Segmentation of Tooth Crack on Dental CT-Scan Images

[1,*] *Feng Lin*（林鋒）, [2,*]*Y. T. Augustine Tsai*（蔡岳廷）, [3]*Chih-Chia Huang*(黃智嘉), [1,*] *Chiou-Shann Fuh*（傅楸善）,

[1]Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan,
[2]Institute for Information Industry,
[3]Department of endodontics, Cardinal Tien Hospital, Ankang. New Taipei City, R.O.C. Taiwan

[*]E-mail: r12922197@ntu.edu.tw atsai@iii.org.tw chihchiahuang0216@gmail.com fuh@csie.ntu.edu.tw

ABSTRACT

This paper presents a comprehensive method for tooth crack segmentation on dental Micro Computed Tomography (Micro CT)-scan images, adapting a Convolutional-Transformer Network for Crack Segmentation with Boundary Awareness proposed by Tao et al. [1]. Our approach encompasses data preparation, model modification, training, program development, and crack location extraction. Our model was trained and evaluated on a dataset of 850 annotated Micro CT images. The results demonstrate the model's potential to significantly enhance diagnostic accuracy and efficiency in clinical settings, providing a valuable tool for dental professionals. Future work will focus on improving the model's robustness and real-time integration capabilities.

Keywords: *Tooth Crack Segmentation, Dental CT-scan, Vertical Root Fractures, Convolutional-Transformer Network, Boundary Awareness, Model Training, Image Processing, IPPR, CVGIP 2024.*

1. INTRODUCTION

Vertical root fractures (VRFs) represent a critical and often challenging aspect of dental pathology, characterized by longitudinal cracks that extend along the length of a tooth's root. These fractures are significant not only due to their potential to compromise the structural integrity of the tooth but also because of the complex diagnostic and treatment challenges they pose. VRFs can arise from various etiological factors, including traumatic injuries, excessive forces during endodontic procedures, and prolonged occlusal stress. Despite their relatively low prevalence, the impact of VRFs on dental health is profound, often necessitating tooth extraction due to the difficulty in achieving successful restorative outcomes [2][3].

The clinical presentation of VRFs is often insidious, with symptoms such as pain on biting, sensitivity to thermal changes, and localized swelling being easily mistaken for other dental conditions. This symptomatology, coupled with the subtle radiographic signs, makes early diagnosis of VRFs particularly challenging. In addition to removing the affected tooth, a treatment method using mineral trioxide aggregate (MTA), a dental repair material, to fill the cracks achieved a 94.3% survival rate in 52 patients [4]. In addition, there are many new treatment methods for VRFs [5][6][7], which also require accurate identification of their location. Therefore, identifying the location of VRFs is crucial. Advanced imaging techniques, such as Cone Beam Computed Tomography (CBCT), have therefore become invaluable tools in the identification and management of these fractures [8].

Tooth crack segmentation is essential for precise dental diagnostics and treatment planning. Automatic crack detection on dental CT-scan images can significantly streamline this process, enabling clinicians to make informed decisions. Dental professionals often rely on visual inspection and manual examination to identify cracks, which can be time-consuming and subject to human error. This challenge underscores the need for automated systems that can enhance diagnostic accuracy and efficiency.

Recent advancements in deep learning and medical imaging have paved the way for sophisticated models capable of handling complex segmentation tasks. Convolutional Neural Networks (CNNs) and Transformer architectures have emerged as powerful tools in this domain, offering the potential to significantly improve the accuracy and reliability of image-based diagnostics. In this context, our work focuses on adapting

