References

1. F. Schwendicke, A. Chaurasia, T. Wiegand, S. E. Uribe, M. Fontana, I. Akota, O. Tryfonos, J. Krois *et al.*, "Artificial intelligence for oral and dental healthcare: core education curriculum," *Journal of Dentistry*, vol. 128, p. 104363, 2023.
2. C. Huang, J. Wang, S. Wang, and Y. Zhang, "A review of deep learning in dentistry," *Neurocomputing*, p. 126629, 2023.
3. F. P. Mahdi, K. Motoki, and S. Kobashi, "Optimization technique combined with deep learning method for teeth recognition in dental panoramic radiographs," *Scientific Reports*, vol. 10, no. 1, p. 19261, 2020.
4. S. Risnes, Q. Khan, E. Hadler-Olsen, and A. Sehic, "Tooth identification puzzle: a method of teaching and learning tooth morphology," *European Journal of Dental Education*, vol. 23, no. 1, pp. 62–67, 2019.
5. L. Wang, J. Mao, Y. Hu, and W. Sheng, "Tooth identification based on teeth structure feature," *Systems Science & Control Engineering*, vol. 8, no. 1, pp. 521–533, 2020.
6. A. Jan, R. Albenayan, D. Alsharkawi, and F. Jadu, "The prevalence and causes of wrong tooth extraction," *Nigerian Journal of Clinical Practice*, vol. 22, no. 12, pp. 1706–1714, 2019.
7. Y. Miki, C. Muramatsu, T. Hayashi, X. Zhou, T. Hara, A. Katsumata, and H. Fujita, "Classification of teeth in cone-beam ct using deep convolutional neural network," *Computers in Biology and Medicine*, vol. 80, pp. 24–29, 2017.

8. A. Sonavane and R. Kohar, "Dental cavity detection using yolo," in *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 2*. Springer, 2022, pp. 141–152.
9. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 3431–3440.
10. W. Sun and R. Wang, "Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with dsm," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 3, pp. 474–478, 2018.
11. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer, 2015, pp. 234–241.
12. Z. Chen, S. Chen, and F. Hu, "Cta-unet: Cnn-transformer architecture unet for dental cbct images segmentation," *Physics in Medicine & Biology*, vol. 68, no. 17, p. 175042, 2023.
13. K. Panetta, R. Rajendran, A. Ramesh, S. P. Rao, and S. Agaian, "Tufts dental database: a multimodal panoramic x-ray dataset for benchmarking diagnostic systems," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 4, pp. 1650–1659, 2021.
14. X. Xu, C. Liu, and Y. Zheng, "3d tooth segmentation and labeling using deep convolutional neural networks," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 7, pp. 2336–2348, 2018.
15. C. Fares and M. Feghali, "Tooth-based identification of individuals," *International Journal of New Computer Architectures and their Applications (IJNCAA)*, vol. 3, no. 1, pp. 22–34, 2013.
16. M. Maddalone and M. Gagliani, "Periapical endodontic surgery: a 3-year follow-up study." *International Endodontic Journal*, vol. 36, no. 3, 2003.
17. R. Izzetti, M. Nisi, G. Aringhieri, L. Crocetti, F. Graziani, and C. Nardi, "Basic knowledge and new advances in panoramic radiography imaging techniques: A narrative review on what dentists and radiologists should know," *Applied Sciences*, vol. 11, no. 17, p. 7858, Aug 2021. [Online]. Available: http://dx.doi.org/10.3390/app11177858
18. D. D. Patil and S. G. Deore, "Medical image segmentation: a review," *International Journal of Computer Science and Mobile Computing*, vol. 2, no. 1, pp. 22–27, 2013.
19. Z. Rezaei, "A review on image-based approaches for breast cancer detection, segmentation, and classification," *Expert Systems with Applications*, vol. 182, p. 115204, 2021.
20. H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, and J. Wu, "Unet 3+: A full-scale connected unet for medical image segmentation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1055–1059.
21. F. Oztekin, O. Katar, F. Sadak, M. Aydogan, T. T. Yildirim, P. Plawiak, O. Yildirim, M. Talo, and M. Karabatak, "Automatic semantic segmentation for dental restorations in panoramic radiography images using u-net model," *International Journal of Imaging Systems and Technology*, vol. 32, no. 6, pp. 1990–2001, 2022.
22. A. Sakuma, H. Saitoh, Y. Makino, G. Inokuchi, M. Hayakawa, D. Yajima, and H. Iwase, "Three-dimensional visualization of composite fillings for dental identification using ct images," *Dentomaxillofacial Radiology*, vol. 41, no. 6, pp. 515–519, 2012.

23. M. C. KAYA, "Dental panoramic and bitewing x-ray image segmentation using u-net and transformer networks," Master's thesis, Middle East Technical University, 2023.
24. N. K. Tomar, D. Jha, M. A. Riegler, H. D. Johansen, D. Johansen, J. Rittscher, P. Halvorsen, and S. Ali, "Fanet: A feedback attention network for improved biomedical image segmentation," *IEEE Transactions on Neural Networks and Learning Systems*, 2022.
25. L. Yuan, J. Song, and Y. Fan, "Fm-unet: Biomedical image segmentation based on feedback mechanism unet," *Mathematical Biosciences and Engineering*, vol. 20, no. 7, pp. 12 039–12 055, 2023.
26. E. Shibuya and K. Hotta, "Feedback u-net for cell image segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 974–975.
27. A. E. Yüksel, S. Gültekin, E. Simsar, Ş. D. Özdemir, M. Gündoğar, S. B. Tokgöz, and İ. E. Hamamcı, "Dental enumeration and multiple treatment detection on panoramic x-rays using deep learning," *Scientific Reports*, vol. 11, no. 1, p. 12342, 2021.
28. S. HELLİ and A. HAMAMCI, "Tooth instance segmentation on panoramic dental radiographs using u-nets and morphological processing," *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, vol. 10, no. 1, pp. 39–50.
29. H. Li, G. Sun, H. Sun, and W. Liu, "Watershed algorithm based on morphology for dental x-ray images segmentation," in *IEEE 11th International Conference on Signal Processing*, vol. 2, 2012, pp. 877–880.