PRAD: Periapical Radiograph Analysis Dataset and Benchmark Model Development

Zhenhuan Zhou[1] *, Yuchen Zhang[2] *, Ruihong Xu[1], Xuansen Zhao[3], and Tao Li[1,4](✉)

[1] College of Computer Science, Nankai University, Tianjin, China
[2] Department of stomatology, Tianjin Union Medical Center, The First Affiliated Hospital of Nankai University, Tianjin, China
[3] College of Mechanical Engineering, Taiyuan University of Technology, Shanxi, China
[4] Haihe Lab of ITAI, Tianjin, China
litao@nankai.edu.cn

Abstract. Deep learning (DL), a pivotal technology in artificial intelligence, has recently gained substantial traction in the domain of dental auxiliary diagnosis. However, its application has predominantly been confined to imaging modalities such as panoramic radiographs and Cone Beam Computed Tomography, with limited focus on auxiliary analysis specifically targeting Periapical Radiographs (PR). PR are the most extensively utilized imaging modality in endodontics and periodontics due to their capability to capture detailed local lesions at a low cost. Nevertheless, challenges such as projection angle and artifacts complicate the annotation and recognition of PR, leading to a scarcity of publicly available, large-scale, high-quality PR analysis datasets. This scarcity has somewhat impeded the advancement of DL applications in PR analysis. In this paper, we present PRAD-10K, a dataset for PR analysis. PRAD-10K comprises 10,000 clinical periapical radiograph images, with pixel-level annotations provided by professional endodontists for nine distinct anatomical structures, lesions, and artificial restorations or medical devices. We also include classification labels for images with typical conditions or lesions. Furthermore, we introduce a DL network named PRNet to establish benchmarks for PR segmentation tasks. Experimental results demonstrate that PRNet surpasses previous state-of-the-art medical image segmentation models on the PRAD-10K dataset. The code and dataset will be released at https://github.com/nkicsl/PRAD.

Keywords: Periapical radiographs · Dental dataset · Medical image segmentation

1 Introduction

Radiological imaging is vital in dentistry for diagnosis, treatment planning, and outcome evaluation [15]. As shown in Fig. 1, the main modalities include

* Zhenhuan Zhou and Yuchen Zhang contribute equally.

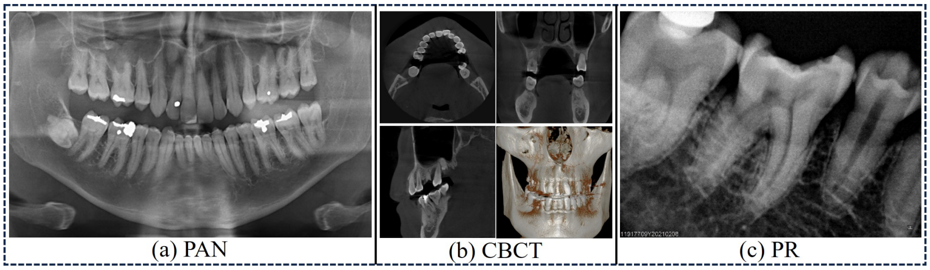

Fig. 1. Examples of three dental radiographic imaging modalities. (a) PAN, (b) CBCT and (c) PR.

Panoramic Radiography (PAN), Cone Beam Computed Tomography (CBCT), and Periapical Radiographys (PR). PAN offers a broad dental arch view but lacks detail for individual teeth [26]. CBCT provides high-resolution 3D images but is less common in routine endodontics due to cost and radiation [23, 26]. As a common imaging modality in endodontics [22], PR mitigate these challenges by providing high-resolution intraoral images of one or two teeth, aiding in the diagnosis of apical periodontitis and root canal treatment planning [24]. While PR offer detailed views of local structures, its limited field of view and sensitivity to angulation and anatomical overlap lead to interpretative difficulties and inter-observer variability. Thus, applying Deep Learning (DL) to PR analysis has the potential to enhance treatment outcomes, reduce subjectivity, and support computer-aided diagnosis and treatment in endodontics.