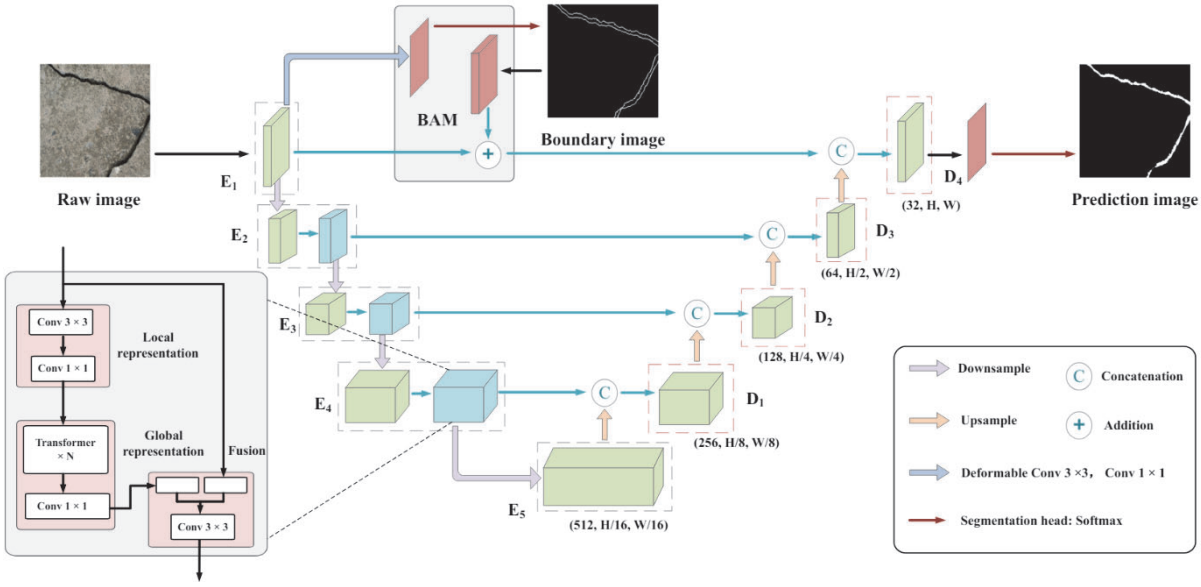


Fig. 1. An overview of the proposed model. E and D denote encoder and decoder respectively. Green feature maps are generated via Dilated Residual Block while blue feature maps are generated via MobileViT Block. From Tao, H., Liu, B., Cui, J., & Zhang, H. (2023). "A Convolutional-Transformer Network for Crack Segmentation with Boundary Awareness. "In 2023 IEEE International Conference on Image Processing (ICIP) (pp. 86-90). IEEE. [1]

a convolutional-transformer network crack segmentation model for application in dental imaging.

The specific challenges associated with tooth crack segmentation include the subtlety of crack features, variations in tooth anatomy, and the presence of noise and artifacts in CT images. These factors necessitate a robust and sensitive approach capable of distinguishing fine details and accurately delineating crack boundaries.

In recent years, various approaches have been proposed for crack detection and segmentation in medical and non-medical images. For instance, Yang et al. (2019) introduced a Feature Pyramid and Hierarchical Boosting Network for pavement crack detection, demonstrating the effectiveness of hierarchical structures in capturing multi-scale features [9]. Guo et al. (2022) provided a comprehensive review of imaging modalities and the evolving role of AI in the diagnosis of cracked teeth, highlighting the importance of integrating AI-based analysis into clinical workflows [10]. Liu et al. (2019) and Zou et al. (2018) both propose a deep convolutional neural network for automatic crack detection [11][12].

In this paper, we adapt the Convolutional-Transformer Network with Boundary Awareness proposed by Tao et al. [1] to the domain of dental CT-scan images. This model integrates multiple advanced components—Dilated Residual Block (DRB), Boundary Awareness Module (BAM), and MobileViT Block—to effectively capture both local and global features essential for precise crack segmentation. By leveraging these components, our approach aims to address the specific challenges posed by dental crack detection and provide a reliable tool for clinical use.

In summary, this paper presents a method for tooth crack segmentation on dental CT-scan images, utilizing an adapted Convolutional-Transformer Network. We detail the steps involved in each stage of the process. Through rigorous evaluation and performance analysis, we demonstrate the effectiveness of our approach and highlight its potential benefits for dental diagnostics and treatment planning.

