IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 10 December 2023, accepted 22 December 2023, date of publication 1 January 2024, date of current version 25 January 2024.

Digital Object Identifier 10.1109/ACCESS.2023.3348788

RESEARCH ARTICLE

Detection of Tooth Position by YOLOv4 and Various Dental Problems Based on CNN With Bitewing Radiograph

KUO-CHEN LI[1], YI-CHENG MAO[2], MU-FENG LIN[3], YI-QIAN LI[3], CHIUNG-AN CHEN[4], TSUNG-YI CHEN[5], (Member, IEEE), AND PATRICIA ANGELA R. ABU[6], (Member, IEEE)

[1]Department of Information Management, Chung Yuan Christian University, Taoyuan 320317, Taiwan
[2]Department of General Dentistry, Chang Gung Memorial Hospital, Taoyuan 33305, Taiwan
[3]Department of Electronic Engineering, Chung Yuan Christian University, Zhongli, Taoyuan 32023, Taiwan
[4]Department of Electrical Engineering, Ming Chi University of Technology, New Taipei City 243303, Taiwan
[5]Department of Electronic Engineering, Feng Chia University, Taichung 40724, Taiwan
[6]Department of Information Systems and Computer Science, Ateneo de Manila University, Quezon City 1108, Philippines

Corresponding author: Tsung-Yi Chen (tsungychen@fcu.edu.tw)

This work was supported in part by the National Science and Technology Council, Taiwan, under Grant 111-2221-E-033-041, Grant 111-2823-8-033-001, Grant 111-2622-E-131-001, Grant 110-2223-8-033-002, Grant 110-2221-E-027-044-MY3, Grant 110-2218-E-035-007, Grant 110-2622-E-131-002, Grant 109-2622-E-131-001-CC3, Grant 109-2221-E-131-025, Grant 109-2410-H-197-002-MY3, Grant 112-2410-H-197-002-MY2, Grant 112-2222-E-035-004, and Grant 112-2410-H-033-014.

ABSTRACT Periodontitis is a high prevalence dental disease caused by bacterial infection of the bone that surrounds the tooth. Early detection and precision treatment can prevent more severe symptoms such as tooth loss. Traditionally, periodontal disease is identified and labeled manually by dental professionals. The task requires expertise and extensive experience, and it is highly repetitive and time-consuming. The aim of this study is to explore the application of AI in the field of dental medicine. With the inherent learning capabilities, AI exhibits remarkable proficiency in processing extensive datasets and effectively managing repetitive tasks. This is particularly advantageous in professions demanding extensive experiential knowledge, such as dentistry. By harnessing AI, the potential arises to amplify process efficiency and velocity. In this study, bitewing radiographs are used as the image source, and there are two major steps to detect the dental symptoms including 1) tooth position identification; and 2) symptom identification. The study combines image enhancement techniques and tooth position identification using Gaussian filtering and adaptive binarization for data preprocessing, facilitated by the YOLOv4 model to precisely mark tooth positions. The subsequent step enhances symptom area visibility via contrast enhancement, utilizing a CNN model, particularly the AlexNet model, with significant improvements in caries recognition accuracy (92.85%) and restorations recognition accuracy (96.55%) compared to prior research. Moreover, the inclusion of periodontal disease symptoms achieves an accuracy of 91.13%. By harnessing deep learning techniques based on CNN models, this research enhances diagnostic precision, reduces errors, and increases efficiency for dentists, thereby providing meticulous and swift patient care. This innovation not only saves time but also has the potential for widespread implementation in remote and preventive medicine, aligning with the aspiration of universal health care accessibility.

INDEX TERMS Biomedical image, bitewing radiographic, contrast limited adaptive histogram equalization, tooth segmentation, tooth position, YOLO, CNN.

I. INTRODUCTION

Oral illnesses are the most prevalent of 300 common diseases, according to the WHO's worldwide oral health status

The associate editor coordinating the review of this manuscript and approving it for publication was Ines Domingues.

report. Over 3.5 billion people worldwide suffer from oral diseases [1]. One of the most prevalent dental illnesses, dental caries affects almost 2 billion individuals worldwide. The severe periodontal disease affects 1 billion people, accounting for about 30% of total population. Therefore, dental caries and periodontal disease are the major issues in dentistry nowadays. Periodontal disease is a severe dental disease usually caused by bacterial infection in the periodontal tissues. Poor oral hygiene habits and smoking are the main causes. Common symptoms of periodontal disease include bleeding or gingivitis, bad breath, sensitive teeth, gum recession, and severe periodontal disease can lead to tooth loss. The first step in diagnosing periodontal disease is to visually inspect the color and shape of the gums to check if there is redness or inflammation. Next, a periodontal probe is used to detect the depth of the periodontal pocket to assess the extent of periodontal damage. An X-ray examination is performed to assess the shape of the bone surrounding the teeth, diagnose whether the tooth gap is normal, and confirm the presence of gum recession. The process of identifying symptoms is quite time-consuming. In recent years, AI has drawn attention in solving these problems. Some researchers have utilized deep learning technology in the field of dental medicine to assist dentists in quickly determining the presence of dental disease. Most previous studies used panoramic X-ray for diagnosis, with fewer researchers utilizing bitewing radiographs as the image source. According to a market research, the demand for periodontal disease surgery services is rapidly increasing at a Compound Annual Growth Rate (CAGR) of 11%. The market size is projected to reach a value of $6.6 billion [2]. Early exploration of this market is expected to exert a significant impact on the future development of medicine.

At the 5th World Health Assembly in 2005, the World Health Organization (WHO) adopted Resolution WHA28 [3], which urged member states to develop various types of eHealth-related services. In the same year, the WHO established the Global Observatory for eHealth (GOe), which is studying the development and impact of eHealth on countries around the world. Research has found that eHealth is rapidly changing medical services and systems worldwide, particularly in underdeveloped and developing countries.

Artificial Intelligence (AI) is widely used in the medical field, including mechanical arm in precision surgery to aid doctors in performing high-precision operations. There is also object detection software used to diagnose symptoms and most symptoms can be identified through AI learning. Additionally, intelligent medical auxiliary equipment that assists in mobility is also quite promising. Real-time, large-scale calculations using AI can simplify the operation of auxiliary equipment. Object Detection is particularly common and is mainly applied in diagnosing symptoms because AI can perform repetitive administrative tasks or diagnostic tasks that demand extensive experience. In radiology, ophthalmology, dentistry, and other fields, there exists a significant amount of medical image data that requires manual processing. For example, in ophthalmology, Fundus Fluorescein Angiography (FFA) is used to locate bleeding and ascertain the presence of diabetic retinopathy [4]. Cardiology uses electrocardiogram simulations [5] to predict potential causes of sudden death and can initially identify high-risk individuals for doctors to focus their attention on them. In neurology, AI can automatically analyze the MRI image data [6], label, and record the location, size, and number of brain tumors. In dentistry, X-rays can be used to confirm periodontal conditions and detect diseases such as tooth decay and periodontitis. Utilizing deep learning technology in the field of medicine has the potential to address the issue of inadequate doctor availability, decrease diagnosis time, and enhance treatment efficiency.

Over the recent years, researches of using deep learning for image recognition have been growing rapidly This advancement holds the potential to aid dentists in diagnosing symptoms during clinical practice. AI models can be used to detect tooth position, periodontal disease, dental caries, and restorations. Such developments can reduce the time on diagnosis and enable dentists to focus more on complex conditions or treatments. Tooth position identification is one of the most fundamental works in the dental field. Tooth detection using panoramic X-ray images as the source of the model successfully located the teeth, including missing teeth, with an accuracy of 99.7% [7]. The results demonstrate that using CNN models to locate tooth positions stands out as a remarkably efficient approach. For another tooth localization method using YOLOv3, image enhancement techniques were utilized, reaching 95.58% and 94.90% for precision and recall [8]. This indicates the beneficial impact of this approach on tooth positioning and suggests its viability for continued refinement. Moreover, the YOLO model demonstrates a high degree of accuracy in object detection. The research of dental caries detection [9] combined an iPhone7 camera with YOLOv3, RetinaNet, and other models to construct a caries detection system during the treatment of dental caries. However, the accuracy is compromised by the low image quality. The research also noted that the performance of YOLO models is excellent and worth developing in the future. The other study [10] found that the neural network built upon deep learning technologies has similar performance to experienced dental experts, which can be used for treatment decisions and symptom diagnosis.