Recently, there has been a surge of research and applications of DL in the field of dentistry. For instance, Wang et al. [30] introduced a deep multi-task learning framework designed to segment teeth and root canals in dental CBCT images. Jang et al. [13] developed a fully automated system for integrating intraoral scans (IOS) with CBCT through individual tooth segmentation and identification. Wang et al. [28] employ weakly supervised tooth instance segmentation with multi-label learning, facilitating accurate segmentation of teeth in dental 3D models. Regarding PR image analysis, several studies have applied DL methods. For example, previous works have employed UNet [21] based DL approaches to analyze PR images [2, 9, 14], enabling effective feature extraction and improving the accuracy of identifying anatomical structures and lesions. Chen et al. [6] and Ozsari et al. [17] applied DL techniques for the early diagnosis of periodontal bone loss and the detection of vertical root fractures. However, most of these studies were conducted using small, private datasets, and publicly available, high-quality benchmark datasets for PR image analysis remain scarce.

In this paper, to address the above issues, we introduce PRAD-10K, a large-scale PR dataset featuring expert annotations and serving as a potential benchmark for research in DL-based PR image analysis. Additionally, to tackle the multi-scale challenges inherent in PR image segmentation tasks, we present PR-Net, a DL network that integrates the Multi-scale Wavelet Convolution Network

Table 1. Comparison between PRAD-10K and some publicly available 2D dental datasets, OR and IR stands for Occlusal and Intraoral Radiograph, respectively.

Dataset	Modality	Year	Size	Task
PDSM [1]	PAN	2020	116	Mandibles segmentation
TDD [18]	PAN	2021	1000	Tooth segmentation
PRD [20]	PAN	2021	598	Tooth segmentation
OCI [3]	RGB Photo	2022	131	Oral cancer classification
DC1000 [29]	PAN & OR	2023	597+2389	Caries segmentation
Thalji et al. [25]	PR	2023	929	Classification
IO150K [33]	IR	2024	1,500,000	Tooth segmentation
PRAD-10K (Ours)	PR	2025	10,000	Multi-structure segmentation Disease classification

(MWCN) and the Channel Fusion Attention (CFA) mechanism. Extensive experiments demonstrated that PRNet surpasses previous state-of-the-art (SOTA) medical image segmentation networks on the PRAD-10K dataset. Ablation experiments also confirmed the effectiveness of each component of PRNet.

2 PRAD-10K

2.1 Overview

As shown in Table 1, PRAD-10K is compared with existing 2D dental datasets. The introduction of PRAD-10K provides a benchmark for DL-based PR image analysis. Fig. 2 presents four PRAD-10K images with expert pixel-level annotations covering all annotated categories, alongside a detailed label index. PRAD-10K includes nine pixel-level annotation types for anatomical structures, lesions, restorative materials and medical devices. Additionally, classification labels are provided for images with periodontitis, apical periodontitis, or inadequate root canal fillings.

2.2 Collection and Annotation

The PRAD-10K dataset is sourced from real clinical data from the Department of Endodontics of The First Affiliated Hospital of Nankai University, ensuring that no personally identifiable information was included, and the data were used exclusively for research purposes. The study was approved by the hospital ethics committee (approval number 2025-B114). A radiology expert selected 10,000 high-quality PR images based on criteria such as clear tooth visibility, no severe artifacts, and excluding master cone or working length radiographs. Images featuring rich structures such as implants, orthodontic appliances, or dental restorations were also intentionally included to enhance dataset diversity.

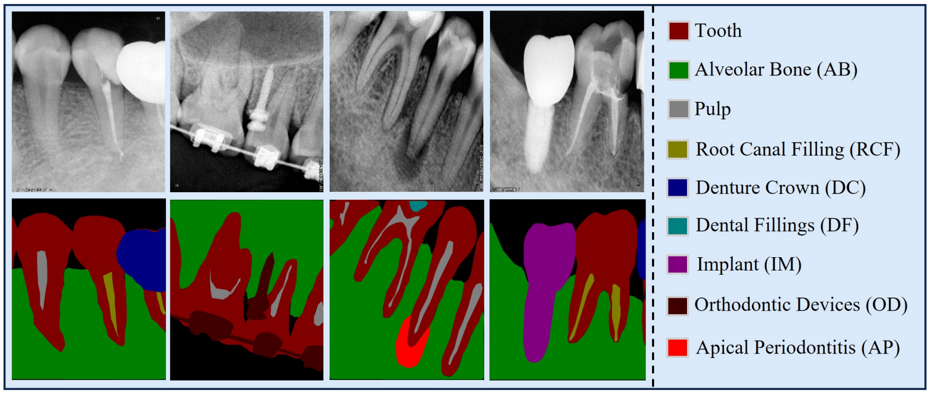

Fig. 2. Examples from the PRAD-10K dataset (top left) along with their corresponding pixel-level annotations (bottom left). The right side displays the label distribution.