2. METHOD

Our methodology consists of several key steps:

2.1. Data Preparation

We collected a dataset comprising dental Micro CT images with corresponding crack location annotations. First, we annotated the locations of cracks on the Micro CT images. Then, the annotated parts were extracted to form the ground truth binary masks, indicating the presence or absence of cracks at each pixel location. The dataset used includes 850 Micro CT images, which were partitioned into three sets: training, validation, and test sets (595, 85, 170).

2.2. Model Adaptation for Tooth Crack Segmentation

The original model by Tao et al. [1] was used to address the specific challenges of tooth crack segmentation. This model leverages a Convolutional-Transformer Network with three key components: Dilated Residual Block (DRB), Boundary Awareness Module (BAM), and

MobileViT Block. Figure 1 provides an overview of the proposed model.
Dilated Residual Block (DRB): The DRB captures local details of cracks through dilated convolutions, enhancing segmentation accuracy for fine-grained features typical of dental cracks.
Boundary Awareness Module (BAM): The BAM focuses on learning boundary features from dilated crack labels, which is particularly useful in dental images where the contrast between the crack and the surrounding tooth material may be subtle.
MobileViT Block: The MobileViT Block encodes global contextual information with fewer parameters, aiding in accurate segmentation across varying dental images.

2.3.Model Training

The model was trained using the prepared dataset with several adjustments to enhance performance:
Training Parameters: The model was trained with a learning rate of 0.001, a batch size of 2, and for a total of 100 epochs. We utilized the Adam optimizer, and early stopping was implemented based on the validation loss to prevent overfitting.
Data Augmentation: To improve the model's robustness, data augmentation techniques such as rotation, flipping, scaling, and contrast adjustment were applied to the training data. This helped the model generalize better to unseen data.
Loss Function: To address the class imbalance in the dataset, we employed a combination of weighted binary cross-entropy and dice loss. The weights for the binary cross-entropy were adjusted to emphasize the minority class (crack pixels), ensuring the model focuses adequately on crack features.

2.4. Crack Location Extraction

The trained model was used to predict crack locations on unseen dental CT-scan images. We evaluated the model's performance on a separate test set by compared the predictions with ground truth annotations. Post-processing techniques, such as morphological operations, were applied to refine the segmentation output.

2.5. Experimental Setup

Datasets: The model was trained and evaluated on a dataset of annotated dental images, including 850 CT images of cracked teeth.
Training Details: The network was trained using the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 2 over 100 epochs. Early stopping and model check-pointing were employed to prevent overfitting
Evaluation Metrics: Precision, recall, and F1-score were used to evaluate the model's performance.

3. RESULTS

Figure 2 illustrates sample predictions generated by the model. The model outputs a value between 0 and 1 for each pixel, representing the likelihood of that pixel being a crack. The model prediction in Figure 2 displays the model's output as grayscale values ranging from 0 to 255, corresponding to the range [0,1]. For the model's output, we can choose different thresholds to serve as the criteria for identifying cracks.
Figure 3 illustrates the variation in predicted results when using different thresholds (ranging from left to right: 0, 0.4, 0.9).
With a low threshold, the model's predictions cover most of the crack locations. However, since the predicted crack width is broader than the ground truth, and some non-crack part are included, the precision performance is suboptimal.
Figure 4 illustrates the comparison between model predictions and ground truth under three different thresholds (0, 0.4, 0.9). The highlighted areas are true positive (white), false positives (red), false negatives (blue), and true negative (black).
To assess the efficacy of our tooth crack segmentation model, we conducted a rigorous performance evaluation using various metrics at different threshold levels. Table 1 summarizes the precision, recall, and F1 score achieved at different threshold values, which illustrates model performance under different thresholds.
The precision, recall, and F1 score serve as crucial indicators of the model's ability to accurately identify tooth cracks across varying thresholds. Precision measures the proportion of true positive predictions among all positive predictions, indicating the model's ability to minimize false positives. Recall, on the other hand, calculates the proportion of true positives identified correctly, indicating the model's capacity to capture all actual positives. Lastly, the F1 score, the harmonic mean of precision and recall, provides a balanced measure that considers both false positives and false negatives.

Table 1. Prediction scores with different thresholds.