In the field of dental restorations, the study of dental restorations detection [11] noted that errors often occur when teeth are mispositioned, but the model based on artificial intelligence demonstrates strong ability to locate restorations. In the research of detecting different restorations, the study [12] investigated how to distinguish restorations composed of different materials. The accuracy of detecting restorations with amalgam material was 0.82, composite materials were 0.75, and metal ceramics were 0.73. The study demonstrated that using bitewing radiographs to develop a CNN model based on deep learning is a promising technique.

In previous research on periodontal disease, the study [13] used panoramic X-rays to detect period bone loss and achieved an accuracy of 84% in diagnosing periodontal disease. However, the accuracy of diagnosis is affected by the imaging characteristics of panoramic X-rays, primarily due to the inherent deficiency in image clarity. Using different X-rays could potentially obtain a better result on diagnosis of periodontal disease. In the research of detecting periodontitis by CNN model [14], the accuracy of using ResNet was only 77.12%, indicating that ResNet's performance is suboptimal and a more proficient CNN model should be suggested.

In summary, this study used the YOLOv4 model to crop images of individual teeth. Previous studies find that manually cropping images of teeth is a complex process prone to exceptions. Therefore, marking the position of teeth and training the model to learn automatically can improve the accuracy of tooth position identification. In tooth position identification, image processing techniques such as Gaussian filtering and adaptive binarization algorithms can make bitewings radiographs more easily recognizable to the model, significantly reducing the processing time of tooth localization. This study also proposed a technique for using bitewings radiographs to identify symptoms of periodontal disease based on a CNN model, providing dentists with faster and more accurate reference information to assist with diagnosis. After testing multiple models, this study chose to use AlexNet as a basis model and modified related parameters to achieve the desired results. In addition, contrast enhancement techniques were used to make the symptoms more distinct, effectively improving the accuracy of detecting caries and dental restorations compared to the previous studies. The contributions of this study are as follows:

1. This study spearheads the incorporation of Convolutional Neural Network (CNN) models in automating periodontal disease identification in dentistry, simplifying diagnostics, and enhancing precision.
2. By employing cutting-edge image preprocessing techniques, Gaussian filtering, and adaptive binarization coupled with YOLOV4, our research achieves a remarkable 98.01% accuracy in tooth position recognition in dental radiographs, concurrently reducing processing time by 61.2%. This substantial improvement establishes a robust foundation for AI-driven diagnostic tools.
3. Progressing beyond traditional diagnostic approaches, our comprehensive framework integrates periodontal disease symptoms using CNN-based deep learning techniques. Comparative analysis demonstrates enhanced disease identification accuracy, with Caries, Periodontitis, and Restorations reaching 92.86%, 92.10%, and 96.51% respectively. This represents a noteworthy improvement of 2.5% to 7% over existing methods, signifying a significant stride in diagnostic precision.

Following the introduction of this study, the second section introduces the method of using YOLO model with bitewing radiographs to predict tooth positions and the materials required for the prediction of caries, restorations, and periodontal diseases by using convolutional neural network (CNN). After introducing the theory and the materials used, the third section exhibits the results of analyzing the experimental data. In the fourth section, the findings of this study and the evaluation are discussed. The final segment concludes by presenting the findings and outlook for the future.

II. METHOD

The procedures of the proposed method can be primarily divided into four parts as shown in Figure 1: (1) Image processing for bitewing images, (2) Tooth detection and cropping using one-stage object detection method, (3) Image processing for the cropped images, and (4) Identification of dental caries, periodontal disease, and dental restorations using CNN-based pattern recognition method. In addition, these procedures are also shown in Figure 2 in the way of pseudocode.

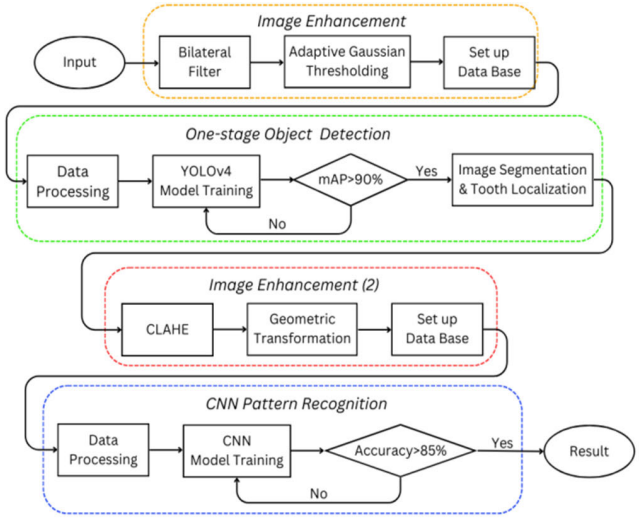


FIGURE 1. The flowchart for the procedure of methods in this study.

The primary aims of this research are to integrate object identification with pattern recognition, as well as to improve pixel-level methods. Image processing techniques are employed as supplementary steps to enhance the performance of detection and recognition. This study presents results of disease recognition using four CNN models and compares them with results of other studies that utilize different image processing techniques.

A. IMAGE ENHANCEMENT

This step of image processing includes noise reduction and binarization of the bitewing images. After the processing, the contours of teeth in the image will be highlighted, which are the regions of interest teeth that need to be localized during tooth detection. The processed bitewing images then can be used for model training.

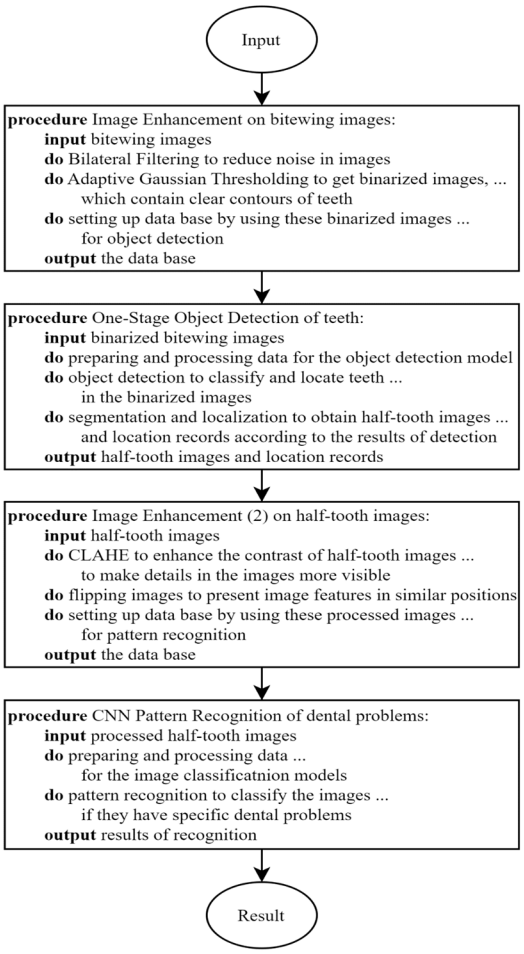


FIGURE 2. The pseudocode for the procedure of methods in this study.

1) BILATERAL FILTER

Low-pass filtering is commonly used to achieve blurring and noise reduction effects. Common low-pass filtering methods include Gaussian filtering and bilateral filtering [15]. The study applies weighted filtering based on the spatial distances between pixels, while the latter is a weighted filtering method that takes into account both the spatial distances and the intensity similarity between pixels, allowing for blurring, denoising, while preserving edge features of the image at the same time.

Therefore, applying bilateral filtering before image binarization can result in a cleaner binary image because the filtering suppresses excessive noise from being converted to small black or white dots while preserving edge features. Bilateral filtering can be expressed as Equation (2) with the weights from Equation (1), where I(x, y) is the pixel that has undergone the noise reduction process.

$$w(x,y,i,j) = exp\left(-\frac{(x-i)^2+(y-j)^2}{2\sigma_s^2} - \frac{I(x,y)-I(i,j)^2}{2\sigma_r^2}\right) \quad (1)$$

$$I(x,y) = \frac{\sum_{i,j} I(i,j)\, w(x,y,i,j)}{\sum_{i,j} w(x,y,i,j)} \quad (2)$$

The parameters of bilateral filtering include the size of the kernel, as well as the spatial parameter and range parameter. The kernel size represents the range of neighboring pixels considered during filtering. The spatial parameter affects the weight of pixels based on their spatial distance. A higher value gives greater weights to pixels with larger spatial distance differences, resulting in a stronger blurring effect. Similarly, the range parameter affects the weight of pixels based on their intensity similarity. A higher value gives greater weights to pixels with larger intensity similarity differences, resulting in a greater blurring effect on edges. Consequently, the filtering effect becomes closer to that of Gaussian filtering.