The data annotation was conducted by two endodontists with over eight years of experience and one computer science researcher. The dataset was divided between the two endodontists, who were trained by the researcher in the use of the annotation software. The researcher also compiled annotations, compiled the annotations, evaluated label usability, and verified the accuracy of the label index. After initial labeling, the endodontists reviewed and corrected each other's work and assigned classification labels to confirmed lesion images. The researcher then reviewed all data for proper formatting. Dataset construction took over eight months. Detailed information on PRAD is available on Github.

3 Method

3.1 Overview

The overall PRNet pipeline, illustrated in Fig. 3, comprises MWCN encoder blocks and the CFA mechanism for skip connections. Given an input image $X \in \mathbb{R}^{H \times W \times C}$, a convolutional stem block generates the initial feature map $X_0 \in \mathbb{R}^{H \times W \times C_s}$. The feature map X_0 is then passed through four MWCN stages, producing hierarchical features $F_i \in \mathbb{R}^{\frac{H}{2^{i-1}} \times \frac{W}{2^{i-1}} \times 2i \cdot C_s}$ $(i \in \{1, 2, 3, 4\})$. Each stage includes L_i MWCN blocks and a MaxPooling layer. F_4 enters the decoder, while X_0, F_1, F_2, and F_3 are processed by the CFA mechanism to produce weighted features, which are fused with decoder features via skip connections. Finally, the decoder output is passed through a segmentation head to generate the final segmentation mask $Y \in \mathbb{R}^{H \times W \times N}$, where N denotes the number of classes.

3.2 Multi-scale Wavelet Convolution Network

Inspired by [10], we leverage WTConv's large receptive field to capture global features in PR images. To address multi-scale challenges through global-local fea-

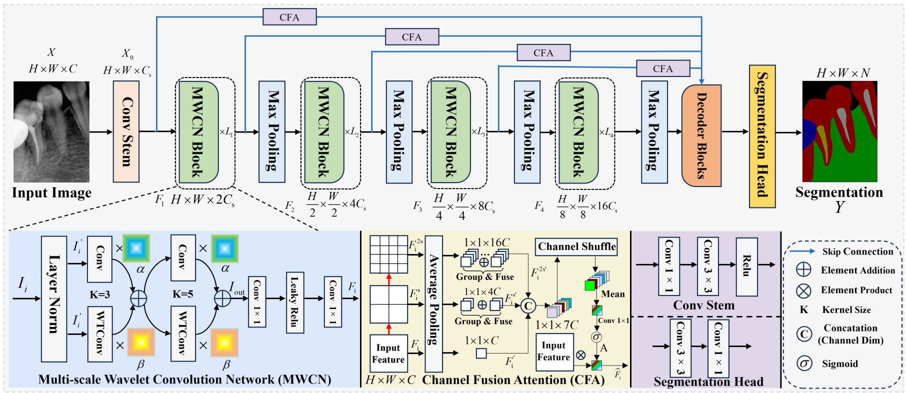

Fig. 3. Overall framework of the proposed PRNet. The structure of the decoder blocks are consistent with UNet.

ture integration, we designed MWCN blocks. As shown in the bottom left blue area of Fig. 3, given the input $I_i \in \mathbb{R}^{H\times W\times C}$, MWCN applies two convolutional layers and two WTConv layers with kernel sizes of 3 and 5. WTConv captures global context, while convolutions extract local details. Outputs from both are fused using two trainable Global-local Feature Weighting Matrices (GFWM) $\alpha, \beta \in \mathbb{R}^{H\times W\times 1}$, initialized as all-ones. GFWM enables adaptive weighting of global and local features. The fused output I_{out} is then passed through MaxPooling and into the next MWCN stage via a feed-forward block, as defined in Equation 1, where I_i' is the LayerNorm output of I.