Threshold	Precision↑	Recall↑	F1 Score↑
0.1	0.5956	**0.9234**	0.7241
0.2	0.6492	0.8942	0.7523
0.3	0.6925	0.8558	0.7655
0.4	0.7231	0.8212	**0.769**
0.5	0.7492	0.7688	0.7589
0.6	0.7676	0.6988	0.7316
0.7	0.7846	0.6488	0.7102
0.8	0.8038	0.5835	0.6762
0.9	**0.8115**	0.4856	0.6076

At lower thresholds (0.1 to 0.3), the model exhibited relatively high recall values, indicating its proficiency in identifying a substantial proportion of true tooth crack pixels. However, precision was comparatively lower at

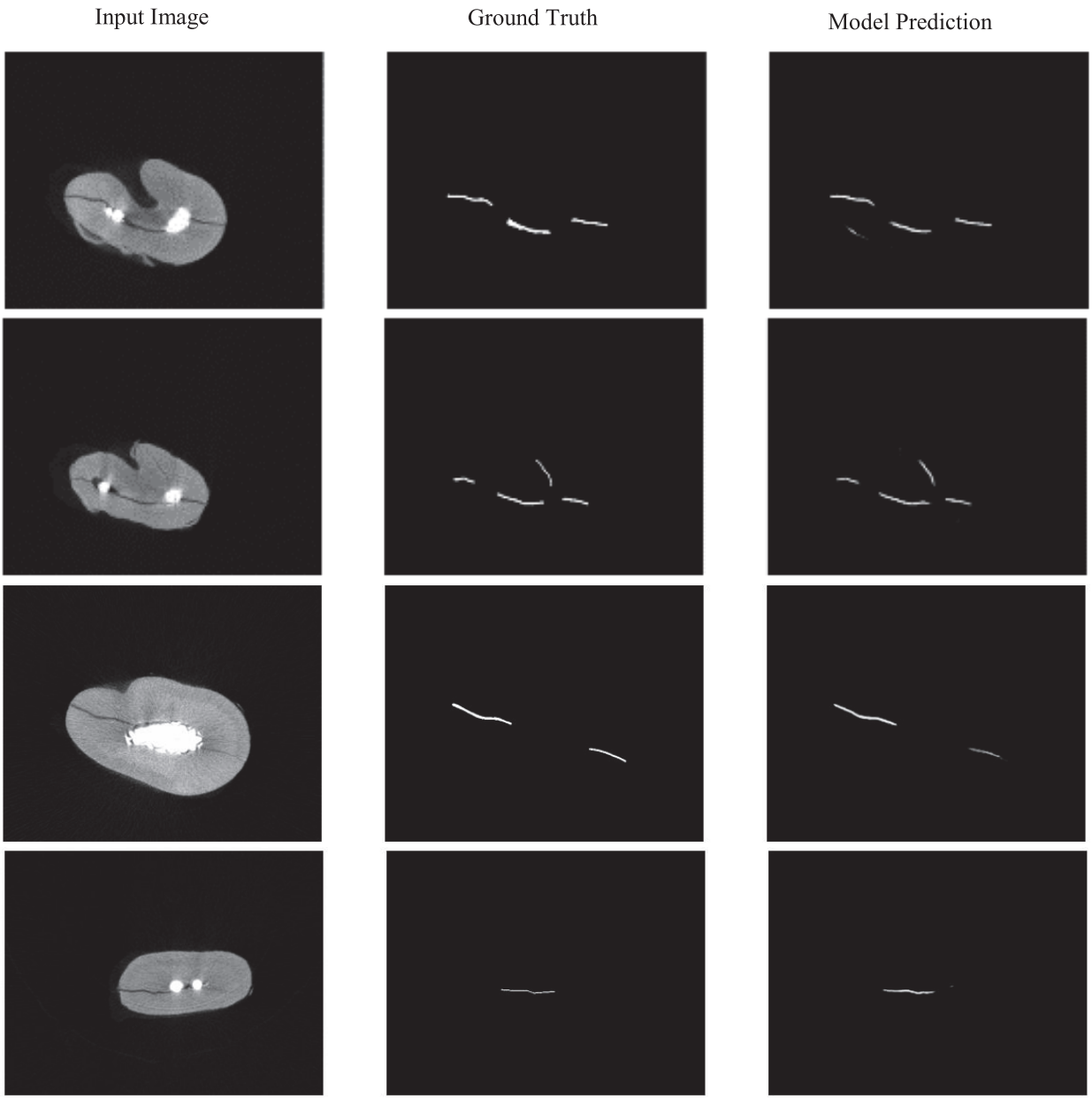


Fig. 2. Examples of predictions generated by the model.

