BIO Web of Conferences 174, 03020 (2025) https://doi.org/10.1051/bioconf/202517403020
ICBB 2025

Artificial Intelligence Precision Recognition and Auxiliary Diagnosis of Dental X-ray Panoramic Images Based on Deep Learning

Riming Liu[1,*], Zhenshan Gao[2]

[1]School of Stomatology, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China
[2]Liushi Oral Hospital, Liushi, 323000, Zhejiang, China

Abstract: Objective: This study aims to explore the application of deep learning algorithms in dental X-ray panoramic images, particularly for the automatic segmentation of dental caries and identification of wisdom tooth types, in order to improve the accuracy and efficiency of dental diagnosis and assist doctors in formulating precise treatment plans. Methods: Multiple classic medical image segmentation network models (including Unet, PSPNet, FPN, Unet++, and DeepLabV3+) were trained and tested on the ParaDentCaries dataset to evaluate their performance in dental X-ray panoramic images. Performance was comprehensively compared using evaluation metrics such as IoU (Intersection over Union), Dice coefficient, sensitivity, accuracy, Hd95 (Hausdorff distance), and model parameters. Additionally, visualization methods were used to display the model's prediction results across different lesion scales (small caries, medium caries, large caries). Results: The experimental results show that the Unet++ model performed best across all evaluation metrics, especially achieving strong results in IoU (58.94), Dice coefficient (72.71), sensitivity (68.03), and accuracy (82.00). Compared to other models (such as FPN, PSPNet, DeepLabV3+), Unet++ demonstrated clear advantages in detail recognition and boundary handling, particularly exhibiting higher accuracy in detecting small and medium caries. Visual analysis showed that Unet++ was able to accurately identify secondary carious areas, while PSPNet and DeepLabV3+ performed poorly in this regard, showing boundary detection deviations. Conclusion: The deep learning-based automatic diagnostic system for dental X-ray panoramic images, especially the Unet++ model, provides high-precision predictive results in dental caries segmentation and wisdom tooth type identification, significantly improving diagnostic efficiency and accuracy.

1 Introduction

In recent years, with economic development and improved living standards, the prevalence of oral diseases has been on the rise, especially among children, with dental caries being particularly prominent. In China, the prevalence of dental caries in permanent teeth among 12-year-old children is as high as 34.5%, while the prevalence of dental caries in primary teeth among 5-year-old children has reached 70.9%. As oral health issues become an increasing global concern, the prevention, diagnosis, and treatment of oral diseases have gradually gained attention, particularly dental caries, which, as one of the most common oral diseases, has a widespread impact on people's quality of life and overall health. Therefore, improving the efficiency of early diagnosis of oral diseases and reducing the diagnostic pressure on doctors has become an urgent issue that needs to be addressed [1].

Traditional oral diagnostic methods rely on visual inspection and the experience of doctors. However, due to the complexity of the oral structure, and the often irregular and progressive nature of dental caries, misdiagnosis and missed diagnoses can occur during the diagnostic process [2]. Oral X-rays, especially panoramic X-rays, are an important auxiliary diagnostic tool that provides imaging information for the entire mouth and is widely used for caries detection and localization. Despite this, the image quality of oral X-rays is often affected by noise and artifacts, and doctors must process large amounts of imaging data within a limited time, making accurate and efficient diagnosis of oral X-rays a significant challenge.

In recent years, with the rapid development of artificial intelligence, particularly deep learning technologies, deep learning algorithms have made significant progress in the field of medical image processing, especially in tasks such as image classification, object detection, and image segmentation, showing superior performance. By analyzing oral X-rays using deep learning algorithms, precise identification and classification of carious lesions can be achieved, effectively improving the accuracy and efficiency of diagnosis. Therefore, conducting research on artificial intelligence-based precise recognition and auxiliary diagnosis of dental panoramic X-rays using deep learning has important practical significance and application value [3-4].

* Corresponding author: c18357757158@126.com

2 Materials and Methods

2.1 Introduction to CNN

Convolutional Neural Networks (CNNs) are a class of deep learning models widely used in fields such as image processing, speech recognition, and natural language processing. The network structure of CNN is illustrated in Figure 1.

