Mühendislik Bilimleri ve Tasarım Dergisi
14(2), 153-169, 2026
e-ISSN: 1308-6693

Journal of Engineering Sciences and Design
DOI: 10.21923/ jesd.1779687

Araştırma Makalesi Research Article

PATCH-BASED TOOTH SEGMENTATION ON CHILDREN'S DENTAL PANORAMIC RADIOGRAPHS

Nida KUMBASAR[1,2*], Emin Argun ORAL[2], Ibrahim Yucel OZBEK[3]

[1]TÜBİTAK, Informatics and Information Security Research Center (BİLGEM), Kocaeli, Türkiye
[2]Atatürk University, Faculty of Engineering, Department of Electrical Electronics Engineering, Erzurum, Türkiye
[3]Atatürk University, Faculty of Engineering, Department of Artificial Intelligence and Data Engineering, Erzurum, Türkiye

Keywords

Tooth Segmentation, Panoramic Radiograph, Patchify Image, Deep Learning, Artificial Intelligence.

Abstract

Dental Panoramic Radiography (DPR) is a widely utilized imaging modality for assessing oral health. However, accurate tooth segmentation in pediatric DPRs remains a significant challenge due to the variable morphology of deciduous teeth and the inherent complexity of overlapping anatomical structures in 2D projections. This study was conducted to enhance segmentation accuracy in pediatric DPRs by employing a patch-based deep learning (DL) approach. The models were trained using the relatively recent CDPR dataset, which consists of 193 pediatric DPRs and their corresponding ground truth masks. To investigate the optimal input size, the images were divided into 2, 4, 8, 16, and 32 patches, in addition to the full-resolution images. Five prominent DL models—PSPNet, DeepLabv3+, R2 U-Net, U-Net, and U-Net++—were trained and evaluated across these different patch configurations. A subject-based patch reconstruction method was subsequently implemented to rebuild the full-size segmented images for comprehensive evaluation. Experimental results indicated that the optimal segmentation performance for PSPNet, DeepLabv3+, and U-Net was achieved with the 8-patch approach, while R2 U-Net performed best with the 16-patch configuration. Notably, the U-Net++ model using a 2-patch approach outperformed all other models and configurations, achieving the highest Dice score of 0.9222. These findings demonstrate that patch-based training significantly improves segmentation performance in pediatric DPRs, which is a particularly valuable outcome under limited data conditions. This novel approach offers promising results for clinical integration, effectively supporting pediatric dental care by enhancing segmentation accuracy and overall workflow efficiency.

ÇOCUKLARIN DENTAL PANORAMİK RADYOGRAFİLERİNDE YAMA TABANLI DİŞ SEGMENTASYONU

Anahtar Kelimeler

Diş Segmentasyonu, Panoramik Radyografi, Görüntü Yamalama, Derin Öğrenme, Yapay Zeka.

Öz

Dental Panoramik Radyografisi (DPR), ağız sağlığını değerlendirmek için yaygın olarak kullanılan bir görüntüleme yöntemidir. Bununla birlikte, süt dişlerinin değişken morfolojisi ve 2 boyutlu projeksiyonlarda üst üste binen anatomik yapıların doğasında var olan karmaşıklık nedeniyle, pediatrik DPR'lerde doğru diş segmentasyonu önemli bir zorluk olmaya devam etmektedir. Bu çalışma, yama tabanlı derin öğrenme (DÖ) yaklaşımı kullanarak pediatrik DPR'lerde segmentasyon doğruluğunu artırmak amacıyla yapılmıştır. Modeller, 193 pediatrik DPR ve bunlara karşılık gelen gerçek maskelerden oluşan nispeten yeni CDPR veri seti kullanılarak eğitilmiştir. Optimal giriş boyutunu araştırmak için, görüntüler tam çözünürlüklü görüntülere ek olarak 2, 4, 8, 16 ve 32 yamaya bölünmüştür. Beş önde gelen DÖ modeli—PSPNet, DeepLabv3+, R2 U-Net, U-Net ve U-Net++—bu farklı yama konfigürasyonlarında eğitilmiş ve değerlendirilmiştir. Daha sonra, kapsamlı değerlendirme için tam boyutlu segmentlenmiş görüntüleri yeniden oluşturmak üzere özne tabanlı bir yama yeniden yapılandırma yöntemi uygulanmıştır. Deneysel sonuçlar, PSPNet, DeepLabv3+ ve U-Net için en iyi segmentasyon performansının 8-yama yaklaşımıyla elde edildiğini, R2 U-Net'in ise 16-yama konfigürasyonuyla en iyi performansı gösterdiğini ortaya koymuştur. Özellikle, 2-yama yaklaşımını kullanan U-Net++ modeli, diğer tüm modelleri ve konfigürasyonları geride bırakarak 0,9222'lik en yüksek Dice skoruna ulaşmıştır. Bu bulgular, yama tabanlı eğitimin pediatrik DPR'lerde segmentasyon performansını önemli ölçüde iyileştirdiğini göstermektedir; bu da özellikle sınırlı veri koşullarında son derece değerli bir sonuçtur. Bu yeni yaklaşım, segmentasyon doğruluğunu ve genel iş akışı verimliliğini artırarak pediatrik diş bakımını etkili bir şekilde desteklemek için klinik entegrasyon açısından umut verici sonuçlar sunmaktadır.

*Corresponding author: nida.kumbasar@tubitak.gov.tr, +90-262-675-3254

Alıntı / Cite		
Kumbasar, N., Oral, E. A., Ozbek, I. Y., (2026). Patch-Based Tooth Segmentation on Children's Dental Panoramic Radiographs, Journal of Engineering Sciences and Design, 14(2), 153-169.		
Yazar Kimliği / Author ID (ORCID Number)	**Makale Süreci / Article Process**	
N. Kumbasar, 0000-0001-5497-4618	**Başvuru Tarihi / Submission Date**	08.09.2025
E. A. Oral, 0000-0002-8120-9679	**Revizyon Tarihi / Revision Date**	26.04.2026
I. Y. Ozbek, 0000-0002-5734-7430	**Kabul Tarihi / Accepted Date**	30.04.2026
	Yayım Tarihi / Published Date	30.06.2026

PATCH-BASED TOOTH SEGMENTATION ON CHILDREN'S DENTAL PANORAMIC RADIOGRAPHS

Nida KUMBASAR[1,2†], Emin Argun ORAL[2], Ibrahim Yucel OZBEK[2,3]
[1]TÜBİTAK, Informatics and Information Security Research Center (BİLGEM), Kocaeli, Türkiye
[2]Atatürk University, Faculty of Engineering, Department of Electrical Electronics Engineering, Erzurum, Türkiye
[3]Atatürk University, Faculty of Engineering, Department of Artificial Intelligence and Data Engineering, Erzurum, Türkiye

Highlights

- Tooth segmentation on DPR images was performed using an innovative patch-based approach.
- Various DL models were benchmarked for pediatric tooth segmentation during the mixed dentition period.
- Patch-based training significantly enhances deep learning performance when utilizing datasets of limited size.

Purpose and Scope

Dental Panoramic Radiography (DPR) is a widely utilized imaging modality for assessing oral health. However, accurate tooth segmentation in pediatric DPRs remains a significant challenge due to the variable morphology of deciduous teeth and the inherent complexity of overlapping anatomical structures in 2D projections. This study aims to significantly enhance segmentation accuracy in pediatric DPRs by employing a patch-based deep learning (DL) approach using the relatively recent CDPR dataset.

Design/methodology/approach

The models were trained using the CDPR dataset, comprising 193 pediatric DPRs and their corresponding ground truth masks. To investigate the effect of input scale, the images were systematically divided into 2, 4, 8, 16, and 32 patches, in addition to being evaluated as full-resolution images. Five prominent DL models—PSPNet, DeepLabv3+, R2 U-Net, U-Net, and U-Net++—were trained and evaluated across these different patch configurations. A subject-based patch reconstruction method was subsequently implemented to rebuild the full-size segmented outputs for comprehensive final evaluation.

Findings

Performance was evaluated using the Dice score. Experimental results indicated that the optimal segmentation performance for PSPNet, DeepLabv3+, and U-Net was achieved with the 8-patch approach, while R2 U-Net performed best with the 16-patch configuration. Notably, the U-Net++ model using a 2-patch approach outperformed all other models and configurations, achieving the highest Dice score of 0.9222. Overall, the results indicate that patch-based training consistently and significantly improved segmentation performance.

Originality

These findings demonstrate the efficacy of patch-based training in substantially improving segmentation performance in pediatric DPRs, which is a particularly valuable outcome under limited data conditions. This novel approach offers promising results for clinical integration, effectively supporting pediatric dental care by enhancing segmentation accuracy and overall diagnostic workflow efficiency.

†Corresponding author: nida.kumbasar@tubitak.gov.tr, +90-262-675-3254

1. Introduction

Dental panoramic radiography (DPR) is a widely utilized preliminary diagnostic tool in dental practice. DPR enables clinicians to evaluate oral health in a comprehensive manner. DPR facilitates the analysis of various structures—including teeth, the mandibular canal, condyles, the jawbone, implants, and bridges—along with their associated pathologies (Güller, 2026). Low radiation rates, low cost, and offering fast and practical solutions are among the advantages of DPR. Distortions and overlaps occur in DPRs due to the two-dimensional projection of three-dimensional anatomical structures (Perschbacher, 2012). Therefore, interpreting the DPR and making an accurate and definitive diagnosis requires experience. Factors such as long working hours, high patient density, and variations in clinical experience may lead to challenges or inaccuracies in manual diagnostic interpretation.