Fig. 3. Model predictions with different thresholds. When the predicted value of a pixel by the model exceeds the threshold, that pixel is marked as a crack, represented by white in the image; otherwise, it is marked as black.

Threshold = 0 Threshold = 0.4 Threshold = 0.9

Fig. 4. Comparison between model predictions and ground truth under different threshold. White pixels represent correct predictions of cracks (true positives), red pixels represent non-cracks incorrectly identified as cracks (false positives), blue pixels represent cracks incorrectly identified as non-cracks (false negatives), and black pixels represent correct predictions of non-cracks (true negatives).

these thresholds, suggesting a higher rate of false positives.

As the threshold increases (0.4 to 0.9), precision improved while recall decreased. This trend signified a trade-off between precision and recall, with higher thresholds leading to reduced false positives but potentially missing some true tooth crack pixels. Reference to Figure 4, it effectively demonstrates this trend.

Figure 5 is Precision-Recall Curve. Each point represents the Recall and Precision obtained by setting different threshold values. Which give a better view of relationship between precision and recall.

Recall and Precision are both suitable for use in datasets with class imbalance because their formulas do not consider True Negatives. They focus solely on correctly identifying positive samples. Therefore, even if the number of negative samples in the dataset is much larger than the number of positive samples, Recall and Precision remain effective metrics. In our samples, cracks usually only occupy a small portion of the entire image.

The F1 score, representing the harmonic mean of precision and recall, peaked around the 0.4 threshold, indicating an optimal balance between precision and recall at that point.

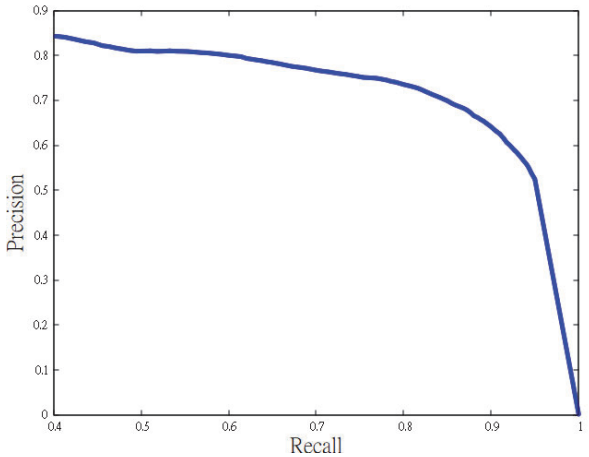


Fig. 5. Precision-Recall Curve.

Another metric for evaluating model performance is Intersection over Union (IoU). IoU is a key metric used to evaluate the accuracy of image segmentation models. It measures the overlap between the predicted segmentation and the ground truth segmentation. IoU is calculated as the ratio of the area of overlap to the area of union between the predicted and ground truth regions. An IoU of 1 indicates a perfect match, while an IoU of 0 indicates no overlap. IoU is widely used in tasks like object detection and semantic segmentation to assess model performance. Higher IoU values signify better segmentation accuracy, making it an essential metric for comparing and improving models.

Figure 6 shows the IoU curve of our model. In the Figure, it can be observed that the model performs best around threshold = 0.4 based on the IoU metric. This aligns with the results obtained when evaluating using the F1 score as the metric.

Overall, these results highlight the model's ability to segment tooth cracks across different threshold levels, offering insights into the trade-offs between precision and recall inherent in the segmentation process.