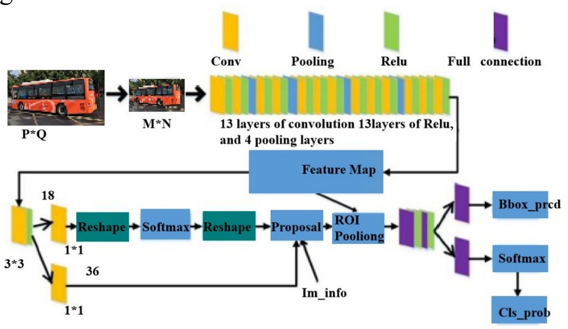

Figure 1. Network Structure of CNN

The mathematical formula of Convolution layer is shown in Formula 1.

$$x_i^l = f\left\{\sum_{j\in M_i} x_i^{l-1} * k_i^{l-1} + b_i^{l-1}\right\} \tag{1}$$

Among them, x_i^l represents the set of input characteristic graphs. *Is the convolution operation, b_i^{l-1} is the corresponding offset term.

The mathematical formula of Activation layer is shown in Formula 2.

$$f(x) = \begin{cases} x, if & x \geq 0 \\ 0, if & x < 0 \end{cases} \tag{2}$$

The mathematical formula of Pool layer is shown in Formula 3. The pooling layer can be abstractly represented by mathematical expression.

$$x_i^l = f(\beta_i^l down(x_i^{l-1}) + b_i^{l-1}) \tag{3}$$

Among them, $down$ is the downsampling function, x_i^l is the sampling result, b_i^{l-1} is the corresponding offset term. The mathematical formula of Full connection layer is shown in Formula 4.

$$x_i^l = f\left\{\sum_{j\in M_i}^{n} x_j^{l-1} w_{i,j}^{l-1} + b_i^{l-1}\right\} \tag{4}$$

2.2 Deep Learning Algorithm Models

U-Net is a deep learning model for medical image segmentation, originally proposed by Olaf Ronneberger and others. It uses an encoder-decoder structure, where the encoder gradually extracts features from the image, and the decoder restores the spatial information through upsampling and skip connections. Skip connections help retain detailed information, enabling U-Net to perform well on small sample images. U-Net is primarily applied in medical image segmentation (e.g., tissue or tumor segmentation in CT and MRI images) [5].

PSPNet is a network that captures multi-scale contextual information through a pyramid pooling module. The model combines global information and local details by pooling the image at different scales, enhancing its ability in semantic segmentation in complex scenarios. PSPNet focuses on scene parsing and can handle multi-level and multi-scale image features. It is mainly applied in urban street scene analysis, remote sensing images, and semantic segmentation [6].

FPN enhances multi-scale features for image segmentation and object detection through a feature pyramid. It combines features extracted from different levels to create an up-to-bottom pyramid structure, allowing the model to learn from features ranging from low to high resolution. The goal of FPN is to provide precise multi-scale predictions and improve the performance of traditional methods in handling objects of various sizes. FPN is mainly applied in object detection and instance segmentation.

U-Net++ is an improved version of U-Net, which introduces more skip connections and deep supervision mechanisms between the decoder and encoder. This allows the network to capture detailed information better when processing complex images. The architecture of U-Net++ is more complex than that of U-Net, but it improves the accuracy of image segmentation, particularly in complex scenarios and imbalanced data. U-Net++ is primarily applied to medical image segmentation and remote sensing image analysis [7].

DeepLabV3+ is the latest version of the DeepLab series, combining dilated convolution and encoder-decoder structures to capture rich contextual information and enhance segmentation accuracy. DeepLabV3+ introduces refined boundary processing and global context modeling, especially excelling in handling boundaries and details. DeepLabV3+ is mainly applied to image segmentation, particularly in scene parsing, street image segmentation, and road detection [8].

2.3 Deep Learning-Based Image Segmentation Methods

Deep learning-based image segmentation methods mainly include the following three types:

1) Semantic Segmentation: Semantic segmentation assigns each pixel in the image to a specific category, aiming to achieve automatic understanding of the entire image. Each pixel is labeled to indicate its belonging to a particular object category (e.g., background, person, etc.). Semantic segmentation is suitable for tasks that require distinguishing between different object categories but cannot differentiate between objects of the same category [9].