In orthodontic treatments, factors such as tooth alignment, angulation, and positioning—along with dental caries, missing or supernumerary teeth, periapical lesions, cysts, and tumors—significantly influence the treatment process and strategy (Kumbasar et al., 2025; Güller et al., 2024). Accurate tooth segmentation and the clear determination of dental boundaries facilitate the analysis of odontogenic problems and the extraction of critical clinical information. Manual segmentation performed by experts varies depending on individual knowledge and experience; furthermore, it represents a time-consuming preparation phase (Ammar et al., 2011; Tosun et al., 2025). Consequently, the automated segmentation of teeth using artificial intelligence (AI) is of paramount importance in clinical dental applications.

In recent years, convolutional neural networks (CNN) have outperformed traditional image segmentation techniques (Sengun et al., 2023; Sedik et al., 2024) in terms of both accuracy and speed. The application of deep learning (DL) architectures to DPR images for tooth segmentation in dentistry (Almalki et al., 2023; Vural and Kumbasar, 2025), mandibular canal segmentation (X. Zhao et al., 2023; Cha et al., 2021), caries segmentation (Kumari et al., 2022; Day et al., 2023; X. Wang et al., 2023), lesion-cyst-tumor segmentation (Tosun et al., 2025; Yu et al., 2022), and jaw segmentation (Nafiiyah et al., 2022) are among the prevalent areas of study.

Inter-individual variability in oral anatomy and dimensions are among the factors that negatively affect model performance. Whether the patient is a child, adolescent, or elderly individual affects the number and alignment of teeth, thereby challenging the generalizability of computer vision models (Xu et al., 2018). To overcome this problem, categorizing training data by age group can enhance pattern recognition capabilities. DL architectures require substantial datasets to ensure robust learning and high performance. The fact that data labeling is a time-intensive process for experts, coupled with the restrictions on sharing medical and dental data publicly, poses challenges for supervised learning (Miloğlu et al., 2025, Kumbasar et al., 2026). In the literature, publicly available datasets in the dental field remain limited. For instance, Zhang et al. (2023) relabeled existing datasets—originally intended for different purposes—for tooth segmentation tasks to support the research community. Specifically, the CDPR dataset, containing children's DPRs from subjects between the ages of 2 and 13, was also labeled for tooth segmentation and introduced into the literature.

The remainder of this study is organized as follows: Section 2 provides a comprehensive literature survey. Section 3 presents the materials and methods, including the technical details of the implemented DL models. Section 4 details the performance analysis and experimental results. Finally, Section 5 discusses the findings, summarizes the proposed tooth segmentation approach, and outlines future research directions.

2. Literature Survey

Oral health is a critical domain of both dentistry and stomatology. Through image segmentation, radiographic images are partitioned into meaningful pixel clusters, each representing distinct anatomical or pathological features. Image segmentation is essential for enabling experts to efficiently analyze pathological regions across various radiographic modalities.

In recent years, algorithmic developments in CNN architectures, coupled with high-performance computational hardware, have significantly increased the performance of DL applications. As in many sectors, tooth segmentation in dentistry has shown promising results. Specifically, Koch et al. (2019) used a fully convolutional network (FCN) based on U-Net for tooth segmentation using DPR images, while Sheng et al. (2023) applied the transformer-based Swin-Unet. While Silva et al. (2018) utilized Mask R-CNN for tooth segmentation, Wirtz et al. (2018) proposed a fully automatic approach combining a statistical shape model with a neural network. Furthermore, an attention-based approach was adopted by Lai et al. (2020), whereas Y. Zhao et al. (2020) suggested a two-stage attention mechanism. In another study, Oktay and Gurses (2021) implemented the Mask R-CNN approach for the simultaneous detection, localization, and numbering of teeth. Mohan et al. (2023) observed that performing tooth segmentation with VGG-UNet before tooth recognition enhanced overall performance. Notably, no distinct distinction between pediatric and adult subjects was observed in the datasets used in the aforementioned studies. The presence of primary teeth in children creates unique patterns in DPRs; therefore, Zhang et al. (2023) adjusted existing DPR datasets for tooth segmentation and introduced the CDPR dataset, comprising images from children between the ages of 2 and 13. They utilized R2 U-Net, PSPNet, DeepLabv3+, and U-Net in their experiments on adult and child datasets, both independently and in combination.

In the literature, the patch-based segmentation approach has been applied to various challenges in the medical field. Specifically, Cordier et al. (2013) performed brain tissue segmentation based on similarities between multi-channel patches. Cui et al. (2016) utilized a patch-based CNN for brain MRI segmentation, while Xie et al. (2016) proposed the "patch forest" approach, which combines patch-based segmentation with random forests for hippocampus segmentation. Furthermore, retinal vessel segmentation was conducted by Wang et al. (2019) using a strategy based on Dense U-Net and patch-based learning. Similarly, Kim et al. (2020) implemented a 3D-patch-based deep CNN for abdominal multi-organ segmentation.

In this study, we investigated binary semantic segmentation of teeth using the publicly available CDPR dataset, which comprises DPR images of children aged 2 to 13 years. To effectively address the limitations of a small dataset and augment the training sample size, the DPRs were systematically divided into subject-based patches (at scales of 2, 4, 8, 16, and 32). Following the training process, a patch reconstruction method was implemented to combine the outputs and obtain full-size segmented images. The efficacy of the patch-based segmentation approach was rigorously tested on key DL architectures, including those previously evaluated in the original CDPR dataset study (Zhang et al., 2023)—namely PSPNet, DeepLabv3+, U-Net, and R2 U-Net—with the U-Net++ model additionally included in our analysis. Segmentation results encompassing a total of six different input configurations (comprising five patch sizes and the original resolution) and five distinct models are presented and compared using various performance metrics. Given the scarcity of dedicated tooth segmentation studies specifically targeting pediatric DPRs in the literature, this work holds significant potential for clinical application. The proposed methodology promises to enhance the diagnostic accuracy and efficiency of clinicians, thereby accelerating the overall treatment planning process.
The contributions of this study are summarized as follows:

- An innovative patch-based training approach was applied for the first time to input images for tooth segmentation in DPR.

- A comparative analysis of prominent DL models in medical imaging was conducted, specifically focusing on pediatric tooth segmentation during the mixed dentition period.

- This research demonstrates that for DL applications constrained by limited datasets, segmentation performance can be significantly enhanced by employing a patch-based data splitting strategy.

3. Material and Method

Accurate tooth segmentation in DPRs is crucial for comprehensive oral health analysis and treatment planning in both clinical and research settings. Standard CNN-based DL models often employ downsampling operations during segmentation, which can inherently result in the loss of fine-grained spatial details when processing full-resolution images. To address this limitation and enhance detail preservation, the current study utilized a patch-based training strategy. Specifically, each DPR image was divided into subject-based, non-overlapping patches using a patchification process. This approach served the dual purpose of preserving pixel-level details and effectively augmenting the dataset size for model training. During the testing phase, the segmented patches from the test data were then recombined on a subject basis, allowing the final performance evaluation to be conducted on the fully reconstructed DPR image.

The following subsections detail the dataset utilized in this study and the specific DL models applied for tooth segmentation.

3.1. Children's Dental Panoramic Radiographs, CDPR

Zhang et al. (2023) emphasized that DPRs, which capture the entire lower and upper jaw regions in a single image, serve as a valuable diagnostic tool for pediatric dental diseases; consequently, they introduced the CDPR dataset—the first publicly available dataset of children’s DPRs—to the literature. CDPR consists of 193 pediatric DPRs, including 100 cases with dental pathologies and 93 healthy cases, obtained from pediatric patients aged 2–13 who visited Hangzhou Xiasha Dental Hospital between March and June 2022. In addition to tooth segmentation, the dataset includes labels for the detection of caries and other dental diseases. Within the scope of this study, only the tooth segmentation problem was addressed, and AI models specific to pediatric tooth patterns were comparatively analyzed. An example DPR image and its corresponding mask from the dataset are shown in Figure 1.