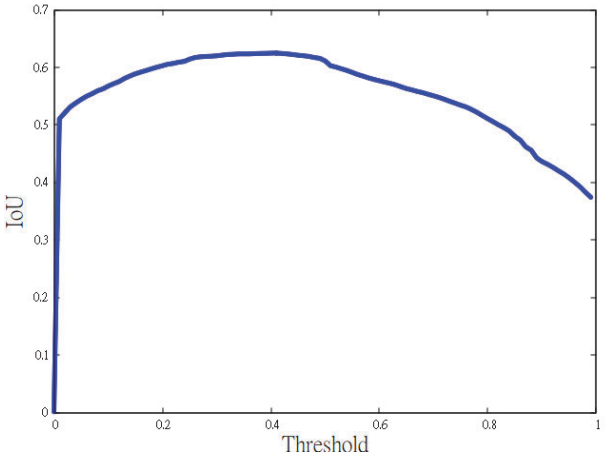


Fig. 6. Intersection over Union (IoU) curve.

In the primary application scenario, Diagnosis and treatment of VRFs, dentists need to identify the widest part of the crack to inject filling material. Therefore, we believe a threshold of around 0.4 is more suitable. Even though the precision is not optimal, it covers most of the cracks and provides a relative width for reference.

The Micro CT images we used were taken from extracted tooth samples, allowing the use of higher radiation doses to achieve higher resolution and detect cracks of 5–20 μm. In contrast, CBCT taken directly from patients can only detect cracks of 50–300 μm. Figure 7 shows a comparison of the two. Therefore, more testing and improvement are needed in current practical applications.

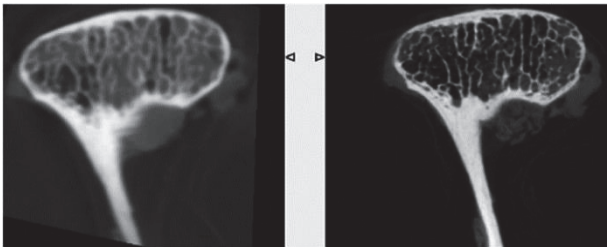
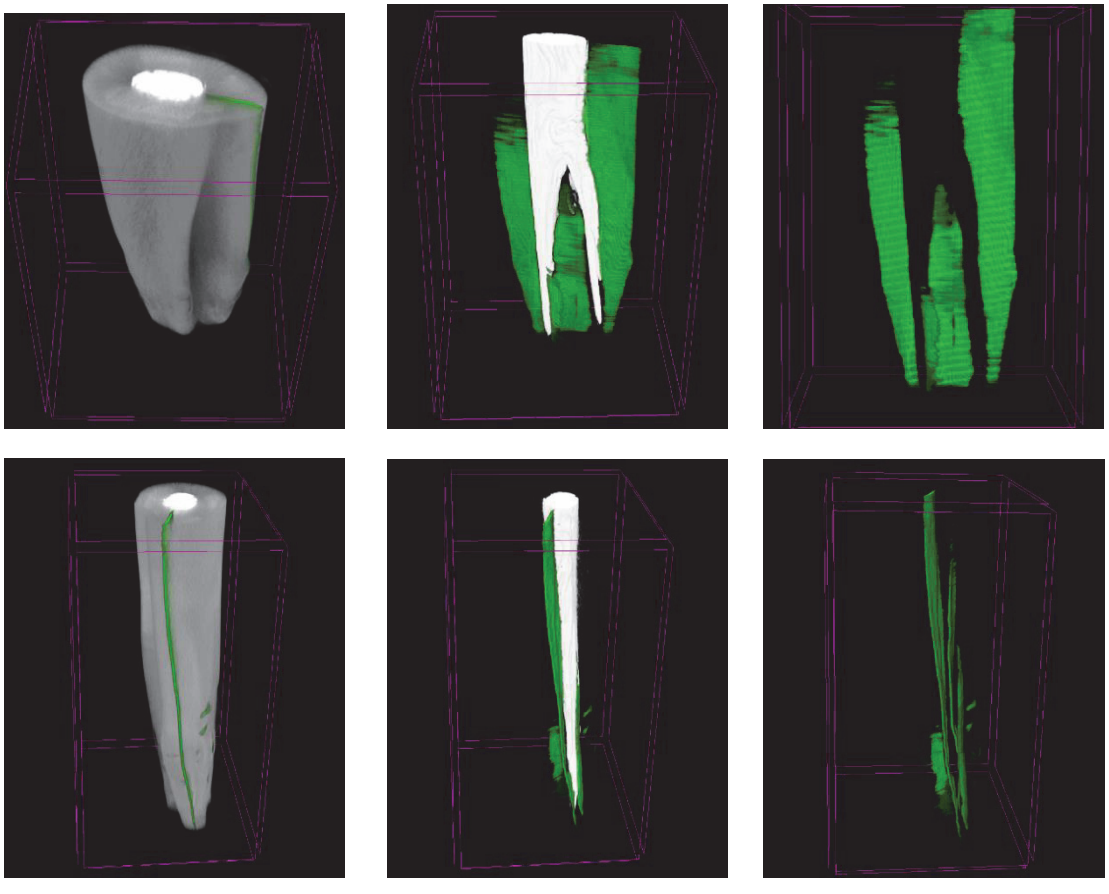


