Tooth Numbering and Condition Recognition on Dental Panoramic Radiograph Images Using CNNs

SZU-YIN LIN[ID], **(Member, IEEE), AND HAO-YUN CHANG**[ID]

Department of Computer Science and Information Engineering, National Ilan University, Yilan City, Yilan County 260, Taiwan

Corresponding author: Szu-Yin Lin (szuyin@niu.edu.tw)

ABSTRACT Dentists and medical personnel strive to provide patients with prompt medical services. In the past, Dental Panoramic Radiograph (DPR) was often used to diagnose and understand the dental condition of patients. In recent years, many machine learning and deep learning methods have been applied to medical image recognition problems. Moreover, when combined with deep learning methods, data augmentation and image pre-processing methods can also give positive feedback. This study aims to combine data augmentation and data pre-processing methods with advanced deep learning methods to build an innovative and practical two-phase DPR recognition and classification method to assist dentists in diagnosis. It will help to improve the medical quality of dental services by speeding up and saving valuable physician manpower cost and time. Prior to the two-phase recognition based on several effective Convolutional Neural Networks (CNNs), the data augmentation and data pre-processing are processed. In the first phase of this method, the position and numbering of the tooth is automatically classified as one of 32 tooth positions from the DPR tooth images. In the second phase, the dental conditions are automatically recognized from the 6 dental conditions, including orthodontics, endodontic therapy, dental restoration, impaction, implant, and dental prosthesis. The experimental results showed that the trained network, without image pre-processing and augmentation, identified the dental position numbering with an accuracy of 90.93%, and the dental condition with an accuracy of 93.33%. After data augmentation, the accuracy of tooth numbering can be increased to 95.62%, and the accuracy of dental condition can be increased to 98.33%. This is a significant improvement when compared with past research.

INDEX TERMS Medical image processing, panoramic X-ray image, tooth numbering and condition recognition, deep learning, convolutional neural networks.

I. INTRODUCTION

Dental Panoramic Radiograph (DPR) is a scanned wide-angle X-ray radiograph taken from the upper and lower jaw section of the patient. DPR was often used by medical personnel as an important reference diagnostic basis for understanding the dental condition of patients. It assisted dentists to provide patients with the most immediate medical services. However, when the number of patients is much larger than the number of medical professionals, the quality of dental treatment decreases. Currently, professional medical personnel interpret and mark the DPR indirectly increasing their workload. If the image recognition method of deep learning can automatically analyze the patient's panoramic X-ray images, it will help to speed up and save valuable physician

manpower cost and time, and improve the medical quality of dental services, especially when determining the content of a large number of panoramic X-ray images. The purpose of this research is to propose an effective method for recognition and classification of tooth position numbering and tooth condition by combining image pre-processing and data augmentation with several advanced deep learning methods, to accurately and effectively deal with the problems of automatic recognition and classification on DPR.

In recent years, deep learning has performed prominently in the field of machine learning and feature learning, and is usually used to solve image recognition [1], [2], speech recognition [3], and time series problems [4]. The neural networks are the basis for abstracting data through multiple non-linear transformations [5]. Due to the recent technological advances, many data scientists have started to use computers to generate algorithms to solve these different

The associate editor coordinating the review of this manuscript and approving it for publication was Ioannis Schizas[ID].

types of data classification and data processing problems. Among them, N.-H. Lin *et al.* [6] conducted a study for complete DPR identification, and used an image pre-processing method for tooth position numbering identification and a convolutional neural network (CNN) method with data augmentation for automated dental condition identification. In that study, the treatment conditions were identified as Normal, Restored, Missing, Prosthesis, Endodontic treated tooth with Prosthesis, and Endodontic treated tooth without Prosthesis. However, as more and more methods have been proposed in recent years, these methods have yielded good results in various areas of DPR identification, in both dental condition and position of the tooth. For example, to identify a dental condition from DPR, a study by Y.-F. Kuo *et al.* [7] used CNN and pre-processing methods to positively identify the presence of a root canal treated tooth. In the study of identifying tooth position, there is also a study by H. Chen *et al.* [8] using Fast-RCNN and three post-processing methods to automatically distinguish the area where the tooth is located and label the tooth position. Although a complete DPR recognition process like the research of N.-H. Lin *et al.* is already available, H. Chen *et al.* and Y.-F. Kuo *et al.*'s research has shown that the use of better deep learning models can get positive feedback in the field of tooth position numbering and dental condition classification. Studies by C.-H. Wu *et al.*, M. Chung *et al.* and I. Aliaga *et al.* are also conducting similar research on DPR identification or segmentation, which not only shows that DPR has gradually become the main indicator for judging dental conditions, it also shows automatic identification of DPR has become the current major trend [9]–[12].

The automatic image recognition of DPR has not only become the current major trend, but also reduces the time for manual inspection of DPR images. The researchers Motoko, K. *et al.* conducted a method that combines a regional convolutional neural network (R-CNN), single shot multibox detector (SSD), and heuristic methods to detect and number the teeth and implants in a DPR image [13]. The research conducted by Kim, C. *et al.* can detect candidate teeth through faster R-CNN, and then determine the appropriate candidate teeth by optimizing the objective function [14]. In addition, the research conducted by Tuzoff, D. *et al.*, which proposed a faster R-CNN teeth numbering classification method according to the FDI notation [15]. These studies have given positive feedback, which shows that it is possible to apply artificial intelligence to DPR. Many deep learning models have been proposed by experts in the field of data science, such as AlexNet, VGGNet, GoogLeNet, Xception, and ResNet. Since the number of layers in these models is much higher than in a CNN, these five models can obtain more image features than CNN. Because ResNet is a network architecture consisting of residual blocks, the gradient disappearance problem is less likely to occur when training ResNet models. ResNet has four times more model layers than that of AlexNet. Therefore, this study proposes to build a two-phase automatic DPR classification based on ResNet with a data pre-processing method and data augmentation

methods. The conventional AlexNet, GoogLeNet, Xception, and VGGNet are used as the comparison objects of this study.

The goal of this study is to propose a two-phase DPR classification system based on several effective Convolutional Neural Networks (CNNs), image pre-processing, and data augmentation for tooth position numbering and condition classification to reduce the workload of dental medical personnel and assist dentists in diagnosis. The first phase is the automatic identification of 32 tooth positions in the DPR. The second phase is the identification of 6 dental conditions in the DPR, including orthodontics, endodontic therapy, dental restoration, impaction, implant, and dental prosthesis.

For the choice of pre-processing methods, this study refers to the study by R. Kaur *et al.* [16], which used the multiple morphological gradient method to remove redundant features in the DPR. This study uses the edge detection method to remove the redundant image features from DPR, and then inputs the processed DPR into a deep learning model for training to verify whether the deep learning model can get the correct image features. Additionally, data augmentation methods will be incorporated into the model training to verify whether the model can absorb richer picture information from the increased DPRs. This study demonstrates that automated classification of tooth condition classification and tooth position numbering classification will be effective in real-world situations. The first step of the experimental procedure is to process the data using data pre-processing and data augmentation methods. The second step is to input the training data into a deep learning model and train the model. Finally, the testing data are input into the model to obtain the classification results of dental position numbering and condition.

The structure of this study is as follows: The second part introduces the background of DPR, e.g., the dental condition and location of the tooth, and previous studies on DPR. In the third part, the two-phase DPR experimental procedure design and methodology are introduced. The fourth part introduces the evaluation methods of the model, the experimental results, and the comparison of the implementation methods. Finally, the fifth part presents the conclusions and future perspectives.