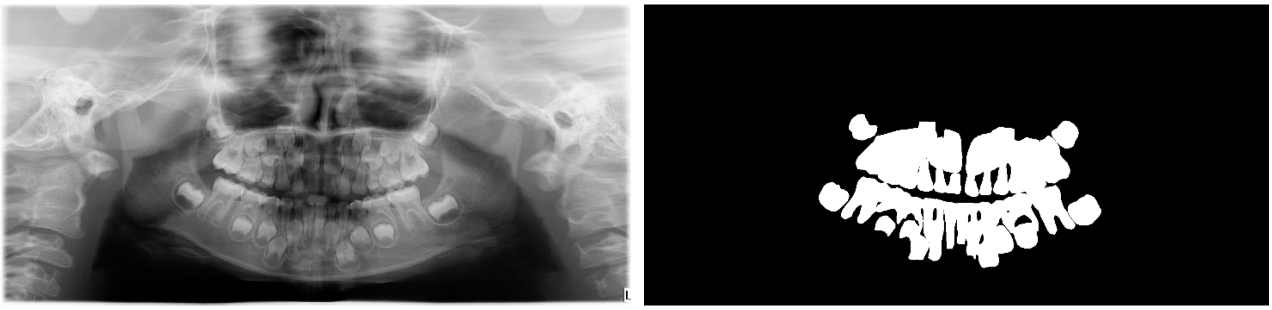


Figure 1. Illustrative example of a pediatric DPR and its corresponding segmentation mask from the CDPR dataset

3.2. Baseline Models

In this study, PSPNet, DeepLabv3+, R2 U-Net, and U-Net—as utilized by Zhang et al. (2023) in the original CDPR dataset publication—were selected to benchmark the patch-based segmentation approach against existing literature. Additionally, the U-Net++ architecture was fine-tuned and evaluated, representing a novel application of this model to the CDPR dataset. Panoramic images were divided into patches to enable the models to focus on local dental structures. Each patch was classified independently, and the subsequent results were combined to produce the final segmentation map. Although the patchification process can lead to minor information loss at boundaries, it enhances segmentation accuracy for small and heterogeneous dental regions. The advantages and potential limitations of this approach for clinical use are discussed in the following sections.

3.2.1 Pyramid Scene Parsing Network (PSPNet)

PSPNet (H. Zhao et al., 2017) is a semantic segmentation model that incorporates a Pyramid Pooling Module (PPM) to aggregate global context information from diverse regions. In PSPNet, local and global features are integrated to enhance the reliability of segmentation results. The PPM—which is based on the fusion of feature maps at multiple resolutions—captures multi-scale object information within the image. These generated feature maps are subsequently concatenated with the original feature map and convolved to produce the final output. Additionally, the architecture utilizes atrous convolution, a technique that expands the receptive field of the network without reducing spatial resolution. Moreover, atrous convolution enables the extraction of multi-scale features and comprehensive context without a significant increase in computational overhead.

3.2.2 DeepLabv3+

DeepLabv3 (Chen et al., 2017b) represents an enhanced iteration of the DeepLabv2 (Chen et al., 2017a) semantic segmentation architecture, incorporating several key modifications. In DeepLabv3, the challenge of segmenting objects at multiple scales is addressed through modules that employ atrous convolution in either cascaded or parallel configurations. Additionally, the Atrous Spatial Pyramid Pooling (ASPP) module from DeepLabv2 is augmented with image-level features, which further enhance segmentation performance. DeepLabv3+ (Chen et al., 2018) extends this architecture by incorporating a simple yet effective decoder module. In DeepLabv3+, the Xception architecture was applied to both the ASPP and decoder modules, resulting in a faster and more robust encoder-decoder network.

3.2.3 U-Net

U-Net (Ronneberger et al., 2015) is a seminal CNN-based architecture originally proposed to enhance the segmentation of biomedical images. The architecture features a characteristic U-shaped structure comprising an encoder (contracting path) and a decoder (expansive path). The encoder side utilizes successive convolution layers activated by ReLU, followed by max-pooling operations for feature extraction. On the decoder side, transposed convolutions (often referred to as deconvolution) are performed to restore spatial resolution. Furthermore, U-Net effectively addresses the information loss associated with the bottleneck by employing skip connections that transfer high-resolution features from the encoder directly to the decoder, thereby preserving fine-grained spatial details.

3.2.4 R2 U-Net

R2 U-Net (Alom et al., 2018) enhances the standard U-Net framework by integrating recurrent residual convolutional units (RRCU) for medical image segmentation. By leveraging recurrent and residual connections, the model facilitates better feature representation and deeper architecture training without the vanishing gradient problem. Furthermore, attention mechanisms incorporated within the architecture improve the model's ability to focus on salient anatomical features while suppressing irrelevant background noise, thereby significantly increasing segmentation performance.

3.2.5 U-Net++

U-Net++ (Zhou et al., 2018) is an advanced encoder-decoder architecture specifically designed to enhance medical image segmentation by bridging the semantic gap between the feature maps of the encoder and decoder sub-networks. This is achieved through a series of nested and intertwined dense skip pathways, which replace the traditional skip connections found in U-Net. These re-designed skip connections mitigate information loss during downsampling and upsampling operations, enabling the model to effectively integrate features across various resolution levels. Consequently, the architecture's ability to capture fine-grained spatial details is significantly improved. Each convolutional block within the network typically consists of a convolution layer followed by an activation function. While U-Net++ generally yields superior segmentation accuracy compared to the original U-Net, its dense connectivity pattern entails increased computational complexity and memory overhead.

4. Experimental Framework

4.1. Performance Metrics

Recall, specificity, pixel accuracy, intersection over union (IoU), and dice score (DS) metrics—which are frequently utilized in medical image segmentation—were employed in this work. These metrics were selected to provide a comprehensive quantitative evaluation of the models' performance. The abbreviations and definitions of the metrics used throughout the remainder of the study are presented in Table 1.

Table 1. Abbreviations used in performance metrics

ABBREVIATION	METRIC NAME	DESCRIPTION IN THIS WORK
Recall	Recall	Measures the proportion of actual tooth pixels correctly identified by the model.
Specificity	Specificity	Measures the proportion of background pixels correctly classified as non-tooth.
PixAcc	Pixel Accuracy	Overall proportion of correctly classified pixels (tooth + background).
IoU	Intersection over Union	Measures the overlap between predicted tooth regions and ground truth (GT), reflecting boundary and shape accuracy.
DS	Dice Score	Measures the overlap between predicted and GT tooth regions, giving more weight to correctly detected pixels.

In pixel-wise segmentation evaluation, True Positive (TP) refers to the number of pixels correctly identified as belonging to the target class. True Negative (TN) represents the pixels correctly classified as background or the non-target class. False Positive (FP) denotes pixels that were incorrectly labeled as the target class despite belonging to the background, while False Negative (FN) corresponds to pixels that belong to the target class but were mistakenly classified as background. These four fundamental values form the basis for calculating the performance metrics employed throughout this study.

They are calculated in terms of TP, TN, FP, and FN test results as defined in Equations 1-5 (Müller et al., 2022).

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{1}$$

$$\text{Specifity} = \frac{\text{TN}}{\text{TN+FP}} \tag{2}$$

$$\text{PixAcc} = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

$$\text{IoU} = \frac{TP}{TP+FP+FN} \tag{4}$$

$$\mathrm{DS} = \frac{2*TP}{2*\mathrm{TP}+\mathrm{FP}+\mathrm{FN}} \quad (5)$$

The metrics defined in Equations (1–5) were employed to quantitatively evaluate the segmentation performance of the proposed method. All metric values reported in the Results section and summarized in the performance tables were calculated using these formulas, based on the pixel-wise TP, TN, FP, and FN counts obtained from the segmentation outputs. In the comparative analysis with the reference study, the Dice Score was prioritized as the primary metric to ensure a consistent and objective performance assessment.

4.2. Experimental Setup

All experimental studies were conducted on a high-performance computing (HPC) unit equipped with an AMD EPYC 7H12 64-core CPU and dual NVIDIA RTX A40 GPUs, each providing 48 GB of VRAM. The proposed algorithms were implemented using the Python 3.8 programming language on the Ubuntu 18.04.4 LTS operating system. Deep learning models were developed and trained using the PyTorch framework, ensuring efficient utilization of the GPU resources for both training and inference.

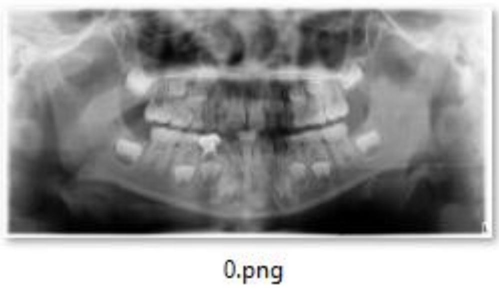
4.3. Experimental Methodology

Experimental studies were conducted following the training and test data splits specified by Zhang et al. (2023). To enhance model tuning, 15 images were randomly withdrawn from the training set to serve as a validation set. We retained the original training and test splits of the CDPR dataset (Zhang et al., 2023) to enable a rigorous and fair comparison with previous studies. This distribution allows for direct benchmarking of our patch-based segmentation approach while maintaining the integrity of the original dataset.

Consequently, the final dataset distribution comprised 148 training, 15 validation, and 30 test images at their original resolution. The DPR and corresponding mask pairs were systematically divided into 2, 4, 8, 16, and 32 patches, in addition to the original full-size images, as illustrated in Figure 2. In our implementation, the patches were extracted without overlapping to maintain a direct mapping to the original image coordinates. Following the model inference, the individual patches were reassembled using a deterministic coordinate-based reconstruction algorithm, which accurately maps each patch back to its original spatial position in the full-resolution DPR.