2) ADAPTIVE GAUSSIAN THRESHOLDING

The binarization method converts an image into a binary image with only black and white colors. Global thresholding employs a single threshold value that is applied to the whole picture. On the other hand, further considering the varying brightness of the image, the adaptive thresholding method is able to apply different thresholds based on the pixels in different regions.

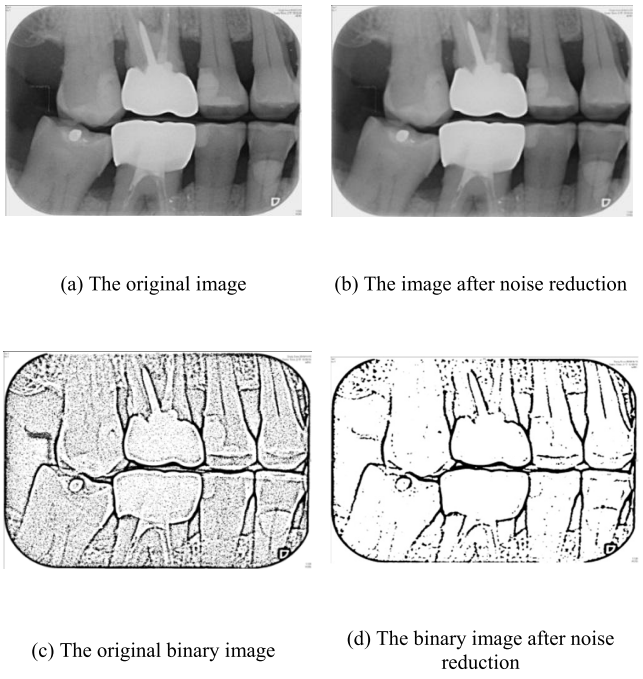
In this study, the adaptive binarization method was used not only for the purpose of obtaining dental contours, but also for the purpose of considering the non-original objects overlaid on the bitewing radiographs, and using global thresholds such as Otsu's global threshold method [16] may introduce bias and affect the accuracy of the binarization. The adaptive thresholding method is able to adapt local features to mitigate the interference of external objects and improve the accuracy of binarization.

In this study, the adaptive Gaussian thresholding method [17] is used, which considers the spatial distance between pixels like Gaussian filtering and applies weighted thresholds with weights from Equation (3). As shown in Equations (4) and (5), each pixel is binarized based on its corresponding threshold value. There are several parameters that need to be set, including the kernel size, standard deviation, and a constant value of C. While the constant value C is a manual parameter used to alter the threshold value and is normally a positive integer, its function is analogous to that of the kernel size and standard deviation as indicated in the filtering stage. Following this step, an example of picture preparation is shown in Figure 3.

$$g(x,y,i,j) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{(x-i)^2+(y-j)^2}{2\sigma^2}\right) \quad (3)$$

$$T(x,y) = \frac{\sum_{i,j} I(i,j)\, g(x,y,i,j)}{\sum_{i,j} g(x,y,i,j)} - C \quad (4)$$

$$I(x,y) = \{255, I(x,y) > T(x,y)\, 0, otherwise \quad (5)$$



(a) The original image (b) The image after noise reduction

(c) The original binary image (d) The binary image after noise reduction

FIGURE 3. Results of the image processing (1).

B. ONE-STAGE OBJECT DETECTION

Object detection methods can directly locate and classify teeth. There are two basic methods for object detection in the present technology: one-stage object detection and two-stage object detection. Region proposal and object recognition are the first two phases of a two-stage object detection process. Common models include the R-CNN series. In this approach, the region proposal step generates potential bounding box proposals, and then the object recognition step classifies these proposals. As opposed to two-stage approaches, one-stage object detection employs a single deep neural network to concurrently identify and categorize items, which speeds up inference. However, in terms of detection accuracy, one-stage methods may not necessarily surpass two-stage methods. Common models include the YOLO series.

Considering the goal of achieving real-time performance, this study adopts the YOLOv4 architecture [18] in the one-stage object detection method. This model replaces pixel-level algorithms and performs the task of detecting teeth. In this study, the training steps for the first-stage object detection model is similar to the pattern recognition model. Both involve data pre-processing, model constructing and training.

1) DATA PREPROCESSING

In this study, the data preprocessing phase encompasses activities such as dataset allocation, image resizing, and data augmentation. In this study, the images are first randomly assigned to the training set, validation set, and test sets in a ratio of 60%, 20%, and 20%, respectively. Additionally, before training the model, the data order is shuffled randomly. Next, the bitewing images (along with the marked bounding boxes used for training) are resized and padded according to the input size of the model. Finally, the images in the training set are randomly subjected to data augmentation techniques including vertical and horizontal flipping and adding mosaic effects.

2) YOLOv4 MODEL TRAINING

As described in [18], with the One-Stage Detector portion shown in Figure 4, this study built a YOLOv4 model with 416 × 416 input size, CSPDarknet53 [19] as the backbone, SPP [20] and PAN [21] as neck, and YOLO [22] as the head. As stated in Table 1, the hyperparameters were as follows: the batch size was set to 4, the quantity of training epochs was set to 400, and the Adam algorithm [23] was employed as the model optimizer with learning rates of 0.001 and 0.9. The effectiveness was observed, and various parameters were adjusted accordingly.

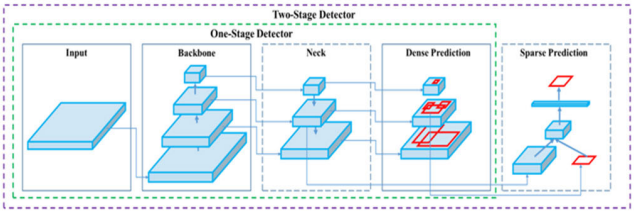


FIGURE 4. Object detection architecture [18].

TABLE 1. YOLOv4 model parameters.

Backbone	CSPDarknet53
Neck	SPP, PAN
Head	YOLO
Input Size	416×416×3
Optimizer	Adam
Learning Rate	0.001
Momentum	0.9
Batch Size	4
Epoch	400

3) IMAGE SEGMENTATION AND TOOTH LOCALIZATION

Since the half side of the tooth is often used as the unit for examining bitewings in clinical practice, the detected teeth are cropped according to the localization boundary, vertically divided into halves, and distinguished into upper and lower, left and right positions based on the relative X and Y-axis coordinates. The cropped images are then numbered in sequence and automatically built into a medical image database.

C. IMAGE ENHANCEMENT (2)

This step of image processing includes contrast enhancement of the half-tooth images and geometric transformation processing. After the processing, the symptoms in the images

can be emphasized, allowing the CNN model to learn these features correctly and improve the accuracy of illness recognition.

1) CLAHE

Based on the experience, caries can cause a darkened tooth surface and a discontinuous tooth margin, and periodontal disease may present with darkened and sunken gums. If the visibility of these symptoms can be improved through image enhancement, it may be helpful in identifying small dental abnormalities in the image.

Contrast enhancement is an image enhancement method that enhances the visibility of an image to display the fine details by the way of histogram processing. It is usually achieved by the histogram equalization (HE) algorithm, which linearizes the cumulative distribution function (CDF) of the histogram and redistributes histogram bins on the histogram to effectively utilize the brightness range, thereby achieving the above goals. HE can be expressed as Equation (7), the Equation (6) is CDF of a pixel value I, and the p(I) is the probability of a pixel value I, which is within the range of 0 to 255.

$$cdf\ (I) = \sum_{j=0}^{I} p\ (I) \tag{6}$$

$$h\ (I) = floor(\frac{cdf\ (I) - cdf_{min}}{cdf_{max} - cdf_{min}} \times 255 + 0.5) \tag{7}$$

The adaptive histogram equalization (AHE) algorithm individually calculates the HE function based on a specific range for each pixel. The algorithm also uses bilinear interpolation to calculate the transformation values for pixels located at the image edges. This approach ensures that each local area of the image is sufficiently enhanced, avoiding the overexposure or underexposure effects that may be caused by global histogram equalization. Contrast Limited Adaptive Histogram Equalization (CLAHE) inherits the benefits of AHE and adds a contrast limiting to lessen the likelihood of amplifying sounds, according to the source [24]. By setting a contrast limit value, the excess pixels are redistributed equally among the other bins of the histogram if there are histogram bins with counts exceeding the clipping limit in a grid, ensuring that the final histogram has equal pixel counts.