$$\begin{aligned} I_1 &= \alpha \cdot (Conv_{k=3}(I_i')) + \beta \cdot (WTConv_{k=3}(I_i')) \\ I_2 &= \alpha \cdot (Conv_{k=5}(I_1)) + \beta \cdot (WTConv_{k=5}(I_1)) \\ F_i &= Conv_{k=1}(LeakyRelu(Conv_{k=1}(I_2))) \end{aligned} \tag{1}$$

3.3 Channel Fusion Attention

As shown in the yellow section at the bottom center of Fig. 3, the main structure of the CFA is illustrated. The primary function of the CFA is to weight the features from a channel perspective by integrating multi-level local features. This process enhances the decoder's ability to recognize objects of various sizes by feeding the feature layer, enriched with multi-scale information, into the corresponding decoder. The inputs fed into the CFA are the X_0 and the hierarchical outputs F_1, F_2 and F_3 from the corresponding MWCN blocks.

Given a hierarchical feature map input $F_i \in \mathbb{R}^{H\times W\times C}$, the first step is to partition the feature map into patches of size s and $2s$, respectively, resulting in feature maps $F_i^s \in \mathbb{R}^{\frac{H}{s}\times\frac{W}{s}\times s^2\cdot C}$ and $F_i^{2s} \in \mathbb{R}^{\frac{H}{2s}\times\frac{W}{2s}\times 4s^2\cdot C}$. In our experiments, s is set to 2. Subsequently, F_i, F_i^s and F_i^{2s} are each passed through an Average Pooling layer, which averages the features along the spatial dimensions H and W. For F_i^s and F_i^{2s}, the channel features after pooling are randomly grouped and summed according to s and $2s$, integrating local features and reducing dimensionality. Then, the three scaled feature maps are concatenated, followed by channel shuffling and averaging. The resulting feature is passed through a pointwise convolution layer followed by a sigmoid activation function to produce the attention map $A \in \mathbb{R}^{1\times 1\times C}$, as defined in Equation 2.

$$A = \mathrm{Sigmoid}(\mathrm{Conv}_{1\times 1}(\mathrm{Mean}(\mathrm{Channel\ Shuffle}(\mathrm{Concate}(F_i^{'}, F_i^{s^{'}}, F_i^{2s^{'}}))))) \quad (2)$$

Here, F_i', $F_i^{s'}$, and $F_i^{2s'}$ represent the outputs after grouping and summing the channel features at the three scales. Then, the attention map A is multiplied by the original features to perform channel weighting incorporating multi-scale information, as shown in equation 3, $\hat{F}_i$ represents the output of the CFA block.

$$\hat{F}_i = A \times F_i \quad (3)$$

4 Experiments

4.1 Implementation Details

We randomly selected 80% of PRAD-10K images as the training set and used the rest as the test set. Network parameters were randomly initialized, and the Adam optimizer with an initial learning rate of 0.0001 and a 'Poly' decay strategy was used. The loss combined Cross-Entropy and DICE losses. Input images were RGB resized to 256×256, trained for 200 epochs with a batch size of 12. Hierarchical feature channels were [64, 128, 256, 512], and MWCN blocks L_i were [1, 1, 2, 1]. Training was done on NVIDIA GeForce RTX 3090 GPUs using PyTorch. All experiments were repeated three times under the same settings, and average results are reported.

4.2 Comparisons with Other Methods

To evaluate the effectiveness of PRNet, we compared it with representative and recent SOTA medical image segmentation models. As shown in Table 2, PRNet achieved the best overall performance on PRAD-10K, with an average DSC of 84.28%. Compared to recent models like ACC-Unet (MICCAI'23), AHGNN (MICCAI'24), and EMCAD (CVPR'24), PRNet outperformed them by 8.78%, 4%, and 5.85%, respectively. PRNet also achieved top results in distinguishing dental crowns and implants, and performed best in identifying pulp, orthodontic devices, and apical periodontitis, demonstrating strong small-target and multi-scale capability. However, from an overall perspective, while PRNet performs well on large anatomical structures such as bones and teeth, its performance is relatively weaker when identifying implants and prosthetic crowns due to feature similarity. Large restorations that resemble crowns further increase the risk of misclassification. Addressing these challenges will be a key focus of future research.