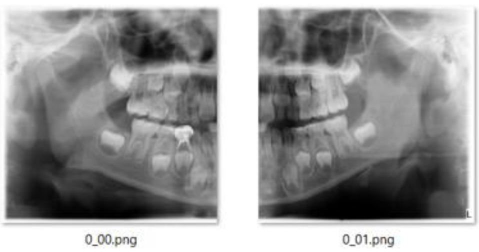
The PSPNet, DeepLabv3+, R2 U-Net, U-Net, and U-Net++ models were evaluated multiple times for each patch configuration. The input image resolution was set to 224×224 pixels. Optimal results across the models were observed using the Adam optimizer with a learning rate of 0.0001 over 100 epochs. The highest performance for original images (patch-1) and patch-2, patch-4, and patch-8 configurations was achieved with a batch size of 4, whereas batch sizes of 16 and 32 were utilized for patch-16 and patch-32, respectively. Furthermore, training was performed using various loss functions; while R2 U-Net yielded superior results with Binary Cross-Entropy (BCE) loss, the PSPNet, DeepLabv3+, U-Net, and U-Net++ models demonstrated higher performance using a hybrid loss function—specifically, the average of Dice Loss and BCE Loss. The optimal models were selected based on the highest DS achieved on the validation set.

Patch-1 DPR

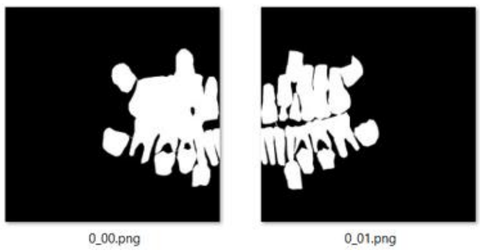


Patch-1 Mask

Patch-2 DPR



Patch-2 Mask



Patch-4 DPR

Patch-4 Mask

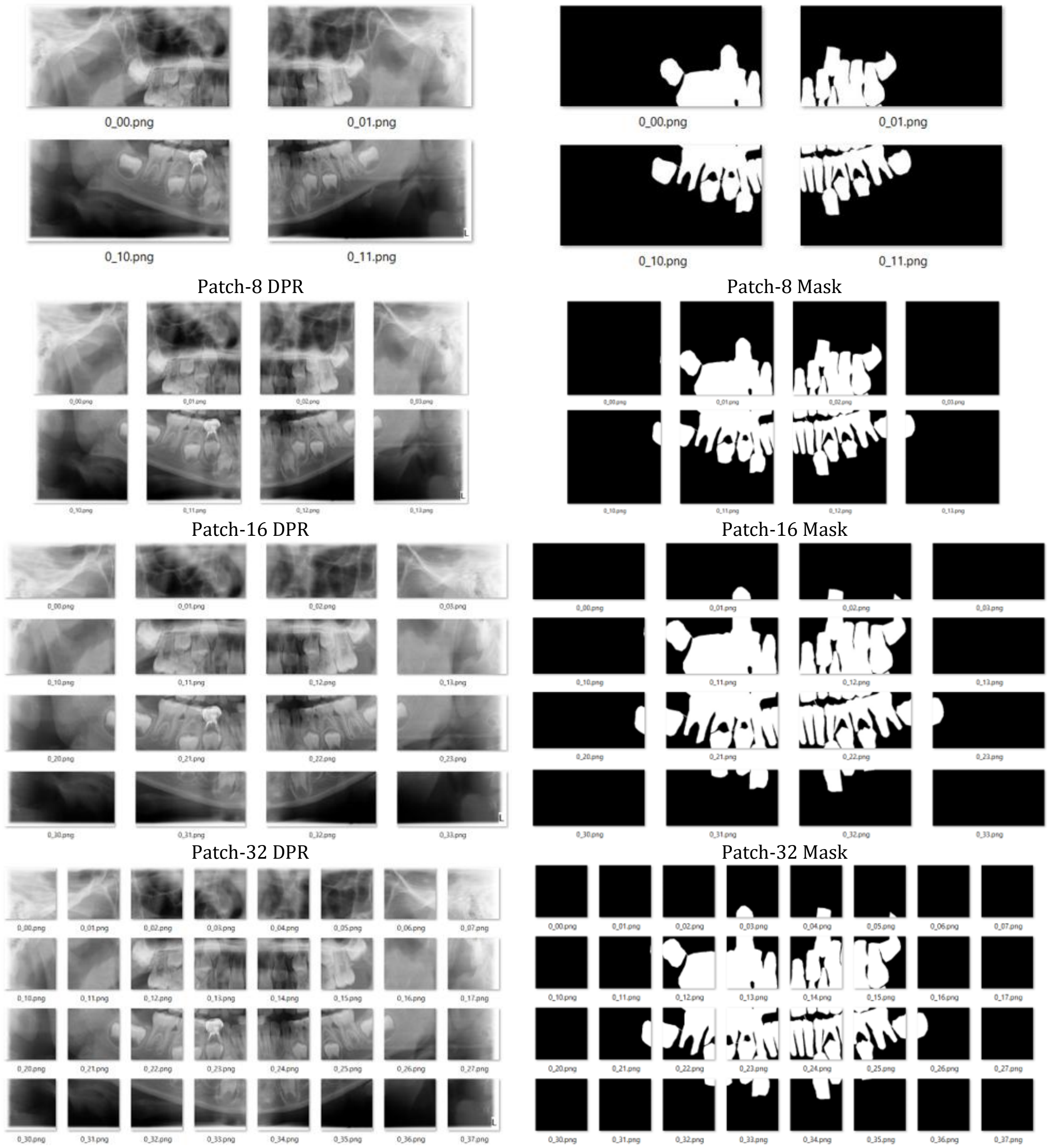


Figure 2. Visual representation of patch generation across different configurations

4.4. Experimental Results

In this section, the comprehensive and comparative performance results for tooth segmentation utilizing the PSPNet, DeepLabv3+, U-Net, R2 U-Net, and U-Net++ models on the CDPR dataset are presented.

Figure 3 provides a visual comparison, showcasing a sample DPR from the CDPR dataset, its GT mask, and the corresponding predicted mask outputs from the optimal configurations of the evaluated models. To highlight segmentation discrepancies, a residual error map was generated via the pixel-wise mathematical subtraction of the predicted mask from the GT mask. In this visual representation, red and white pixels denote FP and FN errors, respectively. Furthermore, a yellow ring is utilized to highlight specific regions of interest, enabling a clearer qualitative comparison across all model outputs.

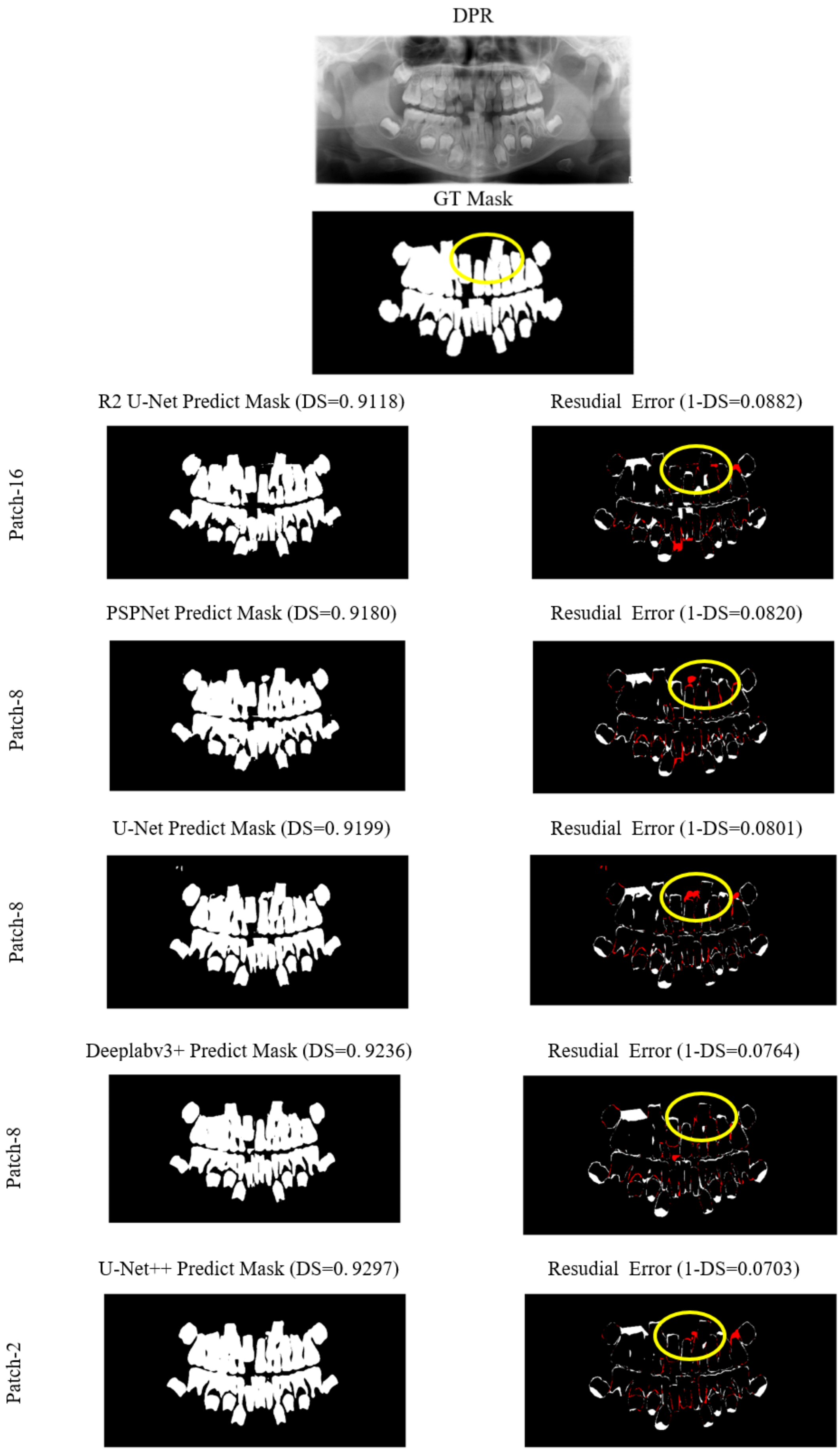


Figure 3. Comparison of predicted segmentation maps versus GT for different models

4.1.1. PSPNet

The segmentation performance results for the PSPNet model are visually detailed in Figure 4. Analysis of these results reveals that the lowest performance was recorded with the patch-2 configuration, yielding a DS of 0.8907. Conversely, the optimal performance was achieved using the patch-8 configuration, which resulted in a DS of 0.9105. When comparing the original resolution input (patch-1) to the patch-8 configuration, the patch-based training strategy yielded a performance enhancement of 0.33% in the DS.