The CLAHE algorithm process is: (1) Select a target pixel, and consider the neighboring region of the target pixel as a sub-image based on kernel size with the target pixel at its center. (2) Choose a clipping limit, and redistribute the portion of histogram bins in the image histogram that exceeds the clipping limit to other bars (repeat step 2 until no excess pixels are clipped). (3) Perform histogram equalization on the sub-image (i.e., compute Equation (7)). (4) Transform the pixel value of the target pixel (repeat steps 1 to 4 for each pixel). (5) Use bilinear interpolation to calculate the transformation values for pixels located at the image edges.

In this study, an 8 × 8 kernel size was used and the clipping limit was set to 4. The half-tooth image's contrast was improved using the CLAHE technique to bring out the details. A sample of photos applied to CLAHE is shown in Figure 5.

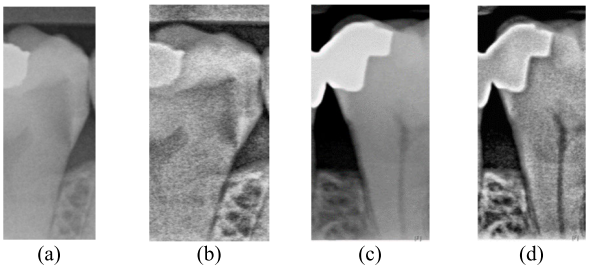


FIGURE 5. A caries half-tooth image (a) after applied CLAHE (b) and a periodontitis half-tooth image (c) after applied CLAHE (d).

2) GEOMETRIC TRANSFORMATION

The four varieties of half-tooth pictures in different orientations are upper-left, upper-right, lower-left, and lower-right when using half-tooth images as the unit. Since the relative position of teeth in the upper-lower and left-right directions is known, this study uses the right part of half-tooth facing upwards as the reference orientation and horizontally or vertically flips all images to present the image features in similar positions.

D. CNN PATTERN RECOGNITION

Deep Neural Networks (DNN), Recurrent Neural Networks (RNN), and Convolutional Neural Networks (CNN) are three of the most widely used artificial neural network designs in deep learning. These networks are used for pattern identification, speech recognition, natural language processing, and other tasks. Among these, CNN structures were chosen for this investigation because they have demonstrated exceptional performance in pattern recognition applications.

Image recognition is usually regarded as a classification problem, and the output of the CNN model is the predicted probability of the image belonging to each category, with the sum of the predicted values equaling 1. Therefore, an image can only be classified into one category. However, the three characteristics of dental caries, restorations, and periodontal disease are not mutually exclusive, which means that a tooth may have multiple features simultaneously and cannot be directly processed by a multi-classifier model. Therefore, it is necessary to establish three CNN models in the form of binary classifiers to independently recognize each disease feature.

1) DATA PREPROCESSING

To obtain better recognition, data preprocessing is employed. Image resizing is performed to ensure that the image size matches the input shape of the model. The image is scaled proportionally and then padded with pixels to match the specified size of the model input shape. Typically, the input shape of the model is smaller than the original image. Properly scaling down the image and input shape, without causing

significant loss of features, can help the model converge correctly.

Considering that three separate CNN models are used to individually identify the three dental features, the system is simplified to only distinguish between the presence and absence of a feature each time. Therefore, a label of 0 or 1 will be assigned to each half-tooth image using Ordinal Encoding (also known as Label Encoding), which maps N categories to decimal value labels ranging from 0 to N-1. Next, the images and labels were used to create a dataset by taking an average sample from each class, and the samples were divided into a training set of 60%, a validation set of 20%, and a testing set of 20% based on the percentage of the total number of samples. This means that 80% of the data is used for model training, with a 3:1 ratio of training to validation data, and 20% of the data is used for final evaluation. The ratio of the dataset can be adjusted according to the number of samples, and the ratio of the validation and testing datasets can be appropriately reduced as the number of samples increases. Due to the limited number of original data samples in this study, a certain level of testing set needs to be maintained to test whether the model has generalization ability. Additionally, random image rotation, contrast adjustment and brightness adjustment were made to augment the image data in this study. The augmented dataset contains four times the amount of data present in the preceding dataset. Moreover, the data augmentation process was only applied to the training dataset to improve the generalization capacity of the model and avoid overfitting. By increasing the quantity of original samples, data augmentation can also address the issue of insufficient picture sample size, enhancing the model's training efficiency.

According to the records provided by the dentists, the number of half-tooth images which were classified and used in this study is shown in Table 2. For single symptom recognition, the amount of images with and without symptoms was controlled to be the same, which also means the number of total used images will be two times of the number of images that contain certain symptoms as shown in Table 2. Table 3 provides the amount of images used in each dataset. The data in the training set is expanded by the data augmentation process, as shown in Table 3 as well.

TABLE 2. Quantity of half-tooth images with each dental symptom.

Caries	Periodontitis	Restorations	Total
105	192	647	944

Before training the CNN model with the datasets, the order of the data is shuffled to avoid an uneven distribution of data from different classes, which may lead to the model learning irrelevant features. Additionally, the pixel values of all images

TABLE 3. Quantity of used half-tooth images in each of datasets.

Dental problem / Dataset	Caries		Periodontitis		Restorations	
	True	False	True	False	True	False
Training (60%)	63	63	115	115	388	388
Validation (20%)	21	21	39	39	130	130
Testing (20%)	21	21	38	38	129	129
Total (100%)	105	105	192	192	647	647
Training (augmented)	252	252	460	460	1552	1552

are normalized to be within the range of 0 to 1, improving the computational efficiency and accuracy of the model.

2) CNN MODEL TRAINING

Considering that the number of samples and image size, overly sophisticated models may not perform as expected. Therefore, in this study, we referred to [25] and attempted to use the AlexNet model structure, as shown in Figure 6.

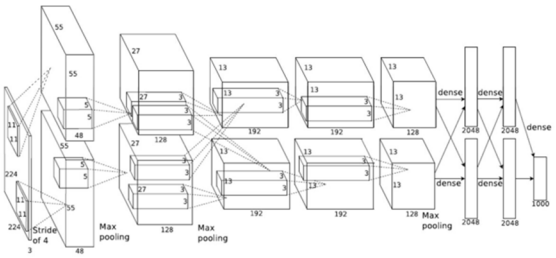


FIGURE 6. AlexNet model structure [25].

The model and hyperparameters shown in Table 4 are set for this study with the hardware and software configuration shown in Table 5. The input port specification of the model is set to 200 × 100 (vertical length x horizontal width), slightly smaller than the original image size, to reduce its size without losing too many features, allowing the model to converge correctly.

TABLE 4. AlexNet model parameters.

Input Size	200×100×3
Optimizer	Adamax
Learning Rate	0.001
Momentum	0.99
Batch Size	64
Epoch	200
Patience	50

Images and ports can use either a single channel or three channels, and in theory, the training results of the two models are the same, because converting a grayscale image to an RGB image means that the pixel value of a single channel is transformed into three repeated sub-pixel values on the same pixel grid. This may only affect the amount of computation

TABLE 5. Hardware and software applied in this study.

Hardware Devices	
CPU	Intel® Core™ i7-11800H
GPU	NVIDIA® GeForce RTX™ 3060
Memory	GDDR6 (6GB), DDR4 (40G)
Software Applications	
OS	Windows 11 Pro
Python	Ver. 3.8.16
VSC	Ver. 1.79.0

and hardware space, and there is no specific notation in this study.

The following hyperparameters are set: The batch size is 64 and is changed based on the hardware load. Its purpose is to divide the dataset into batches for inputting into the model, reducing hardware load and avoiding exceeding hardware computing and access capacity from inputting data at once. The number of the training cycle (epoch) is set to 200, which is sufficient to ensure that the model's accuracy and loss values tend towards saturation. With a learning rate of 0.001 and momentum of 0.99, the model optimizer employs the Adamax algorithm [23], a variation of the Adam algorithm based on the infinite boundary range. The activation function is the Softmax function, which is suitable for both binary and multiclass classifier structure models.

After completing the data preprocessing process and setting the necessary parameters, the model can be trained by inputting the training and validation sets and executing the training and self-validation processes. The best metric of the model's success is the accuracy attained when using the trained model to predict the data from the test set. By observing the model's fitting curve and various evaluation metrics, we can use a trial and error approach to iteratively adjust the parameters, train the model, and test its performance.