Table 2. Quantitative comparison results of PRNet and other SOTA medical image segmentation methods on the PRAD-10K dataset, with DSC as the evaluation metric. Bold indicates the best performance, and underline indicates the second best.

Model	Tooth	Bone	Pulp	RCF	DC	DF	IM	OD	AP	**Avg.**
Unet [21]	91.55	92.17	<u>85.74</u>	84.82	55.33	75.35	63.69	89.71	84.61	80.33
Unet++ [32]	92.56	93.20	85.52	<u>85.31</u>	57.94	<u>78.07</u>	**64.70**	<u>90.38</u>	72.47	80.02
Atten-Unet [16]	<u>92.58</u>	<u>93.31</u>	85.61	84.54	<u>69.14</u>	75.19	63.63	89.14	84.25	<u>81.93</u>
MultiResUnet [12]	91.70	92.64	79.07	82.58	45.16	66.97	61.53	76.02	82.01	75.31
TransUnet [7]	91.34	92.05	75.23	**87.93**	81.50	59.28	61.16	90.19	74.01	79.19
Swin-Unet [4]	89.53	90.43	77.17	69.16	42.35	63.03	56.01	75.67	72.84	70.69
UNEXT [27]	90.41	91.66	74.22	55.71	47.64	64.12	56.89	72.09	76.09	69.87
MGFuseSeg [31]	91.58	92.50	84.89	73.27	48.33	67.96	62.42	70.01	75.54	73.94
ACC-Unet [11]	91.69	92.68	83.34	78.49	41.10	71.68	63.27	79.98	76.92	75.46
EMCAD [19]	91.62	92.40	81.98	80.44	63.26	69.52	62.36	83.48	80.41	78.39
TinyUnet [8]	88.62	89.92	63.47	71.37	40.09	43.69	55.27	68.19	88.64	67.60
AHGNN [5]	**92.60**	**93.45**	85.60	82.49	57.31	75.24	63.78	<u>88.49</u>	83.16	80.24
PRNet (ours)	92.38	93.02	**88.87**	82.89	**78.88**	**78.66**	<u>63.92</u>	**92.46**	**88.83**	**84.24**

Table 3. The quantitative results of the ablation experiments, ✓ indicates inclusion and ○ indicates exclusion.

	CFA	MWCN(k=3)	MWCN(k=5)	GFWM	DSC↑
UNet	○	○	○	○	80.33
	✓	○	○	○	82.89
UNet Decoder	✓	✓	○	○	83.65
	✓	○	✓	○	83.43
	✓	✓	✓	○	83.81
PRNet	✓	✓	✓	✓	**84.23**

4.3 Ablation experiments

We conducted ablation experiments to verify the effectiveness of each PRNet component, with results shown in Table 3. As seen in the second row, adding CFA blocks to the vanilla UNet's skip connections improved segmentation performance, increasing the average DSC by 2.56%. Next, we replaced the vanilla UNet encoder with our MWCN encoder while keeping the decoder and CFA blocks unchanged. To assess the two-scale MWCN encoder, we performed three comparative experiments. In the third and fourth rows, MWCN encoders with kernel sizes of 3 and 5 both outperformed the vanilla encoder. In the fifth row, the two-scale MWCN further improved performance over single-scale versions. Finally, the last row presents results from integrating GFWM into MWCN, forming PRNet. PRNet achieved the highest average DSC, confirming the effectiveness of GFWM. These results demonstrate that all PRNet components work synergistically and are indispensable.

5 Conclusion

This study presents PRAD-10K, the first and largest expert-annotated dataset for multi-class segmentation and classification of anatomical structures and lesions in PR images for endodontics. To establish a benchmark for PRAD-10K, we propose PRNet, a DL network with MWCN encoders and CFA mechanisms. Experiments show that PRNet achieves competitive performance on PRAD-10K. Future work will focus on further expanding and refining the PRAD dataset, as well as developing more efficient models for fully supervised, semi-supervised, or multimodal PR analysis.