II. RELATED WORKS
A. DENTAL PANORAMIC RADIOGRAPH (DPR) UTILIZATIONS

A DPR film shows a panoramic view of the entire upper and lower dental arch and the temporomandibular joint. The DPR film provides a two-dimensional radiograph used for dental analysis. DPR can provide healthcare professionals crucial information (including but not limited to) to indicate the patients' dental condition such as shown in TABLE 1. For example, in the case of endodontic therapy, the feature to be looked for is usually just the shape of the filled root canals that would appear as a brighter area on the DPR. However,

TABLE 1. Identifiable dental conditions from the DPR.

| Dental condition	Explanation
Caries	A progressive destruction of teeth due to activities of bacteria, also known as tooth decay or cavities.
Periodontitis	Inflammation of supporting structures of the teeth and especially the periodontal membrane.
Apical Lesion	Inflammatory lesions that are most commonly caused by microbial invasion or unsuccessful endodontic therapy
Dental Prosthesis	An intraoral prosthesis used to restore defect conditions such as missing dental parts.
Dental Restoration	Also known as dental filling, it is a procedure to restore a tooth damaged by decay back to its normal function and shape by using filling materials.
Missing Tooth	This is the process in which one or more teeth come loose and fall out, or are surgically removed.
Impaction	It refers to the blocking of a tooth by a physical barrier, such as a neighboring tooth, causing an inability to erupt.
Retained Root	Retained root refers to the partial root structure, which remains after the extraction or fracture of the tooth itself.
Implant	It refers to the surgical components that interface with the oral structure to support dental prosthesis (e.g. denture).
Endodontic Therapy	It refers to the treatment of the infected pulp and root of a tooth in sequences such as: eliminating infection, tooth restoration, and protecting it from future microbial invasion through the use of dental filling.
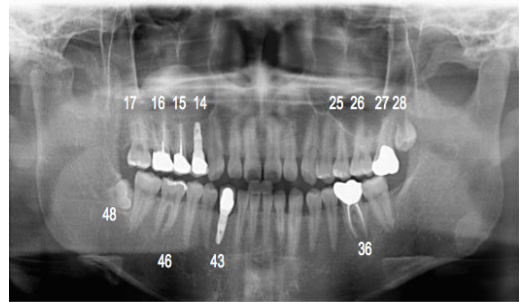
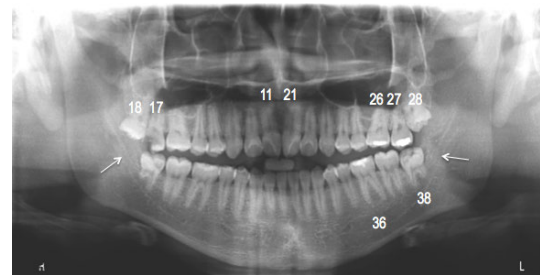

FIGURE 1. Dental panoramic radiograph 1.

FIGURE 2. Dental panoramic radiograph 2.

- Malposition tooth: 11 and 21.
- Restoration: 24, 25, 26, 27 and 46.
- Inferior alveolar canals are indicated by the two arrows above.

in the other cases such as implants or restoration, the shape and size of the feature may differ greatly due to the variances of different types of implants or the degree of restoration for the damaged tooth.

Dental treatment is the way of repairing and artificially filling tooth damage. In general, during the dental treatment, the artificial substance used in the oral panoramic will show a higher brightness value, so it can be identified by this feature. Fig. 1 and Fig. 2 illustrate DPR which provides dental information to dentists, which are marked according to the Fédération Dentaire Internationale (FDI) tooth notation system. FDI notation system is a commonly used system for the numbering and naming of teeth. DPR has become a more popular X-ray examination technique in dentistry, due to its simple execution, relatively low radiation dose, better patient acceptance, and short operation time [16].

Fig. 1 shows the following dental information from panoramic X-ray image:

- Dental implant: 14 and 43.
- Dental crown prosthesis: 16, 15, 14, 27, 36 and 43.
- Impacted tooth: 28 and 48.
- Endodontic treated tooth: 16, 15 and 36.
- Restoration: 17, 46, and 25.

Fig. 2 shows the following information from DPR:

- Dental caries: 17 and 38.
- Impacted tooth: 18 and 28.

B. DEEP LEARNING

Deep learning belongs to a branch of machine learning which is formed by sets of algorithms. The main idea of deep learning would be trying to model data that has highly abstracted information by reducing the data dimensions with multiple processing layers. Many of the deep learning convolutional neural network (CNN) architectures such as AlexNet, VGGNet, GoogLeNet, ResNet, and Xception have been applied to solve computer vision, audio classification, or natural language and numerical processing tasks while yielding state-of-the-art results.

Table 2 shows the comparison of some classic and recent popular deep learning CNN networks with their number of parameters and their accuracy over the ImageNet dataset [17]. The top-1 and top-5 accuracy refers to the model's performance on the ImageNet validation dataset [18]. In this study, we use top-1 accuracy to determine the result, which means

TABLE 2. Deep learning model comparison.

Model	Salient Feature	Top-1 Accuracy	Top-5 Accuracy	Parameters
AlexNet	Deeper	0.633	0.846	62M
VGG16	Fixed-size kernels	0.744	0.919	138M
GoogLeNet	Wider-parallel kernels	0.748	0.922	23M
ResNet	Shortcut connections	0.786	0.943	60M
Xception	Depthwise separable convolution	0.790	0.945	22M

that model prediction (the one with the highest probability) must be exactly the expected answer.

1) ResNet

In 2016, the research team of K. He *et al.* published a study on ResNet [19], because this deep learning model introduced an architecture called residual block. Therefore, compared to other network architectures that do not use residual blocks, the training effect of a network architecture composed of multiple residual blocks is more efficient than other network models. Among them, in 2015, the Kaimei He research team won many championship awards through the ResNet model in competitions such as ILSVRC and COCO. In the ILSVRC competition, the error rate of ResNet was lower than 5%. It was even better than humans in recognizing pictures. Nowadays, there are many models based on ResNet to improve the model, like ResNet-18 and ResNet-34. It can be said that ResNet is the most representative of the deep learning model. Taking the research of [20] as an example, the team used ResNet and object detection to detect and prevent the occurrence of oral cancer, and get positive feedback from it.

2) STRUCTURE OF ResNet

The typical structure of ResNet consists of thirty-three convolutional, pooling layers and one fully connected layers total of thirty-four learning layers as Fig. 3.
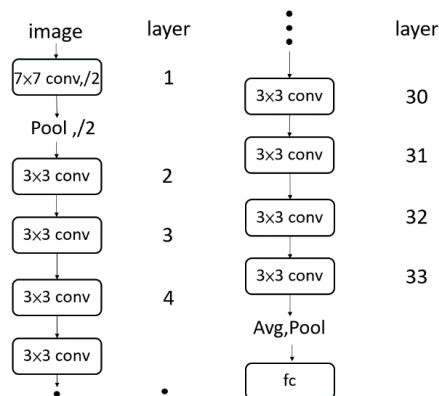

FIGURE 3. ResNet structure.

The convolutional layers consist of rectangular grids of neuron which require the previous layer to be also rectangular grids of neurons. The weights of each neuron at the same convolutional layer would be the same, and these neurons would receive inputs from the previous layer. These weights are specified by the convolution filter, which is where the term of convolutional neural network comes from, because each hidden layer is basically a mathematical convolution of the previous layer but with the possibility of using different filters. The convolutional layer is followed by the pooling layer, which takes small rectangular blocks from the convolutional layer to generate a single output. There are many ways of doing the pooling such as calculating the average of the area

or taking the maximum of that area [21]. The fully connected layer can be found at the end of the network model, and followed by a Softmax layer that generates a sum-normalized distribution. The fully connected layer takes all the neurons in the previous layer, and there will be no convolutional layer after a fully connected layer. Following the fully connected layer, the Softmax layer consist the function of a gradient-log-normalizer of the categorical probability distribution, it is often used for probabilistic classification for multi-classes, which in the case of this study would be to predict the condition in the radiograph.

a: RESIDUAL BLOCK

When the number of layers of the network is superimposed, the accuracy of the model is more likely to reach saturation. And when the number of iterations continues to increase, the classification accuracy of the model will decrease instead. Although the characteristics of this phenomenon are similar to overfitting, it is not caused by overfitting. This phenomenon in which the number of network layers is too large, resulting in an increasing error between the training result and the classification result is called gradient vanishing. To solve the problem of gradient vanishing, the research team of K. He *et al.* proposed the skip connection in 2016. Suppose the underlying mapping is H(x) and the identity mapping is x. Then we let the stack of nonlinear layers to fit the residual mapping as F(x) = H(x) - x. So the underlying mapping can be expressed as H(x) = F(x) + x. In such a network model architecture, if the number of network model layers continues to increase, the subsequent residual mapping will be updated to 0. In this way, the classification accuracy of the model is guaranteed not to decrease as the number of network layers' increases [19].

b: ReLU

The ReLU function is Eq. (1), where x represents the input of the neuron node. In terms of the slope of the training time, the use of saturated nonlinear function is much slower than the unsaturated nonlinear function. Therefore, in the practice of the neural network method, there are many functions which can be used to increase the nonlinear characteristics, such as hyperbolic tangent and sigmoid function. However, according to Nair and Hinton, it is pointed out that the use of ReLU, an unsaturated nonlinear neuron, can improve the training speed of the model without reducing the accuracy of the convolution neural network [22].

$$f(x) = \begin{cases} x, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1}$$

c: DROPOUT

When the deep learning model is trained, the learning goal of the model is defined according to the training data, so that when the extracted feature is too close to the training data, it will let the model learn erroneous features or redundant features, indirectly leading to decreased accuracy. The above

problem is called overfitting. On the issue of dealing with overfitting, dropout is the most commonly used method as a solution. Dropout was Hinton's method in 2012, where in each training case, dropout randomly omitted half of the feature detector to prevent the training information being too complex. In the study of Warde *et al.*, dropout is a very effective integrated learning method, which integrates the results of multiple predictors to find the best solution [23].