2) Instance Segmentation: Instance segmentation not only distinguishes object categories but also identifies different instances within the same category. This means that even if objects overlap or touch, instance

segmentation can accurately segment each object. This method is widely used in object detection and complex scene analysis.

3) Panoptic Segmentation: Panoptic segmentation combines both semantic segmentation and instance segmentation, enabling it to label each object with a semantic tag and accurately distinguish between different object instances within the same category. This technology has been effectively applied in fields like autonomous driving and medical imaging, handling more complex scenarios [10].

2.4 Model Evaluation Metrics

Accuracy, IoU, Dice, Precision, and Sensitivity are commonly used metrics to evaluate the performance of classification models.

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}} \tag{5}$$

$$IoU = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \tag{6}$$

$$Dice = \frac{2 * \text{TP}}{2 * \text{TP} + \text{FP} + \text{FN}} \tag{7}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

Where TP is true positive, TN is true negative, FN is false negative, and FP is false positive.

3 Results

3.1 System Function Demonstration

This system utilizes deep learning algorithms, particularly the VggNet16 model, to accurately identify the types of wisdom teeth in dental X-ray panoramic images and assist doctors in diagnosis and treatment planning. Users can perform different operations based on their permissions. Patients, doctors, and administrators can log in to the system and perform corresponding functions based on their needs. The system can automatically identify the types of wisdom teeth and provide personalized treatment plans based on the patient's preferences and best evidence, which will be confirmed after review by the doctor. All relevant data, such as X-ray images, treatment records, and doctor's suggestions, can be saved and managed for subsequent queries. Through these functions, the system not only improves diagnostic accuracy but also optimizes the treatment planning process, enhancing overall medical efficiency. The process of identification and obtaining preliminary treatment plans is shown in Figure 2.

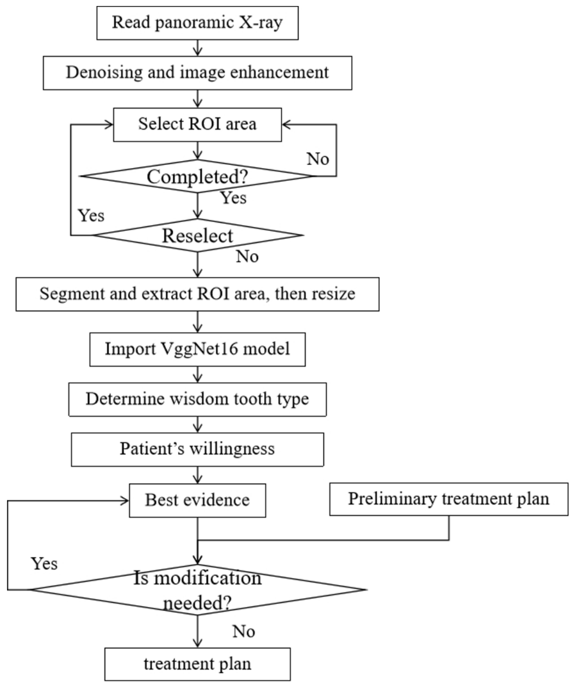

Figure 2. Process of Identification and Obtaining Preliminary Treatment Plans

3.2 Model Training Comparison Experiment Results

Based on the ParaDentCaries dataset, several classic medical image segmentation network models were selected for training and testing to evaluate the performance of different models in the task of dental caries segmentation in panoramic dental X-rays. The selected models include: the pure CNN-based Unet, PSPNet, FPN, Unet++, and DeepLabV3+. To comprehensively assess the effectiveness of each model, a performance comparison was made with the latest caries segmentation model, CariesNet.

The experimental evaluation metrics include IoU (Intersection over Union), Dice coefficient, sensitivity, accuracy, Hd95 (Hausdorff distance), and the model parameter count. All models were tested on the dental panoramic image test set proposed in this paper. The comparison results of the different models are shown in Table 1 for a comprehensive analysis of the strengths and weaknesses of different segmentation methods.