III. RESULTS

A. TOOTH DETECTION RESULTS

A confusion matrix is used to examine a model's performance and assess the outcomes. The confusion matrix, shown in Table 6, assesses the YOLOv4 model's mean average precision (mAP), precision, recall, and F1-score. Equations (8)-(10) outline the calculation methods for these indicators. Figure 7 shows the training process of the YOLOv4 loss function. Additionally, Table 7 displays the comparison of three different YOLO models.

TABLE 6. Confusion matrix.

	Predicted Positive	Predicted Negative
Actual Positive	TP (True Positive)	FN (False Negative)
Actual Negative	FP (False Positive)	TN (True Negative)

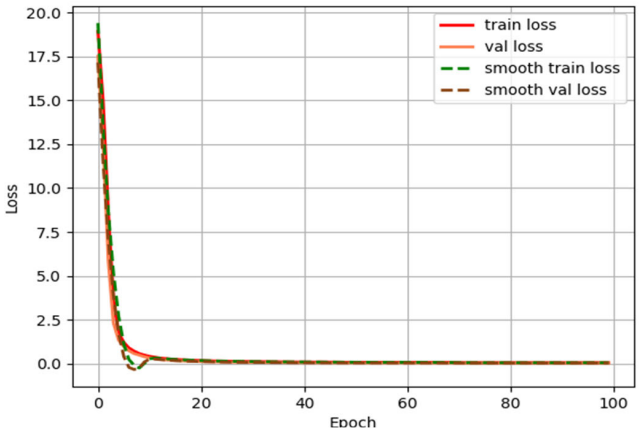


FIGURE 7. Training Process for the YOLOv4 loss function.

TABLE 7. Comparison of different YOLO models.

Epoch = 100	YOLOv3 [26]	YOLOv4	YOLOv5 [27]
mAP	82.77%	99.38%	99.29%
Precision	96.91%	98.01%	94.57%
Recall	82.32%	98.43%	99.43%
F1-score	89%	98%	99%
Processing Time	0.145s	0.134s	0.160s
Data Size	246.5MB	256.3MB	28.3MB

$$\text{Accuracy}: \frac{(TP+TN)}{(TP+TN+FP+FN)} \tag{8}$$

$$\text{Precision}: \frac{TP}{(TP+FP)} \tag{9}$$

$$\text{Recall}: \frac{TP}{(TP+FN)} \tag{10}$$

$$\text{F1-score}: \frac{2\times(Precision\times Recall)}{(Precision+Recall)} \tag{11}$$

In this study, the model was trained by a total of 210 untrained images, which was split into a 9:1 ratio for training and validation sets. Figure 8 shows that YOLOv4 performed better results when the training epoch was set to 100. However, YOLOv5 has the smallest data size and still maintains a promising result as YOLOv4. In addition, to ensure the reliability of our model and mitigate bias towards particular training or testing data, we used 10-fold cross-validation. Tables 7 and 8 illustrate the consistent performance of the YOLOv4, indicating minimal bias in its results. The original and enhancement photos used by YOLOv4 are demonstrated in Figure 9.

According to Table 9, it can be observed that the model using image enhancement techniques performs compara-

TABLE 8. YOLOv4 10-fold cross-validation.

	mAP	Precision	Recall	F1 Score
1	99.25%	97.07%	99.29%	98%
2	99.16%	97.74%	98.71%	98%
3	99.12%	96.93%	99.14%	98%
4	99.32%	97.86%	98%	98%
5	99.23%	97.32%	98.57%	98%
6	99.36%	96.91%	98.71%	98%
7	99.41%	97.85%	97.57%	98%
8	99.17%	98%	98.14%	98%
9	99.24%	97.43%	97.57%	98%
10	99.24%	97.74%	99%	98%
Average	99.25%	97.485%	98.47%	98%

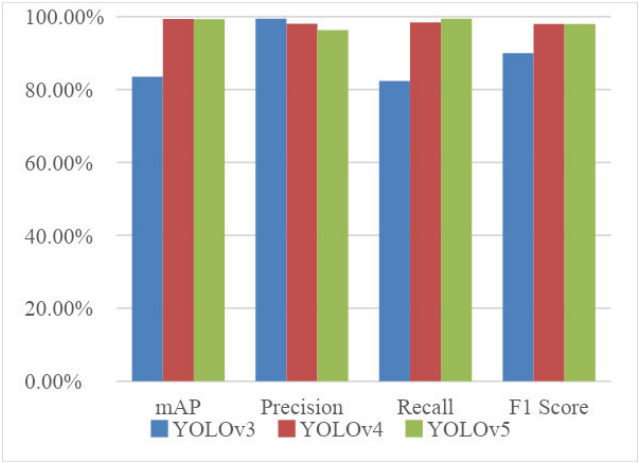


FIGURE 8. Comparison of different YOLO model

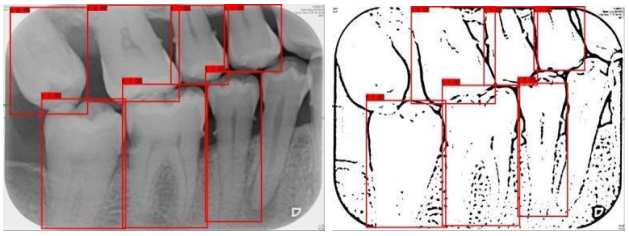


FIGURE 9. Recognition results through YOLOv4 (a) Original Image. (b)Enhancement Image.

ble performance to using original images across various indicators. However, as in Figure 10, the processing time for predicting a single unknown image is significantly reduced by over 61%, demonstrating that enhancement image improves processing efficiency. Additionally, compared to existing studies [26] and [28], the model shows an improvement of 2.85∼9% in precision, 0.95∼16.11% in recall, and 1.1∼5.03% in F1-score in Table 10. As shown in Figure 11, the results show the advantages of using the YOLOv4 model

TABLE 9. Original image compared to enhanced image.

Epoch = 100	Original	Enhanced
mAP	99.38%	99.10%
Precision	98.01%	97.44%
Recall	98.43%	97.86%
F1-score	98%	98%
Processing Time	0.134s	0.052s

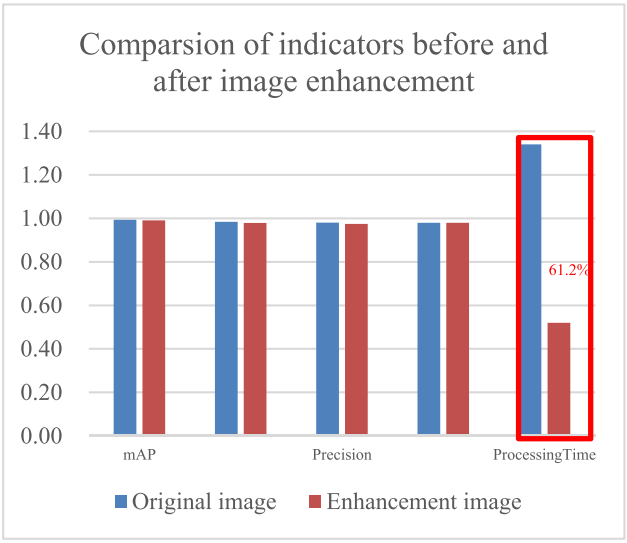


FIGURE 10. A comparison of indications between the original image and the improved version.

TABLE 10. Comparison of the previous and current studies.

Epoch = 100	Method in [28]	Method in [26]	This Study
Precision	92.93%	96.91%	98.01%
Recall	97.48%	82.32%	98.43%
F1-score	95.15%	89%	98%

for tooth position identification. Furthermore, the enhancement images not only maintain the model performance but also enhance model stability by using variations in the training data, ensuring that the model can handle different dental conditions. These advancements contribute to more reliable and effective tooth position identification, serving as a valuable reference for dentists.

After determining the tooth positions, this study successfully obtained individual tooth images and recorded the quantity of images for the upper and lower jaws. The numbers of tooth images are listed as in Table 11.

B. SYMPTOM RECOGNITION RESULTS

After model testing, the study obtained confusion matrices, which are presented in the following tables (Table 12, Table 13, and Table 14). The percentages in the

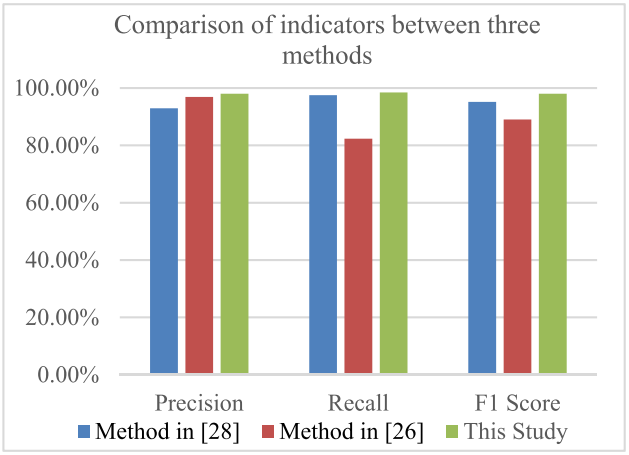


FIGURE 11. Comparison of indicators between three methods.