C. IMAGE PRE-PROCESSING

When using deep learning methods for image classification, most of the cases of low classification accuracy are related to the input data. Usually, the original image data may contain a lot of image information that is not related to the classification items, resulting in the model classification accuracy not as expected. Therefore, before training a model, most data scientists use image pre-processing methods to remove excess image information to improve the classification accuracy. At present, many researchers in the field of computer vision have proposed various image pre-processing methods, which can use linear filtering, median filtering, and edge enhancement filtering in the processing of contours or redundant information. Among them, edge-enhanced filtering is most suitable for highlighting features and removing redundant image information, such as Sobel, Lowpass, Highpass, Laplacian, Sharpen, Spatial and Temporal, etc.

1) SOBEL

Sobel edge detection is an edge-enhanced filtering and was proposed by Irwin Sobel in 1986. The principle of Sobel is to use two 3×3 matrices to convolve the original image to calculate the difference between the horizontal and vertical grayscale values. This is done in order to leave the gradient variation part of the image by convolution. A study by [24] used the Sobel method to solve the problem of automatic cutting of overlapping portions of tooth in DPR. Therefore, it is known from the study of [24] that the Sobel method will be helpful for DPR. The two matrices in the Sobel method can be seen as Eq. (2) and Eq. (3). The matrices of Eq. (2) are used to detect the gradient gx in the x-direction. The matrices of Eq. (3) are used to detect the gradient gy in the y-direction [25].

$$gx = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \tag{2}$$

$$gy = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix} \tag{3}$$

Finally, we use Eq. (4) to square and root the results obtained from the previous two matrixes to obtain the gradient intensity of the images. Therefore, this study intends to use this algorithm to reduce unnecessary image information [25].

$$g = \sqrt{(gx)^2 + (gy)^2} \tag{4}$$

D. DATA AUGMENTATION

In the process of training deep learning models, the overfitting problem often occurs due to data problems. But in addition to dropout, data augmentation methods can also solve the problem of overfitting. Data augmentation solves the problem of overfitting by increasing the diversity of picture data using methods such as randomly panning the picture pixels, randomly enlarging and reducing the size of the picture, or adjusting the color difference of the picture. When these processed data are fed into the model for training, the more data the model absorbs, the lower the probability of overfitting. In addition to solving the overfitting problem, data augmentation can also provide positive feedback in the field of medical research. In the study of [26], if the data-augmented dataset is fed into deep learning for training, the accuracy of the test results can increase from 88.31% to 98.88%. In a study conducted by [27], the cone-beam CT image dataset was trained with data augmentation and input to a deep learning model to increase the test accuracy by up to 5%. As a result of these findings, this study aims to increase the diversity of DPR data using the Keras implementation data augmentation method and expects the classification accuracy of the model to increase accordingly. The data augmentation methods implemented in this study are shown in horizontal displacement and resize image.

1) HORIZONTAL DISPLACEMENT

When shifting the image horizontally, it is necessary to take into account the presence or absence of the identified object in the image. Otherwise, after the image is horizontally shifted, the classification object may be detached from the image. Therefore, the values chosen in this study range from 0.1 to 0.9 to avoid identifying the target objects out of the range of the image display. In the horizontal displacement, using a parameter between 0.1 and 0.9 means that the horizontal displacement distance will be the width of the image multiplied by the specified value (0.1-0.9). For example, if the width of the image is 200 pixels and the specified value is 0.1, the value of the displacement will be 20 (200 ∗ 0.1).

2) RESIZE

When resizing the images by data augmentation, it is also necessary to consider whether the identifiers in the image are present in the range, otherwise the identifiers may also be outside of the image display. Therefore, in order to avoid generating the target objects beyond the range of the image, the range of values chosen in this study is also located between 0.1 and 0.9. If resizing or zooming a float value, then resize will be done in the range [1-value, 1+value]. In the resize, the parameter used in the range of 0.1 to 0.9 means that it is the zoom ratio of resize. If the parameter value is set to 0.3, it means that the zoom ratio of the generated augmented image is between 70% and 130% of the original image.

III. TWO-PHASE DENTAL RECOGNITION METHOD ON DPR

This study presents a two-phase dental panoramic radiograph recognition and classification method based on deep learning. This section discusses the proposed method in detail, along with the definition of variables and the proposed method implementation process. Firstly, the problem is defined. Secondly, the method of data acquisition is described. Then, the model of the two-phase dental panoramic radiograph classification method based on deep learning is introduced in detail. The system architecture is divided into the phases of tooth position numbering and condition classification. Finally, the process, strengths and weaknesses, and research steps of this approach will be discussed.

A. PROBLEM DEFINITION

As mentioned in section 2, DPR classification in actual clinical practices is still being done by trained individuals. There are several limiting constraints when humans perform these tasks, which are described as follows:

(1) The classification task must be done by professionally educated personnel that possess the domain knowledge and professional experience to ensure the result is reliable. Dependency on this human resource is costly.

(2) Classification with human labor is limited to mono-tasking. Although a trained individual would be able to perform multiple classifications at one time, the speed of recognizing and identifying conditions from the radiographs is affected by having to perform the entire task sequentially, thus creating a latent reduction in the performance speed.

(3) The classification result could be subjective due to its heavy reliance on the trained individual's knowledge and experience. It could also be affected by other unstable characteristics of human labor. Eliminating the subjectivity would require more human labor.

As mentioned above, the machine learned classifier, given its natural characteristics, can provide prospective solutions to resolve the problems. A well-constructed classifier can perform multiple tasks in a short period of time with accuracy based on its training data. Its results could provide fast and precise information to the professional individual, who would be using this information to treat the patient.

B. DPR COLLECTION

The dental panoramic radiograph images and labels used in this study are provided by the dental dentist of the Chang Gung Medical Foundation Hospital, and supported by Chang Gung Medical Foundation Institutional Review Board under the IRB NO.201600380B0. The labeled dental panoramic radiographs were marked and identified by two or more professional dentists, while providing the diagnosis of each of the teeth in each dental panoramic radiograph image.
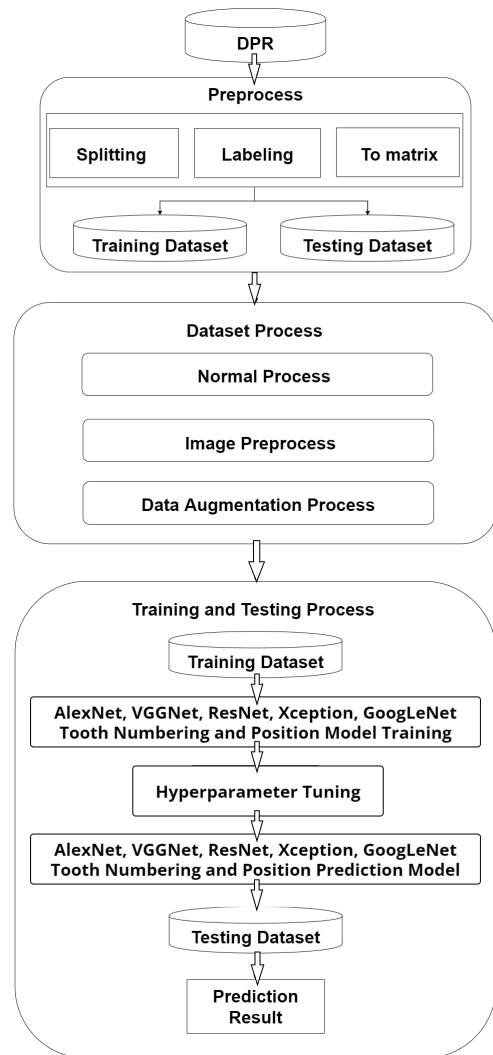

FIGURE 4. Flow chart of the first phase.

C. SYSTEM ARCHITECTURE

In order to solve the problem of dental panoramic radiograph classification described in the first section, this study proposes a two-phase dental panoramic radiograph classification method based on deep learning. The system architecture is divided into two phases: tooth position numbering and tooth condition classification.