Acknowledgments. This work is partially supported by the National Natural Science Foundation of China (62272248), the Natural Science Foundation of Tianjin (23JCZDJC01010, 23JCQNJC00010).

Disclosure of Interests. The authors have no competing interests.

References

1. Abdi, A.H., Kasaei, S., Mehdizadeh, M.: Automatic segmentation of mandible in panoramic x-ray. Journal of Medical Imaging **2**(4), 044003–044003 (2015)
2. Ari, T., Sağlam, H., Öksüzoğlu, H., Kazan, O., Bayrakdar, İ.Ş., Duman, S.B., Çelik, Ö., Jagtap, R., Futyma-Gąbka, K., Różyło-Kalinowska, I., et al.: Automatic feature segmentation in dental periapical radiographs. Diagnostics **12**(12), 3081 (2022)
3. Barot, S.: Oral cancer (lips and tongue) images (2022), https://www.kaggle.com/datasets/shivam17299/oral-cancer-lips-and-tongue-images/data
4. Cao, H., Wang, Y., Chen, J., Jiang, D., Zhang, X., Tian, Q., Wang, M.: Swin-unet: Unet-like pure transformer for medical image segmentation. In: European conference on computer vision. pp. 205–218. Springer (2022)
5. Chai, S., Jain, R.K., Mo, S., Liu, J., Yang, Y., Li, Y., Tateyama, T., Lin, L., Chen, Y.W.: A novel adaptive hypergraph neural network for enhancing medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 23–33. Springer (2024)
6. Chen, I.H., Lin, C.H., Lee, M.K., Chen, T.E., Lan, T.H., Chang, C.M., Tseng, T.Y., Wang, T., Du, J.K.: Convolutional-neural-network-based radiographs evaluation assisting in early diagnosis of the periodontal bone loss via periapical radiograph. Journal of Dental Sciences **19**(1), 550–559 (2024)
7. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
8. Chen, J., Chen, R., Wang, W., Cheng, J., Zhang, L., Chen, L.: Tinyu-net: Lighter yet better u-net with cascaded multi-receptive fields. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 626–635. Springer (2024)
9. Fatima, A., Shafi, I., Afzal, H., Mahmood, K., Díez, I.d.l.T., Lipari, V., Ballester, J.B., Ashraf, I.: Deep learning-based multiclass instance segmentation for dental lesion detection. In: Healthcare. vol. 11, p. 347. MDPI (2023)