TABLE 11. Quantity of tooth images sliced from bitewing radiographs.

Upper jaw	Lower jaw	Total
1086	1028	2114

TABLE 12. Confusion matrix of caries recognition.

Actual\Predicted	True	False
True	45.24% (19)	4.76% (2)
False	2.38% (1)	47.62% (20)

TABLE 13. Confusion matrix of periodontitis recognition.

Actual\Predicted	True	False
True	46.05% (35)	3.95% (3)
False	3.95% (3)	46.05% (35)

TABLE 14. Confusion matrix of restorations recognition.

Actual\Predicted	True	False
True	47.67% (123)	2.33% (6)
False	1.16% (3)	48.84% (126)

TABLE 15. Results of AlexNet models of the dental problem.

	Caries	Periodontitis	Restorations
Accuracy	92.85%	92.10%	96.51%
Precision	95.00%	92.10%	97.61%
Recall	90.47%	92.10%	95.34%
F1-score	92.67%	92.10%	96.46%

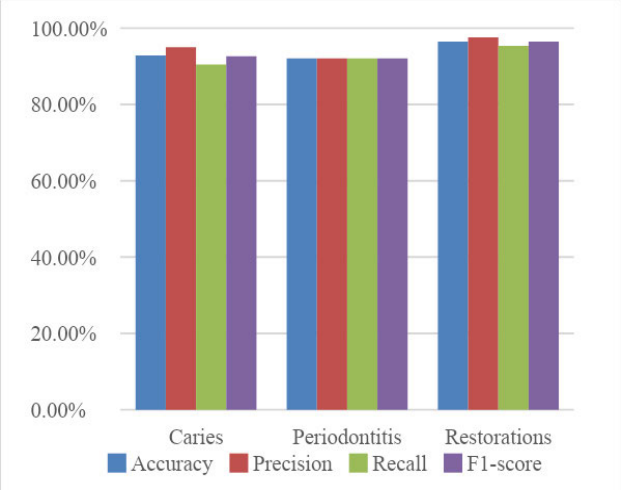


FIGURE 12. Results of AlexNet models of the dental problem.

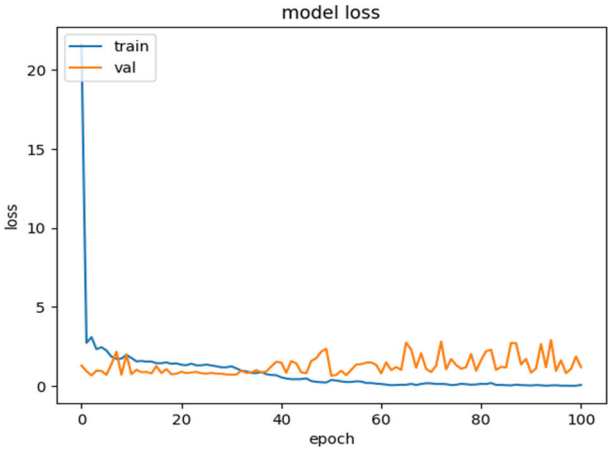


FIGURE 13. Training process of AlexNet loss function for caries recognition.

tables represent the percentage of data belonging to that category out of the total data, and the numbers in parentheses represent the number of images. Moreover, according to these tables, the final training results of AlexNet models of the dental problem can be calculated, as shown in Table 15 and Figure 12. The results of 5-fold cross-validation for each of the AlexNet models are provided in Table 16 as well. In general, the average results are 2% to 4% lower than the best results.

Additionally, Figure 13, Figure 14, and Figure 15 depict the variations in the loss function during the training process for recognizing the three different symptoms using AlexNet. Table 17, Table 18, and Table 19 provide various model training results for the three symptoms using four different models: AlexNet, ResNet50 [30], ResNet101 [30], and EfficientNetV2B0 [31]. These results indicate that as more complex models are used, the accuracy of illness recognition tends to decrease, but the accuracy of restoration recognition tends to increase. Therefore, depending on the available data quantity, choosing different models or more complex models may have the potential to achieve better results.

TABLE 16. AlexNet models 5-fold cross-validation.

		Accuracy	Precision	Recall	F1-score
Caries	1	88.10%	99.99%	76.19%	86.48%
	2	92.85%	95.00%	90.47%	92.67%
	3	90.47%	94.73%	85.71%	89.99%
	4	90.47%	90.47%	90.47%	90.47%
	5	88.10%	86.36%	90.47%	88.36%
	Avg.	89.99%	93.31%	86.66%	89.59%
Periodontitis	1	92.10%	92.10%	92.10%	92.10%
	2	85.89%	83.33%	89.74%	86.41%
	3	88.15%	85.36%	92.10%	88.60%
	4	89.74%	94.28%	84.61%	89.18%
	5	88.15%	89.18%	86.84%	87.99%
	Avg.	88.80%	88.85%	89.07%	88.85%
Restorations	1	94.18%	96.72%	91.47%	94.02%
	2	93.41%	93.07%	93.79%	93.42%
	3	95.38%	94.02%	96.92%	95.44%
	4	96.51%	97.61%	95.34%	96.46%
	5	92.69%	91.72%	93.84%	92.76%
	Avg.	94.43%	94.62%	94.27%	94.42%

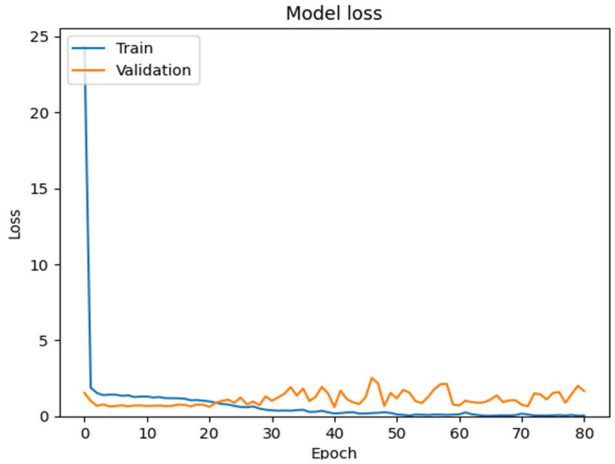


FIGURE 14. Training process of AlexNet loss function for periodontitis recognition.

In this research, a comparative analysis of model performance in recognizing dental restorations, caries, and periodontal diseases is presented, as depicted in Figure 13 and summarized in Table 20. Surprisingly, the accuracy of the AlexNet model was the greatest throughout both the testing and training stages, although accuracy levels were lower when more complex models were used. Notably, ResNet101 demonstrated superior performance in restoration recognition, with a minor decrease in accuracy when adopting the more complex EfficientNetV2B0 model.

The application of image enhancement techniques yielded an approximate 5% improvement in recognition accuracy

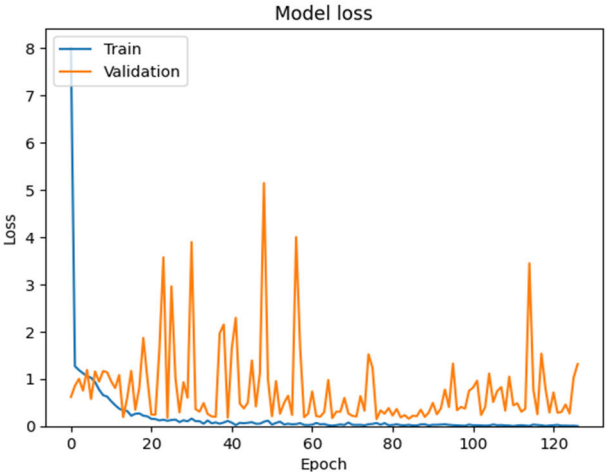


FIGURE 15. Training process of AlexNet loss function for restorations recognition.

TABLE 17. Result of caries recognition with different models.