1) FIRST PHASE - TOOTH POSITION CLASSIFICATION

In the training process, the first phase can be divided into four parts: (1) DPR pre-process, (2) image dataset pre-process, (3) deep learning classification models, and (4) evaluation of the tooth position classification model. The goal of the tooth position classification is to identify 32 different teeth in the dental panoramic radiograph, and to find the corresponding position of each tooth in the mouth. Fig. 4 is the flow chart of the first phase.

a: PRE-PROCESS

A pre-process procedure on the DPR data is required so it can be successfully input into the system in a consistent form by fixing the size and centering the location of the tooth in the DPR. Performing the pre-processing can provide more organized and clearer information to train the network rather than just using the raw DPR. For each DPR image, each tooth in the mouth is considered an individual input data. In Fig. 5, the patient has 28 teeth. The DPR is cut into individual tooth images according to the number of teeth. As shown in Fig. 6, it is converted into 28 independent tooth images.

Chang Gung Hospital collected and provided a total of 895 oral panorama images from adult patients. We first do the image data cleaning, filter out 136 oral panoramic images with better quality and more representative from 895 panoramic oral images, and label them as the training dataset. At the same time, 10 other oral panoramic images are also selected and labeled as the testing dataset. After the image data is cleaned and labeled, the images will be divided into a training dataset and a testing dataset. The training and testing data set will be used in the two-phase tooth recognition process after pre-processing and cutting. The training dataset is used for the feature learning and training of the model, and the testing dataset is used to verify the accuracy of the model that has been trained.

In this study, the original images refer to the oral panoramic images that have not been cut and pre-processed, so that the size of the original oral panoramic image in Fig. 5 is 2816 ∗ 1540 pixels. The original DPR image must be segmented and pre-processed into the individual single teeth. Therefore, the size of the pre-processed split dental image separated from the original DPR image in Fig. 5 is defined as 227 ∗ 227 pixels. Dental image splitting is a part of image pre-processing, the purpose is to take out each individual tooth in the complete DPR image, and then use various deep learning models for image recognition training and learning. Since various models based on the CNN method use the concept of convolution and feature map to achieve the purpose of image recognition and classification. Therefore, as long as the splitting tooth image in the training data is complete and recognizable, it can be helpful for the training and learning of the model.

In this study, we use an image cutting software tool to cut and split the original DPR images under the advice of medical experts. The principle and standard of image cutting and splitting is to make a complete selection of individual teeth of the target. There may be some other teeth next to each other after the cut image, so there will be overlaps in the cut image. The cut tooth images may have different sizes, but they are all smaller than 227 ∗ 227. The image of the tooth after cutting is shown in Fig. 6. Then fill in white edges around this tooth image, without affecting the size of the teeth, and adjust the image to 227 ∗ 227.

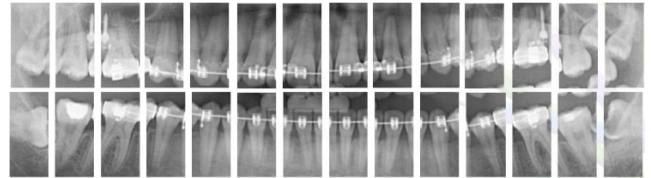**FIGURE 5.** Dental panoramic radiograph.

FIGURE 6. Split dental panoramic radiograph.

b: DATASET PROCESS

In the dataset process section, three types of methods are used to process the data. The processing methods can be divided into (1) normal process, (2) image pre-process, and (3) data augmentation process. In this study, we try to solve the problem of overfitting and other problems that lead to poor test training by these methods. These methods are described and illustrated as follows.

i) NORMAL PROCESS

The normal process uses this dataset as a reference point for comparison with other processing methods. In this step, the data is not processed, so that the suitability of the model structure can be selected and the accuracy of the testing data improved through the use of pre-processing methods.
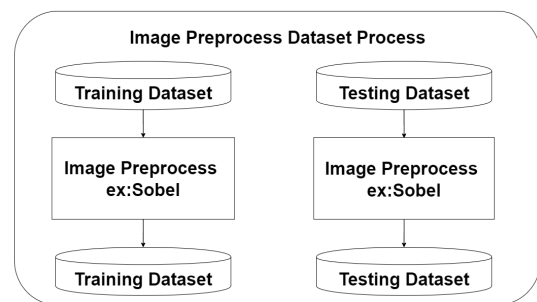

FIGURE 7. The procedure of image pre-process.

ii) IMAGE PRE-PROCESS

In the pre-image processing step, this study uses the Sobel edge detection method to extract the features of the tooth edge contours. The aim is to extract the contours of a single tooth by the Sobel algorithm to remove unnecessary picture information. The processing and distribution of training data

and testing data are shown in Fig. 7. The training data and testing data processed in the previous step need to be processed for image pre-processing. The goal is to obtain classification results similar to the training data after the same processing of the testing data.

iii) DATA AUGMENTATION PROCESS

In total, two data augmentation methods are used in this study. The first method is horizontal displacement and the second method is enlarging or reducing the image. The purpose of using data augmentation methods in this study is to increase the diversity of the training dataset by adding additional data that the machine has not seen. Fig. 8 shows a picture of a single tooth with parameters between 0.1 and 0.9 and after horizontal displacement. Fig. 9 shows a picture of a single tooth with parameters between 0.1 and 0.9 with random zooming in and out. In the processing flow of the data augmentation method, only the training dataset is augmented in this study. The purpose is to allow the model to absorb more information from different pictures, and to increase the adaptability of the model to various kinds of data by this method, so that the model can also be applied to real life. Therefore, in this method, the training data and the testing data are handled as shown in Fig. 10.

c: THE ARCHITECTURE OF DEEP LEARNING MODEL FOR TOOTH POSITION CLASSIFICATION

In recent years, many representative models of deep learning based on CNN, such as AlexNet, VGGNet, GoogLeNet, Xception, and ResNet have been introduced. The best-performing model among these is ResNet [19].
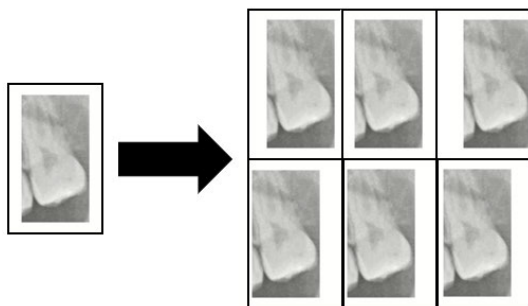

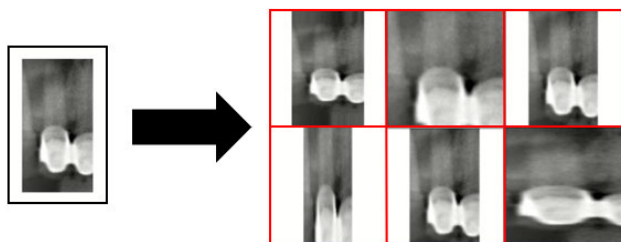**FIGURE 8.** Image of a single tooth after horizontal displacement by data augmentation method.

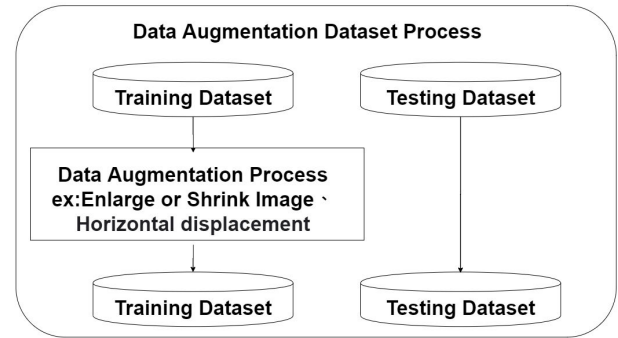**FIGURE 9.** Single tooth image after data augmentation method of resize.

FIGURE 10. The procedure of data augmentation process.

ResNet performs better not only because of the higher number of layers in the model, but also because of the residual block architecture of ResNet. Since this architecture makes the gradient disappearance problem less likely to occur when the number of model layers' increases, it makes ResNet one of the representative deep learning models. A schematic diagram of the residual block is shown in Fig. 11. The residual block contains many convolution layers and batch normalization layers, and it performs even better with activation function-ReLU. This function is added to the model for training in the hope that the model can find the best solution for data classification in a faster and non-linear way.