Table 1. Comparison Results of Different Methods

Method	Unet	PSPNet	FPN	Unet++	DeepLabV3+
IoU	56.78	55.02	57.66	58.94	54.23
Dice	71.23	69.58	71.9	72.71	68.61
Sensitivity	69.22	64.86	68.02	68.03	66.58
Precision	77.69	80.3	79.81	82	75.87
Hd95↓	5.95	6.09	5.79	5.84	6.18
Parameter count	24.4M	21.3M	23.1M	26.1M	22.4M

Based on the results in Table 1, it can be seen that Unet++ is the optimal choice, showing balanced performance across all evaluation metrics, particularly achieving good results in IoU, Dice coefficient, sensitivity, and accuracy. FPN and Unet also have strong competitiveness, performing well in sensitivity and accuracy, making them suitable for applications that require higher sensitivity. Although DeepLabV3+ and PSPNet perform better in terms of the number of parameters and some metrics, they are slightly weaker in sensitivity and accuracy.

3.3 Experimental Visualization Results

To more intuitively demonstrate the prediction results of the model proposed in this chapter, the predicted results of different models on the ParaDentCaries panoramic X-ray test set are shown through visualization methods. The selected models include Unet, PSPNet, FPN, Unet++, and DeepLabV3+. All model tests were conducted on dental slices that underwent overlap cropping treatment, which included caries lesions of different scales, such as small caries, medium caries, and large caries, with medium caries being secondary caries, which appear as erosion regions gradually spreading behind white filling material. Visualization results of different model methods on multi-scale slices of the ParaDentCaries test set are shown in Figure 3.

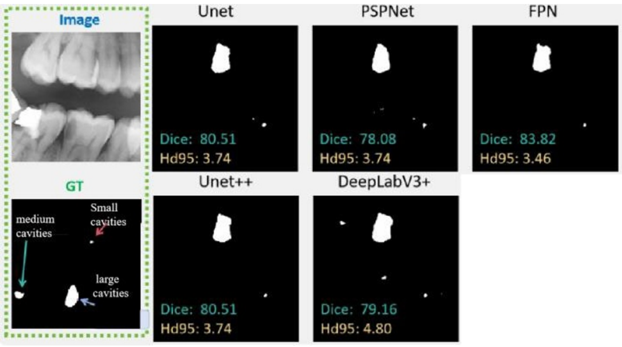

Figure 3. Visualization results of different model methods on multi-scale slices of the ParaDentCaries test set

As shown in Figure 3, the performance of each model differs on different lesion scales. There is a significant performance difference when handling small and medium caries. For medium caries, especially in the secondary caries region, some models like Unet++ and TransUnet can accurately identify the lesion area, while other models like PSPNet and DeepLabV3+ have some deviations in boundary detection, making it difficult to accurately distinguish the erosion area behind the white filling material. The prediction performance for large caries is relatively good, with most models accurately segmenting the lesion area. However, some models still need improvement in handling complex boundaries.

By comparing the experimental results and visualized images, we can further confirm that the model proposed in this chapter demonstrates strong advantages in recognizing and segmenting multi-scale caries lesions, especially in terms of better precision and stability in detail recognition and boundary processing.

3.4 Experimental Results

The visualization of the performance of different algorithms on the pediatric dental dataset is shown in Figure 4.

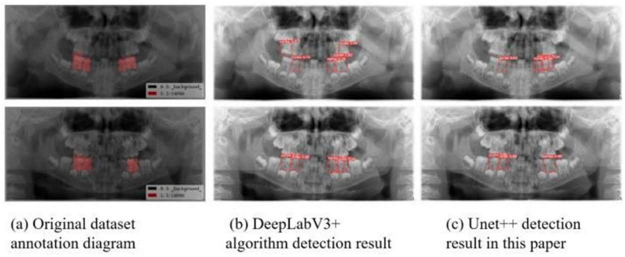

Figure 4. Visualization of the performance of different algorithms on the pediatric dental dataset

To intuitively evaluate the effects of the improvements, a visualization analysis was conducted. As shown in Figure 4, the pink rectangular area in (a) represents dental caries. From the image, it can be seen that the Unet++ algorithm detected results more accurately and comprehensively. This indicates that the improvements made to the algorithm indeed had a significant effect. However, there is still considerable room for improvement in the final result. For example, some caries in the first image were not fully detected, which is primarily due to the characteristics of the dataset. Future research will focus on solving this issue.