Furthermore, Table 2 presents a comprehensive summary of the mean values and standard deviations for all performance metrics, categorized by the various patch configurations utilized for the PSPNet architecture.

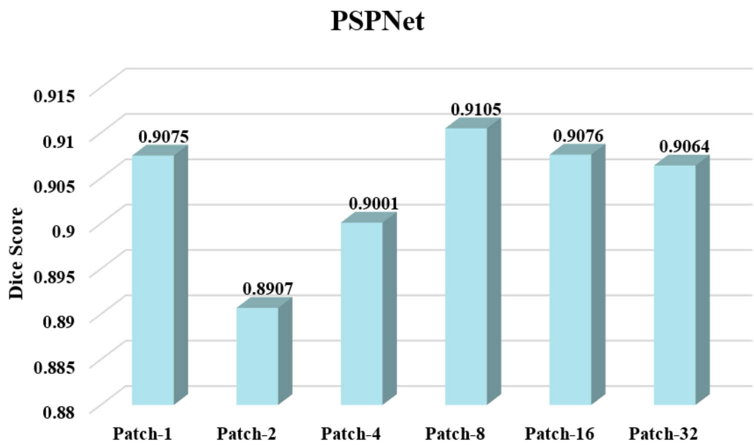


Figure 4. Performance evaluation of PSPNet using different patch-based strategies.

Table 2. PSPNet performance metrics results

Patch Type	Recall	Specificity	PixAcc	IoU	DS
Patch-1	0.8990 ±0.0218	0.9843 ±0.0042	0.9699 ±0.0057	0.8309 ± 0.0199	0.9075±0.0119
Patch-2	0.9020 ± 0.0158	0.9752± 0.0061	0.9628 ± 0.0065	0.8032± 0.0233	0.8907 ±0.0144
Patch-4	0.9093±0.0181	0.9778±0.0052	0.9660±0.0063	0.8185±0.0190	0.9001±0.0116
Patch-8	**0.8927±0.0188**	**0.9866±0.0039**	**0.9705±0.0057**	**0.8361±0.0228**	**0.9105±0.0137**
Patch-16	0.8859±0.0228	0.9869±0.0041	0.9697±0.0062	0.8312±0.0248	0.9076±0.0148
Patch-32	0.8842±0.0280	0.9868±0.0036	0.9695±0.0062	0.8294±0.0307	0.9064±0.0185

4.1.2. DeepLabv3+

Analysis of the segmentation results for the DeepLabv3+ model, as illustrated in Figure 5, reveals that the lowest performance was observed with the original resolution (patch-1) configuration, yielding a DS of 0.8750. Conversely, the optimal performance was achieved using the patch-8 configuration, which resulted in a DS of 0.9170. This comparison underscores that the patch-based training strategy provided a substantial performance gain of 4.8% for the DeepLabv3+ architecture. Furthermore, Table 3 provides a comprehensive summary of the mean values and standard deviations for the evaluated performance metrics across all tested patch configurations for DeepLabv3+.

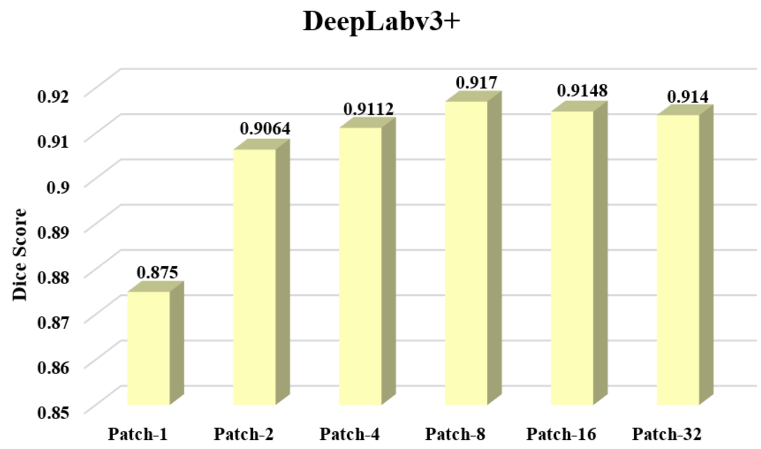


Figure 5. Performance evaluation of DeepLabv3+ using different patch-based strategies

Table 3. DeepLabv3+ performance metrics results

Patch Type	Recall	Specificity	PixAcc	IoU	DS
Patch-1	0.9085±0.0166	0.9670±0.0089	0.9573±0.0087	0.7782±0.0291	0.8750±0.0186
Patch-2	0.8925±0.0181	0.9847±0.0038	0.9692±0.0057	0.8291±0.0253	0.9064±0.0154
Patch-4	0.8994±0.0183	0.9853±0.0041	0.9706±0.0054	0.8372±0.0216	0.9112±0.0129
Patch-8	**0.9020±0.0199**	**0.9872±0.0044**	**0.9725± 0.0062**	**0.8470±0.0230**	**0.9170±0.0136**
Patch-16	0.9018±0.0225	0.9861±0.0045	0.9717±0.0061	0.8433±0.0222	0.9148±0.0130
Patch-32	0.8973±0.0256	0.9891±0.0039	0.9735±0.0058	0.8505±0.0244	0.9140±0.0143

4.1.3. R2 U-Net

The segmentation results obtained for the R2 U-Net model are illustrated in Figure 6. As demonstrated, the lowest performance was recorded with the original resolution (patch-1) configuration, yielding a DS of 0.9041. Conversely, the optimal performance was achieved using the patch-16 configuration, which resulted in a DS of 0.9192. Comparing the patch-1 results to the optimal patch-16 configuration, the patch-based training process provided a significant performance enhancement of 1.67% for the R2 U-Net architecture. Furthermore, Table 4 provides a comprehensive summary of the mean values and standard deviations for all performance metrics across the various patch types evaluated for the R2 U-Net model.

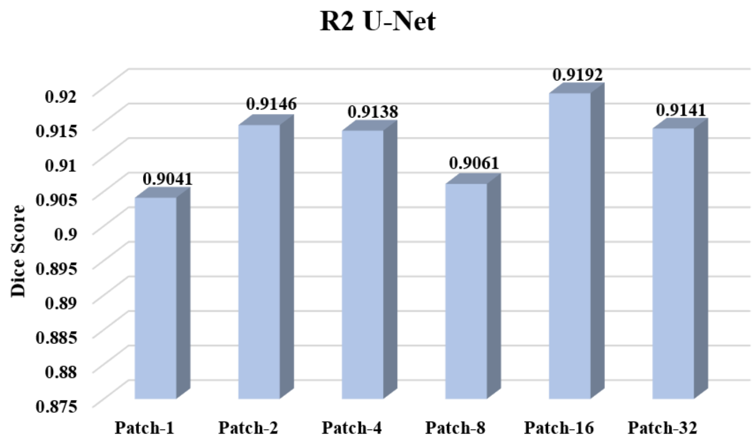


Figure 6. Performance evaluation of R2 U-Net using different patch-based strategies

Table 4. R2 U-Net performance metrics results

Patch Type	Recall	Specificity	PixAcc	IoU	DS
Patch-1	0.9231±0.0183	0.9767±0.0059	0.9679±0.0056	0.8253±0.0234	0.9041±0.0143
Patch-2	0.9430±0.0154	0.9769±0.0051	0.9712±0.0052	0.8430±0.0243	0.9146±0.0145
Patch-4	0.9353±0.0177	0.9783±0.0050	0.9711±0.0054	0.8415±0.0230	0.9138±0.0136
Patch-8	0.8888±0.0573	0.9862±0.0037	0.9704±0.0077	0.8298±0.0506	0.9061±0.0323
Patch-16	**0.9251±0.0212**	**0.9828±0.0048**	**0.9732±0.0060**	**0.8508±0.0209**	**0.9192±0.0123**
Patch-32	0.9035±0.0338	0.9859±0.0043	0.9721±0.0066	0.8423±0.0291	0.9141±0.0174

4.1.4. U-Net

The segmentation results for the U-Net model are illustrated in Figure 7. The findings indicate that the lowest performance was recorded with the original resolution (patch-1) configuration, yielding a DS of 0.9019. Conversely, the optimal performance was achieved using the patch-8 configuration, which resulted in a DS of 0.9193. For the U-Net architecture, the patch-based training strategy provided a notable performance enhancement of 1.92%. Furthermore, Table 5 provides a comprehensive summary of the mean values and standard deviations for the evaluated performance metrics across all tested patch configurations.