This study decided to build a dental position classification system based on ResNet. The ResNet network architecture used in this study is shown in Fig. 12. The input layer defines the dimensions of the data input and the type of data input. The output layer defines the output dimension of the model result and the type of the output result. The middle layer between the input layer and the output layer is mainly made up of residual blocks. The middle layer is composed of residual blocks, which also contain convolution layers and pooling layers. The residual block in Fig. 12 can be seen in Fig. 11.

The traditional method of gradient descent is to deal with all of the information at once, but when the amount of data is too large, and the convergence rate is very slow. Therefore, the learning rate in the model training process is a very important parameter. Learning rate directly affects the convergence rate of the network. If the learning rate is
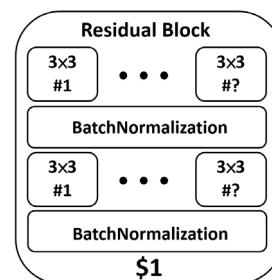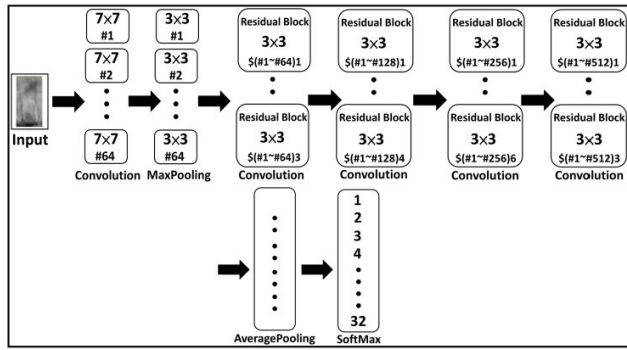

FIGURE 11. Residual block.

FIGURE 12. Deep learning for tooth position.

large, the convergence rate of the network will become faster. On the other hand, the smaller learning rate will slow the convergence rate of the network. In order to optimize the deep learning training process and avoid falling into the slope of the region, the best solution is adding the Adam method into the deep learning model. The Adam method is the combination of RMSprop and Stochastic Gradient Descent (SGD) with momentum. It contains momentum's gradient speed adjustment for the direction of the past gradient and Adam's adjustment for the learning rate of the square value of the past gradient. In this way, Adam is more stable than other optimizers in terms of learning rate updates, and the model is less likely to fall into the problem of regional optimal solutions. Eq. (7) is the Adam weight update equation [28]. The calculation methods of *mt hat* and *vt hat* in Eq. (7) can correspond to Eq. (5) and Eq. (6). *Mt hat* in Eq. (7) stands for momentum, meaning the movement at the previous time point, and *vt hat* stands for velocity, which will be related to the last update. Because Adam uses these parameters for bias correction, the learning rate will have a certain range every time, which will make the parameter update more stable.

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{5}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{6}$$

$$W \leftarrow W - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{7}$$

d: THE EVALUATION OF THE TOOTH POSITION CLASSIFICATION MODEL

In this study, in addition to using ResNet to build the system for the first phase of the experiment, we use AlexNet, VGGNet, and other representative deep learning models to build the system, and use the data processing methods introduced above to process the training data, and then input the training data into the model for training. The model is then evaluated using the testing dataset to obtain the prediction results. The classification results can be divided into two categories: predicting the correct tooth position and predicting the incorrect tooth position.
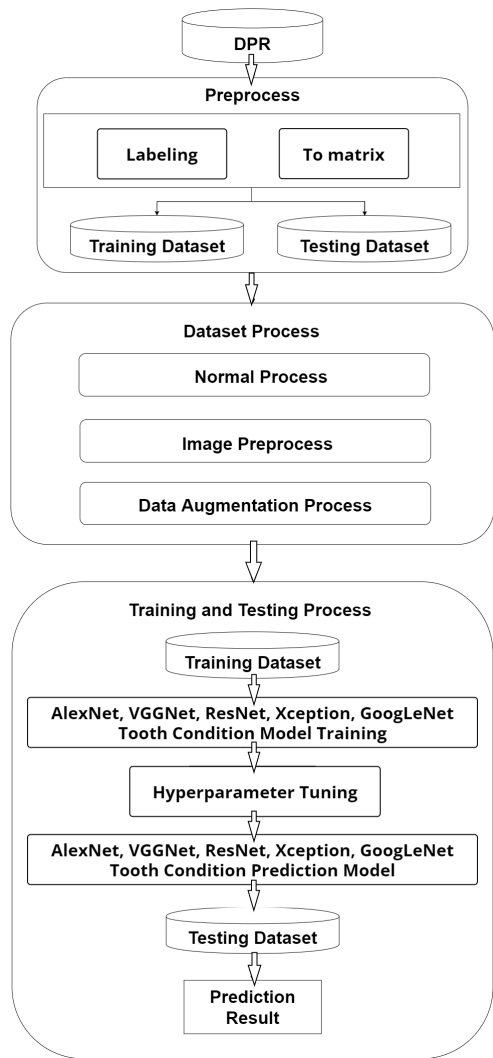

FIGURE 13. Flow chart of the second phase.

2) SECOND PHASE – TOOTH CONDITION CLASSIFICATION

This phase of the training process can be divided into four parts: (1) DPR pre-process, (2) image dataset pre-process, (3) deep learning classification models, and (4) evaluation of the tooth condition classification model. The goal of the dental condition classification is to identify 6 different dental conditions in the dental panoramic radiograph, including orthodontics, endodontic therapy, dental restoration, impaction, implant, and dental prosthesis. Fig. 13 is the flow chart of the second phase.

a: PRE-PROCESS

In order to be able to successfully enter a single tooth image into the system, pre-processing for the image is necessary. First, the data pre-processing fits the input data into a consistent form by fixing the image size and capturing the location of the tooth. Pre-processing the data allows a single dental image to provide a more organized and informative training class of neural networks. The second phase

of data pre-processing consists of two steps: label and matrix. Finally, the pre-processed dental images are divided into the training dataset and the testing dataset. The training dataset is used as the input data for the deep learning model. The testing dataset is used to verify the prediction results of the second phase prediction model.

b: DATASET PROCESS

The steps of the dataset process for phase 2 (dental condition identification), are similar to the steps of phase 1. Dataset process methods can be divided into (1) normal process, (2) image pre-process, and (3) data augmentation process. The processing steps are to process the single subject dental training and testing dataset.

i) NORMAL PROCESS

The data processed by the normal process is the training dataset and the testing dataset. The purpose of this step is to use this dataset as a benchmark for comparison with other methods. This allows us to compare the suitability of deep learning models or other data process methods to improve classification accuracy.

ii) IMAGE PRE-PROCESS

In the field of condition classification, the treated part of the image is whiter than the other parts of the image. Therefore, in this study, we use Sobel to process single tooth images to obtain contour patterns based on this characteristic. After treatment, the image will look like a screw, an oval or other irregular shape of the tooth, etc. Other non-condition related information such as the oral background is also removed by Sobel. It is possible for the model to learn the correct image information.

iii) DATA AUGMENTATION PROCESS

The purpose of using the data augmentation methods for dental condition identification in this study is to reduce the overfitting problem during model training, which leads to poor identification accuracy. In this study, the data augmentation methods used in this step are horizontal displacement and enlarging or reducing the image. As described in the first phase of the dataset process. In order to keep the classification of the target objects in the image range, the parameters set in this study are 0.1 to 0.9. We observe whether the classification accuracy of these values has solved the problem.

c: THE ARCHITECTURE OF DEEP LEARNING MODEL FOR TOOTH TREATMENT CLASSIFICATION

Fig. 14 shows the residual block framework used in this study for the identification of dental conditions, and shows the ResNet framework used in the identification of dental conditions. In the model structure of Fig. 14, the first layer is the input layer of data, which defines the input dimension of data and the type of data input. The last layer of the model is the output layer, which is based on the Softmax function

that defines the output type and the dimension of the output data, and classifies the input data.
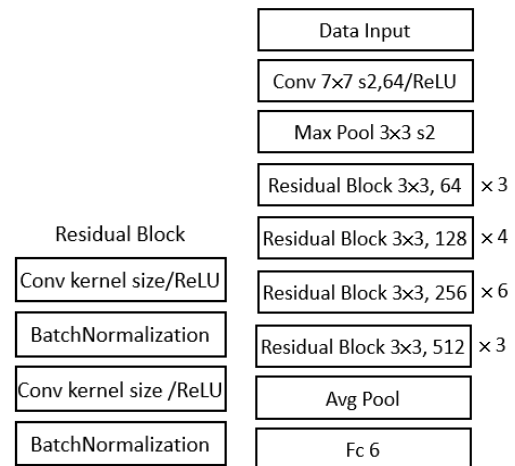

FIGURE 14. Residual block and parameters of ResNet for dental condition.

d: THE EVALUATION OF THE TOOTH CONDITION CLASSIFICATION MODEL

In addition to using ResNet to build an automated dental condition classification system, this study also uses AlexNet, VGGNet, and other representative deep learning models to build the system. The training dataset and the testing dataset are processed using the dataset process methods introduced above, and then the training dataset is input into the model for training. During the training process, the model is allowed to go through several iterations, and the parameters are adjusted so that a complete model of dental condition classification could be obtained in this study. Finally, the experimental model is evaluated using the testing dataset obtained from the dataset process to obtain the results of correct and incorrect dental condition identification.