10. Finder, S.E., Amoyal, R., Treister, E., Freifeld, O.: Wavelet convolutions for large receptive fields. In: European Conference on Computer Vision. pp. 363–380. Springer (2024)
11. Ibtehaz, N., Kihara, D.: Acc-unet: A completely convolutional unet model for the 2020s. In: International conference on medical image computing and computer-assisted intervention. pp. 692–702. Springer (2023)
12. Ibtehaz, N., Rahman, M.S.: Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. Neural networks **121**, 74–87 (2020)
13. Jang, T.J., Yun, H.S., Hyun, C.M., Kim, J.E., Lee, S.H., Seo, J.K.: Fully automatic integration of dental cbct images and full-arch intraoral impressions with stitching error correction via individual tooth segmentation and identification. Medical Image Analysis **93**, 103096 (2024)
14. Khan, H.A., Haider, M.A., Ansari, H.A., Ishaq, H., Kiyani, A., Sohail, K., Muhammad, M., Khurram, S.A.: Automated feature detection in dental periapical radiographs by using deep learning. Oral surgery, oral medicine, oral pathology and oral radiology **131**(6), 711–720 (2021)
15. Lo Giudice, R., Nicita, F., Puleio, F., Alibrandi, A., Cervino, G., Lizio, A., Pantaleo, G.: Accuracy of periapical radiography and cbct in endodontic evaluation. International journal of dentistry **2018**(1), 2514243 (2018)
16. Oktay, O., Schlemper, J., Folgoc, L.L., Lee, M., Heinrich, M., Misawa, K., Mori, K., McDonagh, S., Hammerla, N.Y., Kainz, B., et al.: Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 (2018)
17. Ozsari, S., Kamburoğlu, K., Tamse, A., Yener, S.E., Tsesis, I., Yılmaz, F., Rosen, E.: Automatic vertical root fracture detection on intraoral periapical radiographs with artificial intelligence-based image enhancement. Dental Traumatology (2025)
18. Panetta, K., Rajendran, R., Ramesh, A., Rao, S.P., Agaian, S.: Tufts dental database: a multimodal panoramic x-ray dataset for benchmarking diagnostic systems. IEEE journal of biomedical and health informatics **26**(4), 1650–1659 (2021)
19. Rahman, M.M., Munir, M., Marculescu, R.: Emcad: Efficient multi-scale convolutional attention decoding for medical image segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11769–11779 (2024)
20. Román, J.C.M., Fretes, V.R., Adorno, C.G., Silva, R.G., Noguera, J.L.V., Legal-Ayala, H., Mello-Román, J.D., Torres, R.D.E., Facon, J.: Panoramic dental radiography image enhancement using multiscale mathematical morphology. Sensors **21**(9), 3110 (2021)
21. Ronneberger, O., Fischer, P., Brox, T.: 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. pp. 234–241. Springer (2015)
22. Sarsam, W., Davies, J., Al-Salehi, S.K.: The role of imaging in endodontics. British Dental Journal **238**(7), 448–457 (2025)
23. Shokri, A., Salemi, F., Taherpour, T., Karkehabadi, H., Ramezani, K., Zahedi, F., Farhadian, M.: Is cone-beam computed tomography more accurate than periapical radiography for detection of vertical root fractures? a systematic review and meta-analysis. BMC Medical Imaging **24**(1), 286 (2024)
24. Stera, G., Giusti, M., Magnini, A., Calistri, L., Izzetti, R., Nardi, C.: Diagnostic accuracy of periapical radiography and panoramic radiography in the detection of apical periodontitis: a systematic review and meta-analysis. La radiologia medica **129**(11), 1682–1695 (2024)

25. Thalji, N., Aljarrah, E., Almomani, M.H., Raza, A., Migdady, H., Abualigah, L.: Segmented x-ray image data for diagnosing dental periapical diseases using deep learning. Data in Brief **54**, 110539 (2024)
26. Tibúrcio-Machado, C., Michelon, C., Zanatta, F., Gomes, M.S., Marin, J.A., Bier, C.A.: The global prevalence of apical periodontitis: a systematic review and meta-analysis. International endodontic journal **54**(5), 712–735 (2021)
27. Valanarasu, J.M.J., Patel, V.M.: Unext: Mlp-based rapid medical image segmentation network. In: International conference on medical image computing and computer-assisted intervention. pp. 23–33. Springer (2022)
28. Wang, H., Li, K., Zhu, J., Wang, F., Lian, C., Ma, J.: Weakly supervised tooth instance segmentation on 3d dental models with multi-label learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 723–733. Springer (2024)
29. Wang, X., Gao, S., Jiang, K., Zhang, H., Wang, L., Chen, F., Yu, J., Yang, F.: Multi-level uncertainty aware learning for semi-supervised dental panoramic caries segmentation. Neurocomputing **540**, 126208 (2023)
30. Wang, Y., Xia, W., Yan, Z., Zhao, L., Bian, X., Liu, C., Qi, Z., Zhang, S., Tang, Z.: Root canal treatment planning by automatic tooth and root canal segmentation in dental cbct with deep multi-task feature learning. Medical image analysis **85**, 102750 (2023)
31. Xu, G., Leng, X., Li, C., He, X., Wu, X.: Mgfuseseg: Attention-guided multi-granularity fusion for medical image segmentation. In: 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp. 3587–3594. IEEE (2023)
32. Zhou, Z., Siddiquee, M.M.R., Tajbakhsh, N., Liang, J.: Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE transactions on medical imaging **39**(6), 1856–1867 (2019)
33. Zou, B., Wang, S., Liu, H., Sun, G., Wang, Y., Zuo, F., Quan, C., Zhao, Y.: Teethseg: An efficient instance segmentation framework for orthodontic treatment based on multi-scale aggregation and anthropic prior knowledge. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11601–11610 (2024)