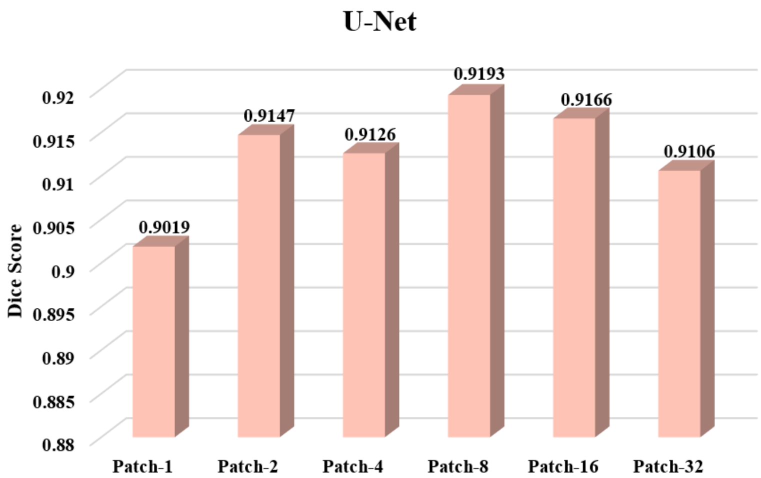


Figure 7. Performance evaluation of U-Net using different patch-based strategies

Table 5. U-Net performance metrics results

Patch Type	Recall	Specificity	PixAcc	IoU	DS
Patch-1	0.8822±0.0221	0.9856±0.0038	0.9686±0.0059	0.8216±0.0255	0.9019±0.0155
Patch-2	0.9012±0.0181	0.9864±0.0041	0.9717±0.0058	0.8430±0.0211	0.9147±0.0124
Patch-4	0.8853±0.0218	0.9892±0.0033	0.9715±0.0059	0.8395±0.0224	0.9126±0.0133
Patch-8	**0.9000±0.0255**	**0.9885±0.0039**	**0.9734±0.0061**	**0.8510±0.0260**	**0.9193±0.0154**
Patch-16	0.8898±0.0215	0.9895±0.0039	0.9727±0.0061	0.8462±0.0219	0.9166±0.0130
Patch-32	0.8858±0.0403	0.9882±0.0038	0.9710±0.0071	0.8366±0.0384	0.9106±0.0234

4.1.5. U-Net++

Analysis of the segmentation results for the U-Net++ model, as illustrated in Figure 8, reveals that the lowest performance was recorded with the original resolution (patch-1) configuration, yielding a DS of 0.9027. Conversely, the optimal performance was achieved using the patch-2 configuration, which resulted in a DS of 0.9222—the highest overall score recorded in this study. For the U-Net++ architecture, the patch-based training strategy provided a significant performance enhancement of 2.16%. Furthermore, Table 6 provides a comprehensive summary of the mean values and standard deviations for the evaluated performance metrics across all tested patch configurations for the U-Net++ model.

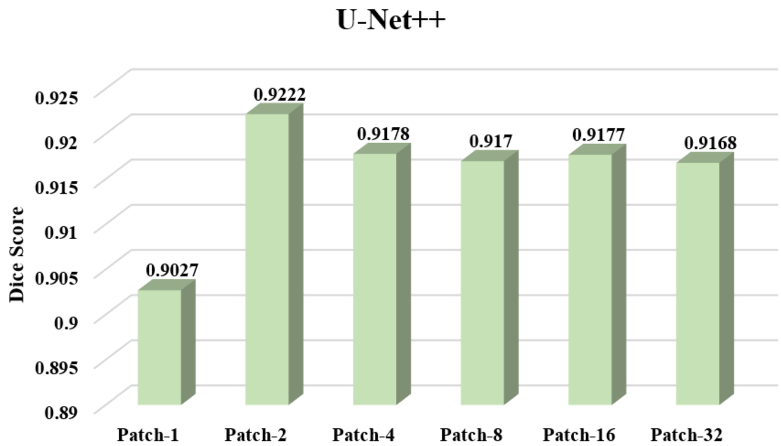


Figure 8. Performance evaluation of U-Net++ using different patch-based strategies

Table 6. U-Net++ performance metrics results

Patch Type	Recall	Specificity	PixAcc	IoU	DS
Patch-1	0.8870±0.0192	0.9850±0.0039	0.9687±0.0056	0.8230±0.0245	0.9027±0.0151
Patch-2	**0.9309±0.0192**	**0.9832±0.0039**	**0.9742±0.0055**	**0.8560±0.0229**	**0.9222±0.0134**
Patch-4	0.8990±0.0198	0.9880±0.0040	0.9729±0.0054	0.8483±0.0201	0.9178±0.0117
Patch-8	0.8906±0.0210	0.9898±0.0036	0.9727±0.0063	0.8470±0.0212	0.9170±0.0124
Patch-16	0.8959±0.0222	0.9887±0.0041	0.9729±0.0060	0.8482±0.0209	0.9177±0.0123
Patch-32	0.8955±0.0245	0.9885±0.0041	0.9727±0.0059	0.8466±0.0248	0.9168±0.0146

5. Result and Discussion

Radiological examinations are pivotal in clinical dentistry, enabling practitioners to comprehensively evaluate critical aspects such as tooth alignment, dental anomalies, impacted teeth, bone-tooth relationships, and the presence of oral cysts, tumors, fractures, or pathologies within the temporomandibular regions. Consequently, precise segmentation of dental structures has become increasingly vital for automated preliminary screening, diagnosis, and treatment planning. While various imaging modalities—including periapical, intraoral, and bite-wing radiographs, as well as cone-beam computed tomography (CBCT)—are utilized, DPR is particularly significant as it captures the entire oral cavity and surrounding maxillofacial structures in a single view. However, as a two-dimensional projection, DPR inherently suffers from the superimposition of anatomical structures, which frequently introduces structural distortions and overlapping artifacts. Given the complexity of these interpretations, relying solely on manual expert analysis carries inherent risks; factors such as clinical experience levels and practitioner fatigue can introduce intra-observer variability, potentially leading to diagnostic discrepancies or suboptimal treatment outcomes.

Automated segmentation in dentistry is becoming an essential tool during the preliminary diagnostic stage, facilitating the systematic examination of various dental pathologies. Several studies have significantly contributed to this field through diverse deep learning approaches. For instance, Jader et al. (2018) utilized the Mask R-CNN architecture to perform tooth segmentation on a dataset of 1,500 DPRs, achieving an F1-score of 0.88. Another pivotal contribution was the introduction of the Tufts Dental Database (TDD) (Panetta et al., 2021), where the authors evaluated a wide array of architectures, including FPN, U-Net, U-Net++, PSPNet, DeepLabv3, DeepLabv3+, nnU-Net, and CE-Net. Through extensive experimentation across different scenarios, they reported IoU values of 0.8642, 0.8667, 0.8662, 0.8608, 0.8621, 0.8641, 0.8611, and 0.8164, respectively.

Furthermore, Sheng et al. (2023) conducted comparative evaluations using U-Net, Link-Net, FPN, and Swin-Unet on two distinct datasets: the PLAGH-BH dataset (100 DPRs) and the PDX dataset (162 DPRs) (Nishitani et al., 2021). Their findings highlighted the impact of dataset complexity on model performance; for the PLAGH-BH dataset, the IoU values were 0.3674, 0.4030, 0.2407, and 0.4689, whereas for the PDX dataset, the models achieved significantly higher IoU values of 0.7455, 0.7820, 0.6597, and 0.6956, respectively.

Zhang et al. (2023) expanded the existing literature by labeling a total of 2,692 adult DPRs—previously utilized for various purposes—specifically for tooth segmentation. Furthermore, they introduced a novel dataset comprising 193 pediatric (children's) DPRs to address various diagnostic tasks, including tooth segmentation. Their study evaluated the performance of U-Net, R2 U-Net, PSPNet, and DeepLabv3+ models across three distinct dataset categories: pediatric, adult, and a combined adult+pediatric cohort.

For the pediatric dataset, the reported DS for U-Net, R2 U-Net, PSPNet, and DeepLabv3+ were 0.9120, 0.9027, 0.9083, and 0.8961, respectively. In the adult cohort (n=1,776), the models achieved higher scores of 0.9392, 0.9411, 0.9299, and 0.9267, respectively. Finally, utilizing the integrated adult+pediatric dataset (n=1,969), the reported DS values were 0.9365, 0.9373, 0.9278, and 0.9237 for the same four architectures.

In this study, tooth segmentation was successfully performed on the publicly available CDPR dataset, which comprises pediatric DPRs as introduced by Zhang et al. (2023). Our methodology systematically divided the DPRs into patches to serve two critical objectives: first, to augment the effective training data volume, and second, to mitigate structural distortions typically caused by resizing high-resolution radiographs to meet the fixed input dimensions of DL models. To facilitate a rigorous and direct comparison with the baseline results established by Zhang et al. (2023), their original architectures (PSPNet, DeepLabv3+, U-Net, and R2 U-Net) were evaluated across five different patch configurations (2, 4, 8, 16, and 32) in addition to the original resolution.