IV. EXPERIMENTS

A. MODEL EVALUATION METHODS AND EXPERIMENTAL SETTINGS

1) MODEL EVALUATION METRICS

In the tooth position numbering identification result or in the tooth condition identification result, the conditions shown in Table 3 are *TP*, *FP*, *TN*, and *FN* respectively. These metrics are related to the model evaluation equations that will be presented next.

The experimental evaluation formulas used in this study are accuracy, precision, recall, and F1-score. The definition of accuracy is the yielded percentage from the number of correct identification results divided by total tested trails in the case. It also means that how closely the predicted answer corresponds to its "true" answer. The accuracy formula is shown in Eq. (8).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{8}$$

TABLE 3. Table for variables used for performance metrics.

Variable	Definition
TP	True Positive
FP	False positive
TN	True Negative
FN	False negative

Precision and recall are also one of the formulas used to evaluate the experiment in this study. Precision represents how much of the data that is predicted to be positive is actually positive. Recall represents the number of data predicted to be positive out of all the positive data. The formulae for *precision* and *recall* are shown in Eq. (9) and Eq. (10).

$$Precision = \frac{TP}{TP + PP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

The F1-score is a new assessment method for evaluating models. This method integrates precision and recall models to evaluate metrics. We hope to achieve a fair judgment of model performance in an unbiased manner. The formula for the *F1-score* is shown in Eq. (11).

$$F1 - score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{11}$$

B. THE CONTENTS OF THE FIRST PHASE EXPERIMENT AND THE SECOND PHASE EXPERIMENT

This study builds a two-phase DPR classification system using deep learning models such as AlexNet, VGGNet, GoogLeNet, Xception, and ResNet, and investigates which model with data processing method is suitable for which phase. The first phase is to identify the position of the tooth in the DPR. The second phase is to identify the dental condition in the DPR. In order to improve the accuracy of the two-phase classification, the Sobel method, the horizontal displacement data augmentation method, and the zoom-in data augmentation method are also used to pre-processing the data in advance. Next, this study divides these data processing methods into two phases and discusses them in the form of cases. Like many deep learning studies, in this study, we tuned many hyperparameters to achieve better performance in neural network models. After many attempts, we found that when the batch size is set between 4 and 8, the learning rate is between 0.0001 and 0.00001, and the dropout rate is between 0.4 and 0.5, better performance can be obtained. Therefore, the experimental hyperparameters of this study are based on this reference setting to obtain experimental results. However, although better model performance can be achieved by using these hyperparameters, it is still necessary to adjust the hyperparameters according to different model architectures.

The following section is a discussion of the DPR image data cut into single tooth datasets and matrixed. The

experiment of this study designed a total of 8 cases to compare and evaluate various situations. It is expressed as follows.

Case 1: phase 1 - original dataset.
Case 2: phase 1 - with Sobel dataset.
Case 3: phase 1 - with augmentation-horizontal dataset.
Case 4: phase 1 - with augmentation-resize dataset.
Case 5: phase 2 - original dataset.
Case 6: phase 2 - with Sobel dataset.
Case 7: phase 2 - with augmentation-horizontal dataset.
Case 8: phase 2 - with augmentation-resize dataset.

1) FIRST PHASE

In the first phase, there are three methods for data processing. We discuss the training dataset and testing dataset according to (1) normal process, (2) image pre-process, and (3) data augmentation process. The contents of the first phase of the dataset can be seen in Table 4.

a: NORMAL PROCESS

In the normal process section of the first phase, the 32 tooth positions are identified. Since each of the labels in the training dataset in this part of the study has 136 images of size $227 * 227$. Therefore, the training dataset without any processing has a total of $32 * 136$ images of $227 * 227$. There are 10 images for each type of label in the testing dataset. Therefore, the total number of testing datasets without any processing is $32 * 10$ images of $227 * 227$. The training data and testing data for this step can be found in Case 1 of Table 4.

b: IMAGE PRE-PROCESS

The image pre-process in the first phase of this study uses the training dataset and the testing dataset from the previous step to perform the Sobel process. In this step, the processing action of the training dataset needs to be the same as the processing action of the testing dataset. In this section, both the training dataset and the testing dataset are treated in the same way, but the number of these two data sets is different because this step uses Sobel to process the data from the previous step. Therefore, the number of training datasets processed by Sobel is the same as the number of training datasets without any processing, which is $32 * 136$. The number of test sets processed by Sobel is also the same as the number of testing sets without any processing, which is $32 * 10$. The contents of the training dataset and the testing dataset processed in this step can be seen in Case 2 in Table 4.

c: DATA AUGMENTATION PROCESS

In total, two data augmentation methods are used in this study, namely horizontal displacement and resize images. In this step, this study also deals with the training dataset and the testing dataset without any processing method. The testing dataset used in the data augmentation step is the same as the testing dataset without any method processing, which is $32 * 10$ testing data set pictures. However, in the training dataset, this study uses horizontal displacement and resize images to process the training data for data augmentation. Therefore,

the process of using data augmentation to process training data is to use horizontal displacement or resize images to process the original image and thus get two processed images. Then add these two images to the original training dataset, and the training data set will become a training data set after data augmentation. In the augmented training dataset, there will be one original image and two augmentation images for each type of label. The number of pictures for each type of label will become 136 * 3 pictures of 227 * 227 size. In this study, a total of 32 types of labels were used in the first phase. The number of pictures in the training dataset will become 136 * 32 * 3 after augmentation. The training data content and testing dataset content of horizontal displacement can be seen in Case 3 in Table 4. The training data content and testing dataset content of resize images can be found in Case 4 of Table 4.

TABLE 4. Training and testing data case detail in phase one.

Tested case	Dataset process method	Image size (pixel)	Training data volume	Testing data volume
Case 1	Normal process	227 * 227	4352	320
Case 2	Sobel process	227 * 227	4352	320
Case 3	Data augmentation for horizontal displacement	227 * 227	13056	320
Case 4	Data augmentation for resize	227 * 227	13056	320

2) SECOND PHASE

In the second phase, there are three methods for data processing. We discuss the training dataset and testing dataset according to (1) normal process, (2) image pre-process, and (3) data augmentation process. The contents of the second phase of the dataset can be seen in Table 5.

a: NORMAL PROCESS

In the normal process section of the second phase, the 6 dental conditions are identified. Each of the labels in the training dataset has 469 images of size 227 * 227, so the training dataset without any processing has a total of 6 * 469 images of 227 * 227. There are 10 images for each type of label in the testing dataset, so the total number of testing datasets without any processing is 6 * 10 images of 227 * 227. The training data and testing data for this step can be found in Case 5 of Table 5.

b: IMAGE PRE-PROCESS

In the second phase of the image pre-process, the Sobel image pre-processing method is also used to process the unprocessed training dataset and the testing dataset. The difference between this part and the first phase is the identification of the six dental conditions and the amount of data. The rest of the processing steps are the same as the first phase of image pre-process. In the training dataset, there are 469 images of 227 * 227 for each type of label. The training dataset, therefore, has

a total of 6 * 469 images of 227 * 227. The number of images for each type of label in the testing dataset is 10 images of 227 * 227, so there will be a total of 6 * 10 images of 227 * 227 in the testing dataset. The training data and testing data for this step can be found in Case 6 of Table 5.

c: DATA AUGMENTATION PROCESS

In the second phase of data augmentation, the processing steps using the data augmentation method are no different from those in the first phase. The difference between this part and the first phase of data augmentation is in the identification of the items into 6 dental conditions and the amount of data. There are a total of 469 images of size 227 * 227 for each type of label in the training dataset, so the augmented training dataset will have a total of 469 * 6 * 3 images of 227 * 227. The testing dataset has 6 * 10 images of 227 * 227. The training data content and testing dataset content of horizontal displacement can be seen in Case 7 in Table 5. The training data content and testing dataset content of resized images can be found in Case 8 of Table 5.

TABLE 5. Training and testing data case detail in phase two.