4 Discussion

This paper explores the technical advancements and challenges in the precise identification and auxiliary diagnosis of oral diseases, particularly dental caries, based on deep learning for panoramic dental X-ray images. From the experimental results, deep learning models have made significant progress in analyzing dental X-ray images, especially in caries detection and segmentation. However, despite the preliminary achievements, there are still areas that require further optimization. The specific reasons can be summarized as follows:

Firstly, the scarcity of dataset quantity remains one of the main factors limiting model performance. The current pediatric dental X-ray dataset is insufficient in both quantity and diversity, especially when compared to adult dental data. Pediatric teeth have significantly different imaging characteristics and are underdeveloped, which poses considerable challenges for model training. The existing datasets lack sufficient sample size and diversity, particularly for complex cases such as different types of dental caries, which reduces the model's generalization ability in practical applications.

Secondly, the low contrast of dental X-ray images is another important factor affecting diagnostic accuracy. Due to the small density differences between oral tissues, the contrast of X-ray images is relatively low, making it more difficult to differentiate between various tissues and structures. This issue is especially prominent when processing pediatric dental images, where the boundaries of caries lesions are often unclear, and some lesion areas are even difficult to distinguish with the naked eye.

Finally, although the model proposed in this paper shows good performance across multiple evaluation metrics, there are still challenges in real-world applications, particularly in clinical environments. Diagnostic accuracy still has room for improvement, especially in handling complex cases, where the model may not yet achieve the precise judgment of clinical doctors. This issue indicates that deep learning models cannot fully replace doctors' judgment but should serve as a supplementary tool to assist doctors in improving diagnostic efficiency and accuracy.

5 Conclusions

The deep learning-based automatic diagnostic system for dental panoramic X-rays, particularly the Unet++ model, can provide high-precision predictions in dental caries segmentation and wisdom tooth type identification, significantly improving diagnostic efficiency and accuracy. Although the current system's performance has shown strong competitiveness, it still faces challenges such as insufficient datasets, low image quality, and high task complexity. To further enhance the application of artificial intelligence in dental disease diagnosis, future research should focus on expanding datasets, improving image processing techniques, and optimizing models. Additionally, the application of AI technology in the medical field needs to be closely integrated with clinical work to create a more efficient and intelligent auxiliary diagnostic system, better serving both doctors and patients and advancing the development of oral health.

References

1. Ding Zhongxi, Zhong Hao, Hu Liefeng. Research on X-ray security inspection image detection and recognition based on improved Faster R-CNN [J]. Microcomputer Applications, 2024, 40(12): 194-198.
2. Zhang Yawen. Application of X-ray examination technology in diagnosing limb fractures [J]. Modern Medical Imaging, 2024, 33(09): 1686-1687+1691.
3. Yang Yihong, Xu Tingxuan, Ma Chengyang, et al. Design and application of an AI-based auxiliary diagnostic mini-program for oral diseases [J]. Computer Programming Techniques and Maintenance, 2024, (09): 58-60+129.
4. Jiang Kaisheng. Research on algorithms for fine segmentation and accurate identification of caries in dental panoramic X-rays [D]. Hangzhou University of Electronic Science and Technology, 2024.
5. Jin Heqi, Gai Shaoyan, Da Feipeng. Chest X-ray image classification algorithm based on convolutional neural networks [J]. Information Technology and Informatization, 2024, (01): 33-41.
6. Zhang Jilong, Zhao Jun, Li Jinlong. Improved YOLOv7 X-ray image hazardous item detection algorithm [J]. Computer Engineering and Applications, 2024, 60(10): 266-275.
7. Xiang Jiao, Li Guoquan, Wu Jian, et al. Improved YOLOv5s algorithm for detecting prohibited items in X-ray security inspection images [J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2023, 35(05): 943-951.
8. Zeng Hongxiang, Wen Zhicheng. X-ray prohibited item detection algorithm based on improved YOLOv5 [J]. Computer Engineering and Applications, 2024, 60(16): 217-227.
9. Li Yongjian, Zhu Huasheng, He Mingzhi, et al. X-ray image detection algorithm based on YOLO-MCA [J]. Journal of Nanchang Engineering Institute, 2023, 42(03): 82-87.
10. Li Jing, Huang Xianbin, Zhang Siqun, et al. Smoothing spectrum algorithm for soft X-ray Dante spectrometer and its application in Z compression diagnosis [J]. High Power Laser and Particle Beams, 2009, 21(11): 1655-1660.