Fig. 7. Sample images from a registered micro-CT/CBCT pair. Micro-CT image (right) vs CBCT image (left) of 0.5 mm slice thickness. From Liang X, Zhang Z, Gu J, Wang Z, Vandenberghe B, Jacobs R, Yang J, Ma G, Ling H, Ma X. “Comparison of micro-CT and cone beam CT on the feasibility of assessing trabecular structures in mandibular condyle.” Dentomaxillofac Radiol. 2017 Jul;46(5):20160435. doi: 10.1259/dmfr.20160435. Epub 2017 Apr 26. PMID: 28350523; PMCID: PMC5595038. [13].

In the ideal diagnostic workflow of VRFs, after capturing the CT images of the patient's teeth and obtaining predictions from the model, we can acquire a series of locations indicating tooth cracks. By integrating these with the original CT images, visualization can be facilitated using software like CTvox. First, load the CT image of the tooth and the predicted crack location by the model into CTvox. Then, adjust the colors to distinguish between the two, and finally, adjust the opacity of the tooth image to observe the interior. We can also individually observe the morphology of the cracks, as demonstrated in Figure 8. This aids dentists in diagnosis and decision-making regarding treatment plans.



(a) (b) (c)

Fig. 8. An example of combining Micro CT images with model-predicted results in CTvox. The part indicating cracks is highlighted in green. (a) is the combined image of both. (b) is the internal structure seen after adjusting the opacity transfer function of (a). The white parts represent the filling material within the tooth pulp chamber. (c) is the structure of cracks.

4. CONCLUSION

In conclusion, we have developed and evaluated a Convolutional-Transformer Network for the segmentation of tooth cracks in dental Micro CT-scan images. The model shows significant potential in aiding dental professionals by providing accurate and reliable crack segmentation.

The developed model has several potential applications in dental practice, including assisting in the diagnosis of other dental conditions that require detailed imaging analysis.

Image-based intelligent auxiliary diagnosis could emerge as a leading approach in clinical applications. While image processing is not extensively employed in diagnosing cracked teeth presently, there's promising potential for AI-driven diagnostic methods to play a significant role in dental diagnosis. Unlike conventional imaging methods rely on human effort, AI offers safer, more precise, and efficient imaging solutions. However, several technical issues may still exist in AI-based detection, such as heavy computational cost, issues of selecting optimal parameters, and pre-processing of the training datasets.

Future research can focus on enhancing the model's robustness and generalizability. This includes training on larger and more diverse datasets to capture a wider range of dental anomalies. We once trained a model using approximately 300 Micro CT images, and its performance was noticeably inferior to the results obtained from training with the current dataset of 850 images. Therefore, increasing the training data is also a crucial aspect. Additionally, exploring advanced post-processing techniques, such as refining segmentation outputs with morphological operations, might further improve accuracy. Integrating real-time processing capabilities will also be a priority to facilitate seamless clinical integration.

We hope that this study will inspire further research on crack segmentation in tooth CT images and other medical application.