Tested case	Dataset process method	Image size (pixel)	Training data volume	Testing data volume
Case 5	Normal process	227 * 227	2814	60
Case 6	Sobel process	227 * 227	2814	60
Case 7	Data augmentation for horizontal displacement	227 * 227	8442	60
Case 8	Data augmentation for resize	227 * 227	8442	60

C. ANALYSIS OF EXPERIMENTAL COMPARISON

1) RESULTS OF PAIRING DEEP LEARNING MODELS WITH DATA PROCESSING METHODS

In the experiments, we first find the model with the best performance through the accuracy indicator in all experiment results from table 6 to table 11, because accuracy is the most commonly used standard indicator for model performance. Then, use three indicators of precision, recall, and F1-score to verify the final result of the best model in tables 12 and 13.

a: FIRST PHASE

The first phase of identification uses AlexNet, VGGNet, GoogLeNet, Xception, and ResNet to build the tooth position numbering identification model for the 32 tooth positions in the DPR. The following are the four cases differentiated for different data processing methods, and the results of each with different models.

- *Case 1: phase 1 - original dataset*

Case 1 is obtained by cutting the DPR so there is a single tooth training dataset and testing dataset. Deep learning models, including AlexNet, VGGNet, ResNet, Xception, and GoogLeNet, are implemented for training a tooth position numbering classification model. The identification accuracy

is shown in Table 6. It can be seen that the classification accuracy of the model increases as the number of model layers increases in the identification of tooth position numbering. This means that as the number of layers of the model increases, the model learns more features of a single tooth picture. The results show that in the performance of the deeper model (ResNet, Xception, GoogLeNet), the ResNet model can perform better.

• *Case 2: phase 1 - with Sobel dataset*

The training dataset and the testing dataset of Case 2 are processed by Sobel. In other words, Case 2 data sets are the same as Case 1 data sets, but Case 2 was processed by Sobel. The accuracy of Case 2 is shown in Table 6. It can be seen that although the Sobel-processed dataset is fed into AlexNet, Xception, and GoogLeNet for training, they result in the poorer classification accuracy. However, when fed into VGGNet and ResNet, the Sobel-processed dataset can improve the classification accuracy. The best performing model here is still ResNet.

TABLE 6. The testing accuracy of Case 1 and Case 2 in phase one.

Tested case	AlexNet	VGGNet	ResNet	Xception	GoogLe Net
Case 1	84.68%	85.62%	**90.93%**	89.68%	85%
Case 2	82.18%	88.125%	**90.93%**	87.81%	82.81%

• *Case 3: phase 1 - with augmentation-horizontal dataset*

The training dataset of Case 3 is the dataset processed by the data augmentation method of horizontal displacement. The testing dataset of Case 3 is the same as that of Case 1. Since the amount of horizontal displacement can be set when using Keras for data augmentation, the horizontal displacement width can be divided into 0.1 to 0.9 for discussion. The meaning of the 0.1 to 0.9 is that if the number is 0.1, then the width is the original width multiplied by 0.1. As shown in Table 7, the classification accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet after training with the training data of Case 3 is higher than that of Case 1. The accuracy of Case 3 is the same or higher than that of Case 1. In Case 1, the classification accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet are 84.68%, 85.62%, 90.93%, 89.68%, and 85% respectively. However, after training the model with the data-augmented training dataset. AlexNet, VGGNet, ResNet, Xception, and GoogLeNet achieved the highest accuracy rates of 90.31%, 91.25%, 95.62%, 95%, and 90% respectively. The accuracy of the models increased by 5.63%, 5.63%, 4.69%, 5.32%, and 5% respectively. It shows that in the first phase of the data augmentation horizontal displacement method, the horizontal displacement data augmentation method is helpful to improve the accuracy of model classification.

• *Case 4: phase 1 - with augmentation-resize dataset*

The training dataset of Case 4 is a dataset that has been scaled up and scaled down by the data augmentation method. The testing dataset of Case 4 is the same as that of Case 1. Because this study can set the value by Keras to determine the

range of image zoom-in and zoom-out, the random zooming range is categorized as 0.1 to 0.9. The meaning of the 0.1 to 0.9 is that if the number is 0.1, then the scaling range is between 1 minus 0.1 and 1 plus 0.1. As shown in Table 8. In Case 1, the testing accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet are 84.68%, 85.62%, 90.93%, 89.68%, and 85% respectively. However, in Case 4, AlexNet, VGGNet, ResNet, Xception, and GoogLeNet achieved the highest accuracy rates of 88.75%, 91.56%, 95.62%, 95%, and 89.06%. The accuracy of these three models improved 4.07%, 5.94%, 4.69%, 5.32%, and 4.06% respectively. Although some of the tests in Case 4 were not more accurate than Case 1, most of the tests were more accurate than or equal to Case 1. Moreover, in Case 4, the testing accuracy increased by up to 5.94%. The growth in testing accuracy is 0.31% higher than the growth in testing accuracy of Case 3. It shows that the data processing in Case 4 is better than Case 3 when combined with model training.

TABLE 7. The testing accuracy of Case 3 in phase one.

Fraction of total width	AlexNet	VGGNet	ResNet	Xception	GoogLeNet
0.1	90.31%	89.06%	91.56%	91.56%	86.56%
0.2	88.75%	87.18%	94.68%	94.06%	86.56%
0.3	88.75%	87.18%	94.06%	93.13%	89.69%
0.4	88.75%	87.5%	95.62%	95%	87.50%
0.5	87.5%	88.125%	94.06%	92.50%	86.56%
0.6	87.18%	91.25%	95.62%	92.81%	90%
0.7	88.12%	87.5%	93.43%	92.50%	86.56%
0.8	88.12%	88.75%	94.06%	95%	86.25%
0.9	86.25%	87.18%	94.06%	93.44%	85.31%

TABLE 8. The testing accuracy of Case 4 in phase one.

Range for random zoom	AlexNet	VGGNet	ResNet	Xception	GoogLe Net
0.1	86.56%	80.62%	90.31%	91%	85.63%
0.2	86.87%	87.5%	90.93%	94.38%	86.56%
0.3	88.43%	89.06%	95.62%	95%	88.75%
0.4	86.25%	88.43%	93.12%	91.88%	86.25%
0.5	88.75%	88.125%	94.06%	91.56%	85.94%
0.6	86.25%	88.125%	92.18%	92.50%	87.81%
0.7	86.87%	90%	94.37%	94.06%	86.88%
0.8	88.43%	91.56%	92.81%	91.25%	88.13%
0.9	87.18%	87.81%	95%	94.37%	89.06%

b: SECOND PHASE

The second phase identifies the six dental conditions in the DPR: orthodontics, endodontic therapy, dental restoration, impaction, implant, and dental prosthesis. In this section, we use AlexNet, VGGNet, GoogLeNet, Xception, and ResNet to build a dental condition identification model, and divided the data into the following Case 5 to Case 8 according to the data processing method.

• *Case 5: phase 2 - original dataset*

The training dataset for Case 5 is a single tooth dataset cut by DPR and converted into an array of single tooth training

dataset and testing dataset. The dental condition classification accuracy for Case 5 is shown in Table 9. It can be seen that as the number of model layers' increases, the classification accuracy of the model becomes more accurate. This means that the higher the level of the model, the better it will perform in terms of dental condition identification. The same with the previous result of the first phase, in the performance of deeper models (ResNet, Xception, GoogLeNet), the ResNet model still performs best.

● *Case 6: phase 2 - with Sobel dataset*

The number of training and testing datasets in Case 6 is the same as that in Case 5. The difference is that the training and testing datasets in Case 6 are processed by Sobel. The accuracy of Case 6 is shown in Table 9. As you can see, Case 6 of Table 9, after the Sobel process and then the training, the accuracy of the model decreases. The reason for this phenomenon in the present study may be that the learning of dental condition characteristics requires the learning of other picture features in order to be useful. Therefore, if some image information is missing, it may lead to poor performance of the deep learning model. In this case, the best performing model is still ResNet.

● *Case 7: phase 2 - with augmentation-horizontal dataset*

Case 7 is the same as Case 5 dataset after the data augmentation method of horizontal shift. This study can set the displacement amount of horizontal displacement by Keras. In this section, the values are divided into 0.1 to 0.9 for this study. The meaning of the 0.1 to 0.9 is that if the number is 0.1, then the width is the original width multiplied by 0.1. In Case 5, the classification accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet are 85%, 88.33%, 93.33%, 91.66%, and 91.66%. However, in Case 7, AlexNet, VGGNet, ResNet, Xception, and GoogLeNet achieved the highest accuracy rates of 95%, 91.66%, 98.33%, 96.66%, and 96.66% respectively. The accuracy improved 10%, 3.33%, 5%, 5%, and 5% respectively. It shows that this data processing method with a deep learning model will help to improve the accuracy of dental condition identification. The experimental results for Case 7 is shown in Table 10.