Furthermore, since the U-Net architecture yielded the best performance in the reference study, its advanced variant, U-Net++, was integrated into this comparative analysis as an architectural enhancement. Each evaluated method demonstrated a unique optimal patch configuration, as illustrated in Figure 9. Specifically, a patch-8 configuration is recommended for PSPNet, DeepLabv3+, and U-Net, whereas patch-16 and patch-2 configurations proved optimal for R2 U-Net and U-Net++, respectively. Notably, square-shaped patches generally outperformed rectangular configurations. This superior performance is attributed to the fact that deep learning models typically process square inputs; consequently, square-like patches undergo minimal aspect ratio deformation during internal resizing, thereby preserving spatial integrity and minimizing data degradation.

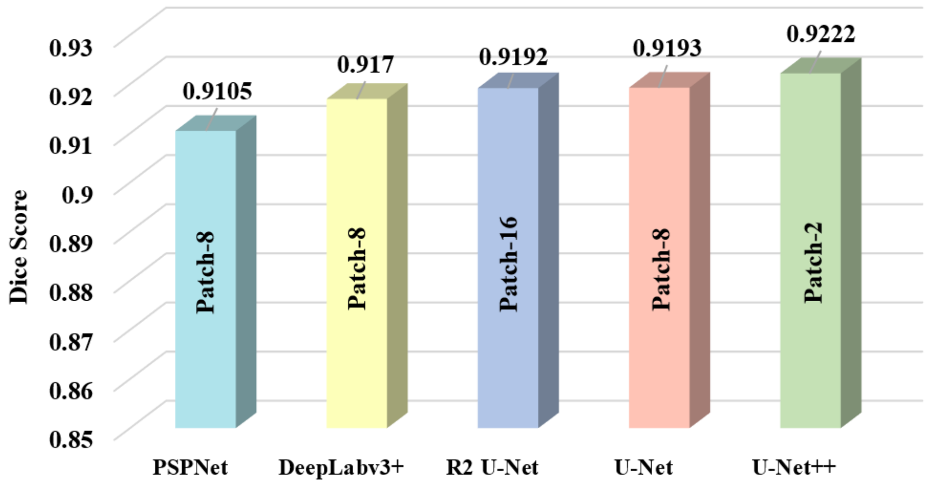


Figure 9. Comparative analysis of the peak performance metrics achieved by each model across its optimal patch configuration

As illustrated in Figure 10, the patch-based approach proposed in this study yielded significant performance gains over the baseline results reported in the literature. Specifically, the DS improvements recorded were 0.24%, 2.33%, 1.82%, and 0.8% for the PSPNet, DeepLabv3+, R2 U-Net, and U-Net models, respectively. Furthermore, the U-Net++ architecture outperformed the most successful model documented in previous studies by a margin of 1.11%, achieving the highest overall accuracy. Consequently, this study proposes the patch-2 based U-Net++ framework as a robust method for automated dental diagnosis. This approach is particularly well-suited for pediatric DPR analysis in clinical settings, offering a significant potential to enhance the precision of preliminary examinations and treatment planning in dental practice.

In this study, a comprehensive comparative analysis for tooth segmentation was successfully conducted using the publicly available CDPR pediatric dataset. To overcome the challenges of limited data availability—a common constraint in DL—a systematic patch-based training strategy was implemented. Final evaluations were performed using a subject-based patch reconstruction approach, ensuring that global spatial consistency was maintained. This strategy effectively minimized the structural distortion typically introduced during the image resizing phase, thereby preserving the fine-grained anatomical details essential for accurate segmentation. The patches were reassembled based on their original spatial coordinates to ensure anatomical continuity. Since patching allows the model to process images at a higher resolution, it captures sharper edge details. This high precision at the edges naturally minimizes the alignment errors and 'seam' artifacts typically seen at the boundaries of reassembled images.

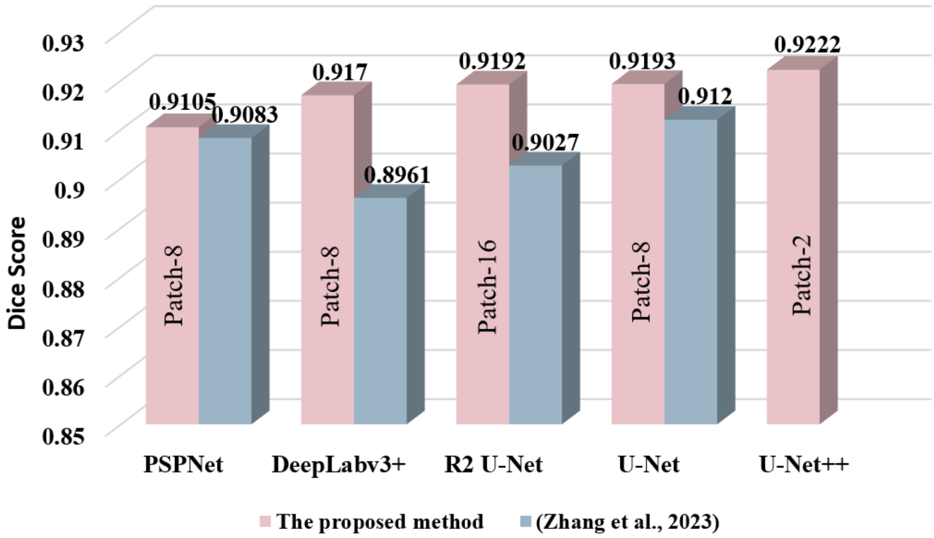


Figure 10. Performance comparison between the proposed patch-based models and existing studies utilizing the same CDPR dataset

The presence of deciduous (milk) teeth in pediatric radiographs introduces distinct morphological patterns and developmental variations that differ significantly from those in adult DPRs. Consequently, this study demonstrates that a dedicated, pediatric-specific AI framework provides more robust and precise solutions for automated segmentation in clinical environments. Through extensive experimentation with six patch configurations (1, 2, 4, 8, 16, and 32) across five state-of-the-art architectures (PSPNet, DeepLabv3+, U-Net, R2 U-Net, and U-Net++), we identified the patch-2 based U-Net++ model as the optimal solution. Achieving the highest overall performance, this proposed model exhibits significant potential for integration into real-world clinical workflows, facilitating fully automated and reliable tooth segmentation for pediatric dental care.

The primary contribution of this study to the dental AI literature lies in demonstrating that a strategic patch-based decomposition significantly overcomes the resolution limitations of standard DL models without losing clinical context. This approach provides a scalable framework that achieves superior precision over full-image training, particularly in specialized datasets like pediatric DPRs where fine morphological detail is paramount. The primary advantages of the patching strategy include enabling high-resolution processing of DPRs, facilitating the extraction of localized anatomical features, improving memory efficiency, and enhancing the model's robustness against image variations. Conversely, potential limitations of this method include the risk of information loss at patch boundaries, the necessity for post-processing reconstruction to recombine the patches, and the challenge of maintaining contextual integrity for clinical evaluation. In this study, we demonstrate that a systematic patching approach does not result in clinically significant information loss; rather, it significantly contributes to improved model performance and segmentation precision.

Future research will focus on investigating the impact of the patch-based approach on larger and more diverse datasets to further validate its scalability. Additionally, we plan to conduct a comparative evaluation of Generative Adversarial Network (GAN)-based architectures against state-of-the-art transformer-based models to explore their respective capabilities in addressing the complexities of the tooth segmentation problem.

Acknowledgement

This work is supported by Ataturk University Scientific Research Projects Grant FDK-2021-9758. Thanks to the Ataturk University High Performance Computing Application and Research Center for computing support.

Conflict of Interest

No conflict of interest was declared by the authors.

References

Almalki, S.A., Alsubai, S., Alqahtani, A., Alenazi, A.A., 2023. Denoised Encoder-Based Residual U-Net for Precise Teeth Image Segmentation and Damage Prediction on Panoramic Radiographs. Journal of Dentistry, 137, 104651.

Alom, M.Z., Hasan, M., Yakopcic, C., Taha, T.M., Asari, V.K., 2018. Recurrent Residual Convolutional Neural Network Based on U-Net (R2U-Net) for Medical Image Segmentation. ArXiv Preprint, ArXiv:1802.06955.

Ammar, H.H., Ngan, P., Crout, R.J., Mucino, V.H., Mukdadi, O.M., 2011. Three-Dimensional Modeling and Finite Element Analysis in Treatment Planning for Orthodontic Tooth Movement. American Journal of Orthodontics and Dentofacial Orthopedics, 139 (1), e59–e71.

Cha, J.-Y., Yoon, H.-I., Yeo, I.-S., Huh, K.-H., Han, J.-S., 2021. Panoptic Segmentation on Panoramic Radiographs: Deep Learning-Based Segmentation of Various Structures Including Maxillary Sinus and Mandibular Canal. Journal of Clinical Medicine, 10 (12), 2577.

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L., 2017a. Deeplab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Transactions on Pattern Analysis and Machine Intelligence, 40 (4), 834–848.

Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017b. Rethinking Atrous Convolution for Semantic Image Segmentation. ArXiv, abs/1706.05587.

Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. Proceedings of the European Conference on Computer Vision (ECCV), 801–818.

Cordier, N., Menze, B., Delingette, H., Ayache, N., 2013. Patch-Based Segmentation of Brain Tissues. MICCAI Challenge on Multimodal Brain Tumor Segmentation, 6–17.