5. REFERENCES

[1] Tao, H., Liu, B., Cui, J., & Zhang, H. (2023). "A Convolutional-Transformer Network for Crack Segmentation with Boundary Awareness. "In 2023 IEEE International Conference on Image Processing (ICIP) (pp. 86-90). IEEE.

[2] Cameron CE. "The cracked tooth syndrome: additional findings." J Am Dent Assoc. 1976;93(5):971–5.

[3] Walton RE, Michelich RJ, Smith GN. "The histopathogenesis of vertical root fractures." J Endod. 1984 Feb;10(2):48-56. doi: 10.1016/S0099-2399(84)80037-0. PMID: 6586962.

[4] Tani-Ishii, Nobuyuki & Mutoh, Noriko & Muromachi, Koichiro & Suzuki, Jiro. (2017). "The clinical evaluation of vertical root fracture after endodontic treatment with mineral trioxide aggregate." Integrative Molecular Medicine. 4. 10.15761/IMM.1000288.

[5] Sugaya, Tsutomu & Natatsuka, Megumi & Motoki, Youji & Inoue, Kana & Tanaka, Saori & Miyaji, Hirofumi & Kawanami, Masamitsu & Sakagami, Ryuji. (2016). "Sealing the Gap of Vertical Root Fracture through the Root Canal." Dentistry. 06. 10.4172/2161-1122.1000354.

[6] Hadrossek PH, Dammaschke T. "New treatment option for an incomplete vertical root fracture--a preliminary case report." Head Face Med. 2014 Mar 26;10:9. doi: 10.1186/1746-160X-10-9. PMID: 24670232; PMCID: PMC3986878.

[7] Okaguchi M, Kuo T, Ho YC. "Successful treatment of vertical root fracture through intentional replantation and root fragment bonding with 4-META/MMA-TBB resin." J Formos Med Assoc. 2019 Mar;118(3):671-678. doi: 10.1016/j.jfma.2018.08.004. Epub 2018 Aug 23. PMID: 30145002.

[8] Kamburoğlu K, Murat S, Yüksel SP, Cebeci AR, Horasan S. Detection of vertical root fracture using cone-beam computerized tomography: an in vitro assessment. Oral Surg Oral Med Oral Pathol Oral Radiol Endod. 2010 Feb;109(2):e74-81. doi: 10.1016/j.tripleo.2009.09.005. PMID: 20031454.

[9] Yang, F., Zhang, L., Yu, S., Prokhorov, D., Mei, X., & Ling, H. (2019). "Feature pyramid and hierarchical boosting network for pavement crack detection." IEEE Transactions on Intelligent Transportation Systems, 21(4), 1525-1535.

[10] Guo J, Wu Y, Chen L, Long S, Chen D, Ouyang H, Zhang C, Tang Y, Wang W. "A perspective on the diagnosis of cracked tooth: imaging modalities evolve to AI-based analysis." Biomed Eng Online. 2022 Jun 15;21(1):36. doi: 10.1186/s12938-022-01008-4. PMID: 35706023; PMCID: PMC9202175.

[11] Yahui Liu, Jian Yao, Xiaohu Lu, Renping Xie, and Li Li, "Deepcrack: A deep hierarchical feature learning architecture for crack segmentation," Neurocomputing, vol. 338, pp. 139–153, 2019.

[12] Qin Zou, Zheng Zhang, Qingquan Li, Xianbiao Qi, Qian Wang, and Song Wang, "Deepcrack: Learning hierarchical convolutional features for crack detection," IEEE Transactions on Image Processing (T-IP), vol. 28, no. 3, pp. 1498–1512, 2018.

[13] Liang X, Zhang Z, Gu J, Wang Z, Vandenberghe B, Jacobs R, Yang J, Ma G, Ling H, Ma X. “Comparison of micro-CT and cone beam CT on the feasibility of assessing trabecular structures in mandibular condyle.” Dentomaxillofac Radiol. 2017 Jul;46(5):20160435. doi: 10.1259/dmfr.20160435. Epub 2017 Apr 26. PMID: 28350523; PMCID: PMC5595038.