	AlexNet	ResNet50	ResNet101	EfficientNet V2B0
Accuracy	92.85%	90.47%	76.19%	71.42%
Loss	0.222	0.503	0.515	0.512
Batch size	64	4	4	4
Epoch	200	200	200	200
Patience	50	50	50	50
Processing time (sec/epoch)	0.980	7.393	12.501	7.837

compared to the use of raw images, emphasizing the effectiveness of contrast enhancement in facilitating symptom recognition. Furthermore, as seen in Figure 13 and described in Table 21, our suggested technique in this work beats state-of-the-art approaches [13] and [29] in both restoration and caries recognition by a margin of 1% to 2.5%. Most notably, the accuracy improvement for the challenging task of periodontal disease recognition reaches an impressive 8.1%. This represents a substantial and impactful achievement, as our method effectively enhances and surmounts the technical intricacies associated with periodontal disease identification. It marks a significant advancement in the integration of dentistry and artificial intelligence. According to the statistical results in Table 22, the identification results of this study showed a strong positive correlation with the physician identification results, with $r = 0.770$, $p < 0.001$; $r = 0.906$, $p < 0.001$; $r = 0.878$, $p < 0.001$, respectively.

IV. DISCUSSION

The model in this study achieved an accuracy of 92.85% in detecting cavities and 96.51% in detecting restorations as shown in Table 19 and Figure 13. Advancing beyond conventional methods, disease identification accuracy for Caries, Periodontitis, and Restorations improved by 2.5%

TABLE 18. Result of periodontitis recognition with different models.

	AlexNet	ResNet50	ResNet101	EfficientNet V2B0
Accuracy	92.10%	86.84%	84.21%	81.58%
Loss	0.379	0.438	0.373	0.617
Batch size	64	4	4	4
Epoch	200	200	200	200
Patience	50	50	50	50
Processing time (sec/epoch)	1.286	13.352	22.087	12.683

TABLE 19. Result of restorations recognition with different models.

	AlexNet	ResNet50	ResNet101	EfficientNet V2B0
Accuracy	96.51%	96.90%	98.06%	97.26%
Loss	0.237	0.147	0.135	0.141
Batch size	64	4	4	4
Epoch	200	200	200	200
Patience	50	50	50	50
Processing time (sec/epoch)	3.645	43.879	79.271	43.811

TABLE 20. Recognition accuracy of the features with different models.

	AlexNet	ResNet50	ResNet101	EfficientNet V2B0
Caries	92.85%	90.47%	76.19%	71.42%
Periodontitis	92.10%	86.84%	84.21%	81.58%
Restorations	96.51%	96.90%	98.06%	97.26%

TABLE 21. Comparisons of accuracy based on AlexNet model.

	Original data	Method in [13]	Method in [29]	Method in this study
Caries	85.71%		90.30%	92.86%
Periodontitis	86.84%	84%		92.10%
Restorations	93.80%		95.56%	96.51%

to 7%, signifying a substantial stride in diagnostic precision and transformative potential for dentistry. Moreover, this study introduces the YOLO model for tooth cropping with a remarkable accuracy rate of 99.38% and a 60% reduction in processing time using enhancement images. The research also pioneers Convolutional Neural Network models, achieving a remarkable 98.01% accuracy in tooth position recognition with a 61.2% reduction in processing time. This has increased our confidence in using the model as a reference. However, there is still a chance for improvement to make the model more valuable in clinical settings. First, more training data is needed to increase the model's experience and performance. Second, a user-friendly interface should be developed to facilitate the utilization by dentists and other users. Finally, hardware devices such as X-ray machines should be integrated into the workflow, allowing the model

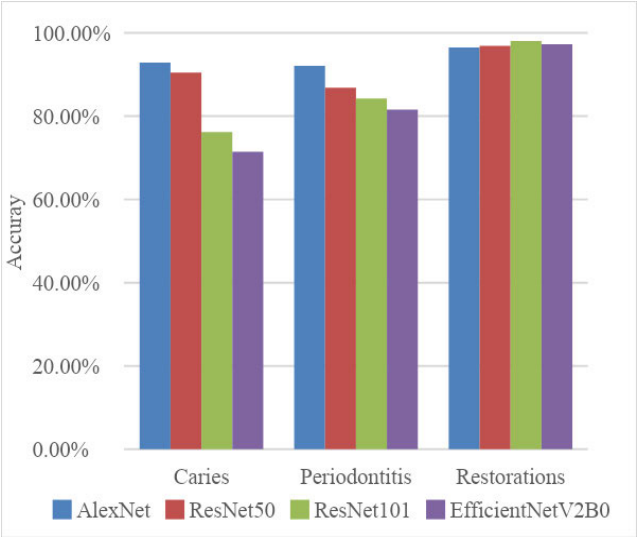


FIGURE 16. Recognition accuracy of the features with different models.

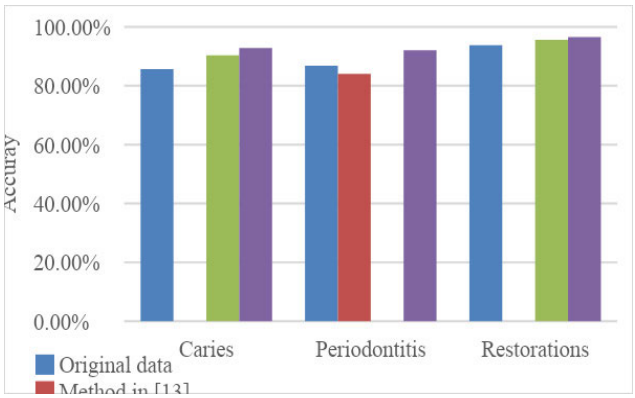


FIGURE 17. Comparisons of accuracy based on AlexNet model.

TABLE 22. Paired samples correlations between AI identification results and clinician results.

	Quantity	Correlations	Significance
Periodontitis	30	0.770	0.0000006417
Restorations	30	0.906	0.0000000000
Restorations	30	0.878	0.0000000002

to automatically label the images and evolve over time. This integration could potentially lead to even higher accuracy.

V. CONCLUSION

This study improved the complex tooth positioning methods used in previous research by using the YOLOv4 model to automatically recognize the position of teeth. Additionally, a new model was added to detect periodontal disease symptoms by using bitewing radiographs, providing more technological advancements in dentistry. Furthermore, contrast enhancement was used as an image processing method, leading to further improvement in the accuracy of the exist-

ing models for detecting cavities and restorations. Based on the experimental results, the proposed workflow was found to be correct and effective. It was also discovered that the more obvious the symptom features were in the image data, the higher the accuracy of the diagnosis. Therefore, future research should focus on making the features of the symptoms more distinct. Additionally, since traditional CNN models suffer from the vanishing gradient problem, it is expected that using models such as YOLO, which reduce computation and improve learning efficiency, or more complex models such as ResNet and EfficientNet, could lead to higher accuracy. The goal is to reach the clinical standards required for practical use by dentists and provide substantial assistance.

REFERENCES

[1] (2022). *The Global Status Report on Oral Health*. World Health Organization. [Online]. Available: www.who.int/team/noncommunicable-diseases/global-status-report-on-oral-health-2022

[2] *Fact.MR—Periodontal Dental Services Market Analysis, by Procedure (Surgical Periodontal Dental Services, Non-Surgical Periodontal Dental Services), by End-Use Industry (Periodontal Dental Services at Hospitals, Dental Clinics) & by Region, Global Insights to 2031*. Market Research Company. Accessed: Jan. 19, 2024. [Online]. Available: https://www.factmr.com/report/periodontal-dental-services-market

[3] World Health Assembly. (2005). *Fifty-Eighth World Health Assembly*. World Health Organization, Geneva, Switzerland. [Online]. Available: https://apps.who.int/iris/handle/10665/20398

[4] S. Wang, Y. Zuo, N. Wang, and B. Tong, "Fundus fluorescence angiography in diagnosing diabetic retinopathy," *Pakistan J. Med. Sci.*, vol. 33, no. 6, p. 1328, Nov. 2017.

[5] K. C. Siontis, P. A. Noseworthy, Z. I. Attia, and P. A. Friedman, "Artificial intelligence-enhanced electrocardiography in cardiovascular disease management," *Nature Rev. Cardiol.*, vol. 18, no. 7, pp. 465–478, Jul. 2021.

[6] A. Aleid, K. Alhussaini, R. Alanazi, M. Altwaimi, O. Altwijri, and A. S. Saad, "Artificial intelligence approach for early detection of brain tumors using MRI images," *Appl. Sci.*, vol. 13, no. 6, p. 3808, Mar. 2023.