Cui, Z., Yang, J., Qiao, Y., 2016. Brain MRI Segmentation with Patch-Based CNN Approach. 35th Chinese Control Conference (CCC), 7026–7031.

Day, B., Üzen, H., Çiçek, İ.B., Duman, Ş.B., 2023. A Novel Deep Learning-Based Approach for Segmentation of Different Type Caries Lesions on Panoramic Radiographs. Diagnostics, 13 (2), 202.

Güller, M.T., Kumbasar, N., Miloğlu, Ö., 2024. Evaluation of the Effectiveness of Panoramic Radiography in Impacted Mandibular Third Molars on Deep Learning Models Developed with Findings Obtained with Cone Beam Computed Tomography. Oral Radiology, 1–16.

Güller, M.T., Kumbasar, N., Miloğlu, Ö., Özbek, İ.Y., Oral, E.A., 2026. Winter Classification of Mandibular Impacted Third Molars Based On Convolutional Neural Networks. Current Research in Dental Sciences, 36 (2), 133–139.

Jader, G., Fontineli, J., Ruiz, M., Abdalla, K., Pithon, M., Oliveira, L., 2018. Deep Instance Segmentation of Teeth in Panoramic X-Ray Images. 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 400–407.

Kim, H., Jung, J., Kim, J., Cho, B., Kwak, J., Jang, J.Y., Lee, S., Lee, J.-G., Yoon, S.M., 2020. Abdominal Multi-Organ Auto-Segmentation Using 3D-Patch-Based Deep Convolutional Neural Network. Scientific Reports, 10 (1), 6204.

Koch, T.L., Perslev, M., Igel, C., Brandt, S.S., 2019. Accurate Segmentation of Dental Panoramic Radiographs with U-Nets. IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), 15–19.

Kumari, A.R., Rao, S.N., Reddy, P.R., 2022. Design of Hybrid Dental Caries Segmentation and Caries Detection with Meta-Heuristic-Based ResneXt-RNN. Biomedical Signal Processing and Control, 78, 103961.

Kumbasar, N., Güller, M.T., Miloğlu, Ö., Oral, E.A., Ozbek, I.Y., 2025. Deep-Learning Based Fusion of Spatial Relationship Classification Between Mandibular Third Molar and Inferior Alveolar Nerve Using Panoramic Radiograph Images. Biomedical Signal Processing and Control, 100, 107059.

Kumbasar, N., Oral, E.A., Ozbek, I.Y., 2026. Fusion Based Collective Intelligence Segmentation Using Panoramic Radiograph Images. Human-Intelligent Systems Integration, 1–18.

Lai, Y., Fan, F., Wu, Q., Ke, W., Liao, P., Deng, Z., Chen, H., Zhang, Y., 2020. Lcanet: Learnable Connected Attention Network for Human Identification Using Dental Images. IEEE Transactions on Medical Imaging, 40 (3), 905–915.

Miloğlu, Ö., Kumbasar, N., Tosun, Z.T., Güller, M.T., Özbek, İ.Y., 2025. Gender Classification With Hand-Wrist Radiographs Using the Deep Learning Method. Current Research in Dental Sciences, 35 (1), 2–7.

Mohan, R., Arunmozhi, R., Rajinikanth, V., 2023. Deep-Learning Segmentation and Recognition of Tooth in Thresholded Panoramic X-Ray. Winter Summit on Smart Computing and Networks (WiSSCoN), 1–5.

Müller, D., Soto-Rey, I., Kramer, F., 2022. Towards a Guideline for Evaluation Metrics in Medical Image Segmentation. BMC Research Notes, 15 (1), 1–8.

Nafiiyah, N., Fatichah, C., Herumurti, D., Astuti, E.R., Putra, R.H., Prakasa, E., 2022. Mandibular Segmentation on Panoramic Radiographs with CNN Transfer Learning. IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), 190–194.

Nishitani, Y., Nakayama, R., Hayashi, D., Hizukuri, A., Murata, K., 2021. Segmentation of Teeth in Panoramic Dental X-Ray Images Using U-Net with a Loss Function Weighted on the Tooth Edge. Radiological Physics and Technology, 14, 64–69.

Oktay, A.B., Gurses, A., 2021. Detection, Segmentation, and Numbering of Teeth in Dental Panoramic Images with Mask Regions with Convolutional Neural Network Features. State of the Art in Neural Networks and their Applications, içinde (s. 73–90). Elsevier.

Panetta, K., Rajendran, R., Ramesh, A., Rao, S.P., Agaian, S., 2021. Tufts Dental Database: A Multimodal Panoramic X-Ray Dataset for Benchmarking Diagnostic Systems. IEEE Journal of Biomedical and Health Informatics, 26 (4), 1650–1659.

Perschbacher, S., 2012. Interpretation of Panoramic Radiographs. Australian Dental Journal, 57, 40–45.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015, 18, 234–241.

Sedik, A., Kolivand, H., Albeedan, M., 2024. An Efficient Image Classification and Segmentation Method for Crime Investigation Applications. Multimedia Tools and Applications, 1–25.

Sengun, B., Iscan, Y., Ozbulak, G.A.T., Kumbasar, N., Egriboz, E., Sormaz, I.C., Aksakal, N., Deniz, S.M., Haklidir, M., Tunca, F., et al.,

2023. Artificial Intelligence in Minimally Invasive Adrenalectomy: Using Deep Learning to Identify the Left Adrenal Vein. Surgical Laparoscopy Endoscopy & Percutaneous Techniques, 33 (4), 327–331.

Sheng, C., Wang, L., Huang, Z., Wang, T., Guo, Y., Hou, W., Xu, L., Wang, J., Yan, X., 2023. Transformer-Based Deep Learning Network for Tooth Segmentation on Panoramic Radiographs. Journal of Systems Science and Complexity, 36 (1), 257–272.

Silva, G., Oliveira, L., Pithon, M., 2018. Automatic Segmenting Teeth in X-Ray Images: Trends, a Novel Data Set, Benchmarking and Future Perspectives. Expert Systems with Applications, 107, 15–31.

Tosun, Z.T., Kumbasar, N., Sumbullu, M.A., Miloglu, O., 2025. Evaluation of the Effectiveness of Artificial Intelligence Models in Radiopaque and Radiolucent Lesions of the Maxillofacial Region on Panoramic Radiographs. Oral Radiology, 1–15.

Vural, Ş.F., Kumbasar, N., 2025. Comparison of VT-Based and CNN-Based Models on Teeth Segmentation. Balkan Journal of Electrical and Computer Engineering, 13 (2), 148–156.

Wang, C., Zhao, Z., Ren, Q., Xu, Y., Yu, Y., 2019. Dense U-Net Based on Patch-Based Learning for Retinal Vessel Segmentation. Entropy, 21 (2), 168.

Wang, X., Gao, S., Jiang, K., Zhang, H., Wang, L., Chen, F., Yu, J., Yang, F., 2023. Multi-Level Uncertainty Aware Learning for Semi-Supervised Dental Panoramic Caries Segmentation. Neurocomputing, 540, 126208.

Wirtz, A., Mirashi, S.G., Wesarg, S., 2018. Automatic Teeth Segmentation in Panoramic X-Ray Images Using a Coupled Shape Model in Combination with a Neural Network. Medical Image Computing and Computer Assisted Intervention–MICCAI 2018, 11, 712–719.

Xie, Z., Gillies, D., 2016. Patch Forest: A Hybrid Framework of Random Forest and Patch-Based Segmentation. Medical Imaging 2016: Image Processing, 9784, 598–605.

Xu, X., Liu, C., Zheng, Y., 2018. 3D Tooth Segmentation and Labeling Using Deep Convolutional Neural Networks. IEEE Transactions on Visualization and Computer Graphics, 25 (7), 2336–2348.

Yu, D., Hu, J., Feng, Z., Song, M., Zhu, H., 2022. Deep Learning Based Diagnosis for Cysts and Tumors of Jaw with Massive Healthy Samples. Scientific Reports, 12 (1), 1855.

Zhang, Y., Ye, F., Chen, L., Xu, F., Chen, X., Wu, H., Cao, M., Li, Y., Wang, Y., Huang, X., 2023. Children's Dental Panoramic Radiographs Dataset for Caries Segmentation and Dental Disease Detection. Scientific Data, 10 (1), 380.

Zhao, H., Shi, J., Qi, X., Wang, X., Jia, J., 2017. Pyramid Scene Parsing Network. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2881–2890.

Zhao, X., Xu, T., Peng, L., Li, S., Zhao, Y., Liu, H., He, J., Liang, S., 2023. Recognition and Segmentation of Teeth and Mandibular Nerve Canals in Panoramic Dental X-Rays by Mask RCNN. Displays, 78, 102447.

Zhao, Y., Li, P., Gao, C., Liu, Y., Chen, Q., Yang, F., Meng, D., 2020. TSASNet: Tooth Segmentation on Dental Panoramic X-Ray Images by Two-Stage Attention Segmentation Network. Knowledge-Based Systems, 206, 106338.

Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J., 2018. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J.M.R.S. Tavares, A. Bradley, J.P. Papa, V. Belagiannis, J.C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, A. Madabhushi (Edt.), Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support içinde (s. 3–11). Berlin: Springer International Publishing.