[7] M. Chung, J. Lee, S. Park, M. Lee, C. E. Lee, J. Lee, and Y.-G. Shin, "Individual tooth detection and identification from dental panoramic X-ray images via point-wise localization and distance regularization," *Artif. Intell. Med.*, vol. 111, Jan. 2021, Art. no. 101996.

[8] T. H. Bui, K. Hamamoto, and M. P. Paing, "Tooth localization using YOLOv3 for dental diagnosis on panoramic radiographs," *IEEJ Trans. Electron., Inf. Syst.*, vol. 142, no. 5, pp. 557–562, 2022.

[9] M. T. G. Thanh, N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap, and D. M. Nguyen, "Deep learning application in dental caries detection using intraoral photos taken by smartphones," *Appl. Sci.*, vol. 12, no. 11, p. 5504, May 2022.

[10] L. Lian, T. Zhu, F. Zhu, and H. Zhu, "Deep learning for caries detection and classification," *Diagnostics*, vol. 11, no. 9, p. 1672, Sep. 2021.

[11] P. Engels, O. Meyer, J. Schönewolf, A. Schlickenrieder, R. Hickel, M. Hesenius, V. Gruhn, and J. Kühnisch, "Automated detection of posterior restorations in permanent teeth using artificial intelligence on intraoral photographs," *J. Dentistry*, vol. 121, Jun. 2022, Art. no. 104124.

[12] O. Karatas, N. N. Cakir, S. S. Ozsariyildiz, H. C. Kis, S. Demirbuga, and C. A. Gurgan, "A deep learning approach to dental restoration classification from bitewing and periapical radiographs," *Quintessence Int.*, vol. 52, no. 7, pp. 568–574, Jun. 2021.

[13] J. Ryu, D.-M. Lee, Y.-H. Jung, O. Kwon, S. Park, J. Hwang, and J.-Y. Lee, "Automated detection of periodontal bone loss using deep learning and panoramic radiographs: A convolutional neural network approach," *Appl. Sci.*, vol. 13, no. 9, p. 5261, Apr. 2023.

[14] D. M. Alalharith, H. M. Alharthi, W. M. Alghamdi, Y. M. Alsenbel, N. Aslam, I. U. Khan, S. Y. Shahin, S. Dianisková, M. S. Alhareky, and K. K. Barouch, "A deep learning-based approach for the detection of early signs of gingivitis in orthodontic patients using faster region-based convolutional neural networks," *Int. J. Environ. Res. Public Health*, vol. 17, no. 22, p. 8447, Nov. 2020, doi: 10.3390/ijerph17228447.

[15] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. 6th Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

[16] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.

[17] N. Khetan, L. Kejriwal, and S. Indu, "Enhancement of degraded manuscript images using adaptive Gaussian thresholding," *Int. J. Future Gener. Commun. Netw.*, vol. 10, no. 1, pp. 47–60, Jan. 2017.

[18] A. Bochkovskiy, C.-Y. Wang, and H.-Y. Mark Liao, "YOLOv4: Optimal speed and accuracy of object detection," 2020, *arXiv:2004.10934*.

[19] C.-Y. Wang, H.-Y. Mark Liao, I.-H. Yeh, Y.-H. Wu, P.-Y. Chen, and J.-W. Hsieh, "CSPNet: A new backbone that can enhance learning capability of CNN," 2019, *arXiv:1911.11929*.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Computer Vision—ECCV*. Berlin, Germany: Springer, 2014, pp. 346–361.

[21] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," 2018, *arXiv:1803.01534*.

[22] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[23] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.

[24] K. Zuiderveld, *Contrast Limited Adaptive Histogram Equalization*. vol. 5. Cambridge, MA, USA: Academic Press, 1994, ch. 5, pp. 474–485.

[25] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017.

[26] Y.-C. Chen, M.-Y. Chen, T.-Y. Chen, M.-L. Chan, Y.-Y. Huang, Y.-L. Liu, P.-T. Lee, G.-J. Lin, T.-F. Li, C.-A. Chen, S.-L. Chen, K.-C. Li, and P. A. R. Abu, "Improving dental implant outcomes: CNN-based system accurately measures degree of peri-implantitis damage on periapical film," *Bioengineering*, vol. 10, no. 6, p. 640, May 2023.

[27] C.-C. Chen, Y.-F. Wu, L. M. Aung, J. C.-Y. Lin, S. T. Ngo, J.-N. Su, Y.-M. Lin, and W.-J. Chang, "Automatic recognition of teeth and periodontal bone loss measurement in digital radiographs using deep-learning artificial intelligence," *J. Dental Sci.*, vol. 18, no. 3, pp. 1301–1309, Jul. 2023.

[28] Y. Yasa, Ö. Çelik, I. S. Bayrakdar, A. Pekince, K. Orhan, S. Akarsu, S. Atasoy, E. Bilgir, A. Odabas, and A. F. Aslan, "An artificial intelligence proposal to automatic teeth detection and numbering in dental bite-wing radiographs," *Acta Odontologica Scandinavica*, vol. 79, no. 4, pp. 275–281, Nov. 2020.

[29] C.-W. Li, S.-Y. Lin, H.-S. Chou, T.-Y. Chen, Y.-A. Chen, S.-Y. Liu, Y.-L. Liu, C.-A. Chen, Y.-C. Huang, S.-L. Chen, Y.-C. Mao, P. A. R. Abu, W.-Y. Chiang, and W.-S. Lo, "Detection of dental apical lesions using CNNs on periapical radiograph," *Sensors*, vol. 21, no. 21, p. 7049, Oct. 2021.

[30] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2015, *arXiv:1512.03385*.

[31] M. Tan and Q. V. Le, "EfficientNetV2: Smaller models and faster training," 2021, *arXiv:2104.00298*.

KUO-CHEN LI received the Ph.D. degree in computer science and engineering from the University of Louisville. He is currently an Associate Professor with the Department of Information Management, Chung Yuan Christian University, Taiwan. His research interests include machine learning, mobile commerce, and data analysis.

YI-CHENG MAO was born in Taipei, Taiwan, in 1990. He received the bachelor's degree in dentistry from Chung Shan Medical University, Taichung, Taiwan, in 2015. He was a resident, who has operative dentistry specialist training with the Department of General Dentistry, Chang Gang Memorial Hospital, Taoyuan, Taiwan, since 2016. He received the Operative Dentistry Specialist Certificate from the Taiwan Academy of Operative Dentistry, in 2020. He has been an Attending Doctor of general dentistry with the Taoyuan Chang Gang Memorial Hospital, since 2021.

MU-FENG LIN is with Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan. His current research interests include image processing, machine learning, and bio-medical signal processing.

YI-QIAN LI has been with Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, since 2020. His current research interests include image processing, machine learning, and bio-medical signal processing.

CHIUNG-AN CHEN received the B.S. degree in electronic engineering from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, in 2005, and the Ph.D. degree in electrical engineering from the National Cheng Kung University, Tainan, Taiwan, in 2013. Since 2017, she has been an Assistant Professor with the Department of Electrical Engineering, Ming Chi University of Technology, Taiwan. Her current research interests include VLSI chip design, image processing, wireless body sensor networks, the Internet of Things, wearable devices, biomedical signal, and image processing.

TSUNG-YI CHEN (Member, IEEE) received the B.S. and Ph.D. degrees from Chung Yuan Christian University, Zhongli, Taoyuan, Taiwan, in 2020 and 2023, respectively, all in electrical engineering. Since 2023, he has been an Assistant Professor with the Department of Electrical Engineering, Feng Chia University, Taiwan. His current research interests include VLSI chip design, image processing, the Internet of Things, fuzzy logic control, wearable devices, machine learning, and biomedical signal processing.

PATRICIA ANGELA R. ABU (Member, IEEE) received the B.S. degree in electronics and communications engineering from Ateneo de Manila University, Philippines, in 2007, the M.S. degree in electronics engineering (major in microelectronics) from Chung Yuan Christian University, Zhongli, Taiwan, in 2009, and the Ph.D. degree in computer science from Ateneo de Manila University, in 2012. She is the Research Laboratory Head of the Ateneo Laboratory for Intelligent Visual Environment (ALIVE) and an Assistant Professor with the Department of Information Systems and Computer Science (DISCS), Ateneo de Manila University. Her current research focuses on image processing and computer vision, with applications that revolve on biomedical imaging and road/transportation, anomaly detection, and IoT systems, which granted ALIVE several best research paper and presentation awards both local and abroad.