● *Case 8: phase 2 - with augmentation-resize dataset*

Case 8 is the Case 5 dataset after the data augmentation method to resize. Because Keras can adjust the zoom-in and zoom-out values, this part of the study is divided into 0.1 to 0.9 for discussion. Table 11 shows the testing accuracy results of the training dataset with scaling up and scaling down and using the model for training. The meaning of the 0.1 to 0.9 is that if the number is 0.1, then the scaling range is between 1 minus 0.1 and 1 plus 0.1. As shown in Table 11.

TABLE 9. The testing accuracy of Case 5 and Case 6 in phase two.

Tested case	AlexNet	VGGNet	ResNet	Xception	GoogLeNet
Case 5	85%	88.33%	**93.33%**	91.66%	91.66%
Case 6	88.33%	86.66%	**91.66%**	86.66%	86.66%

TABLE 10. The testing accuracy of Case 7 in phase two.

Fraction of total width	AlexNet	VGGNet	ResNet	Xception	GoogLeNet
0.1	91.66%	90%	95%	93.33%	93.33%
0.2	90%	88.33%	95%	93.33%	95%
0.3	93.33%	90%	95%	93.33%	95%
0.4	91.66%	88.33%	93.33%	93.33%	95%
0.5	91.66%	88.33%	98.33%	95%	95%
0.6	85%	88.33%	98.33%	93.33%	95%
0.7	91.66%	88.33%	93.33	95%	96.66%
0.8	95%	91.66%	93.33%	93.33%	93.33%
0.9	91.66%	90%	95%	96.66%	95%

TABLE 11. The testing accuracy of Case 8 in phase two.

Range for random zoom	AlexNet	VGGNet	ResNet	Xception	GoogLeNet
0.1	91.66%	88.33%	95%	93.33%	95%
0.2	90%	90%	93.33%	93.33%	95%
0.3	91.66%	88.33%	95%	93.33%	95%
0.4	93.33%	88.33%	95%	93.33%	96.66%
0.5	91.66%	88.33%	93.33%	95%	95%
0.6	91.66%	91.66%	95%	96.66%	93.33%
0.7	88.33%	91.66%	93.33%	95%	95%
0.8	91.66%	91.66%	96.66%	93.33%	95%
0.9	91.66%	88.33%	93.33%	93.33%	95%

In Case 5, the accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet was 85%, 88.33%, 93.33%, 91.66%, and 91.66% respectively. However, in Case 8, the testing accuracy of AlexNet, VGGNet, ResNet, Xception, and GoogLeNet became 93.33%, 91.66%, 96.66%, 96.67%, and 96.66%. The accuracy of these rates increased by 8.33%, 3.33%, 3.33%, 5.01%, and 5%. Although the accuracy of Case 8 has not increased as much as that of Case 7, most of the accuracy rates are higher than or equal to that of Case 5. It shows that this data augmentation method also helps to increase the classification accuracy of deep learning models.

2) DISCUSSION OF PHASE 1 AND PHASE 2

In this section, the best identification accuracy results in the two phases are discussed respectively. In the process, data processing methods and deep learning models are used.

a: FIRST PHASE

Based on the discussion of Case 1 to Case 4, in the first phase, it was found that the accuracy of tooth position numbering classification was up to 95.93%. The deep learning model architecture used is ResNet, and the data processing method used is the horizontal displacement method of the data augmentation method. The horizontal displacement method is set to a value of 0.4 or 0.6, indicating that the displacement interval falls between 0.6 and 1.4 or 0.4 and 1.6. It is shown that the displacement in these two intervals will be the most helpful to improve the accuracy of tooth position numbering classification. It is also shown that ResNet will be more suitable for

tooth position numbering identification. Table 12 shows the accuracy of the precision, recall and F1-score values of the tooth position numbering classification. It can be seen that among the F1-scores from 18 to 48 tooth positions, 29 tooth positions have a classified F1-score higher than 95%. The F1-score was 100% for 21 of the tooth positions. It shows that the ResNet model with the horizontal displacement method in data augmentation can improve the classification efficiency of tooth position numbering identification and is better than the results of previous studies.

Fig. 15 shows the training accuracy and loss against epochs and the validation accuracy and loss against epochs for Table 12. After looking at Fig. 15, we can see that this method stabilizes after 100 epochs of training. The training time for 100 epochs was calculated to be $43 \text{ ms/step} * 100 = 4300 \text{ms}$, so a model with 95% accuracy can be obtained in about 4300ms. And the time required for a picture to go through the model classification is around 18.34s. It shows that the dental position numbering classification method can be practically applied in daily life and improve the efficiency of the dental staff's work.

Fig. 16 shows the confusion matrix of the classification results in the first phase. It can be found from the confusion matrix that although most tooth numbers and positions can be classified correctly, there are still a few prediction errors. In the tooth numbering recognition results of the first phase, the tooth numbers 11 and 12 are most likely to be misclassified. Fig. 17 shows instances of tooth numbers that are easily misclassified in the first phase.

b: SECOND PHASE

From the discussion of Case 5 to Case 8, it can be observed that the highest accuracy of identification of dental conditions was 98.33%. The deep learning model architecture used is ResNet, and the data processing method is data augmented horizontal displacement, and the set horizontal displacement value is 0.5 or 0.6. The data processing results representing the horizontal displacement values of these two values enable
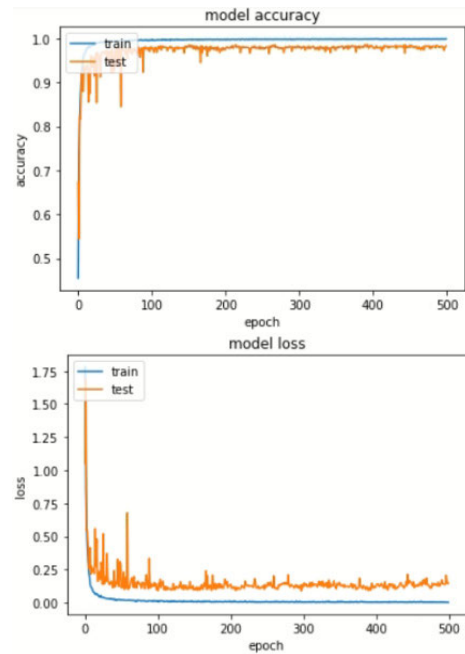

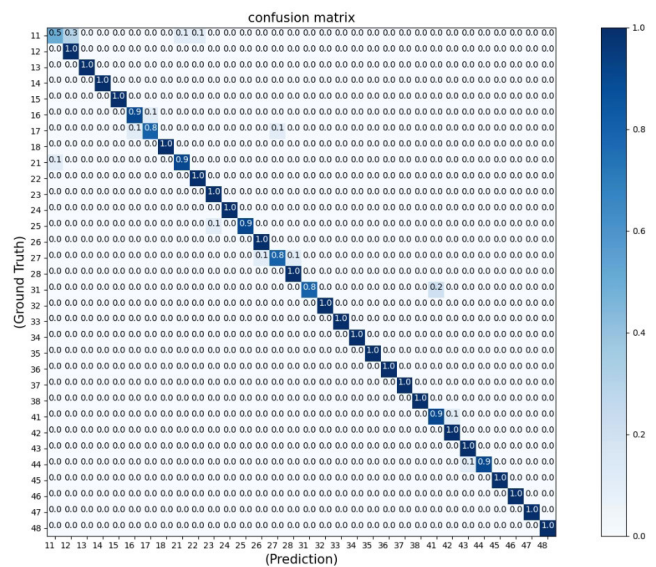**FIGURE 15.** Training and validation of the accuracy and loss curve.

FIGURE 16. Confusion matrix of first phase.

ResNet to perform better. Table 13 shows the six dental condition categories that precision, recall and F1-score obtained using this method. In Table 13, we can see that the F1-score for all six conditions is higher than 95%. The F1-score for four of the conditions was 100%, showing that ResNet models do help to improve the efficiency of dental condition identification when combined with data-augmented horizontal displacement methods.

Fig. 18 shows the confusion matrix of the second phase classification results. From the confusion matrix, we can see that the tooth condition can be completely correctly classified and identified in the majority, but there are still a small
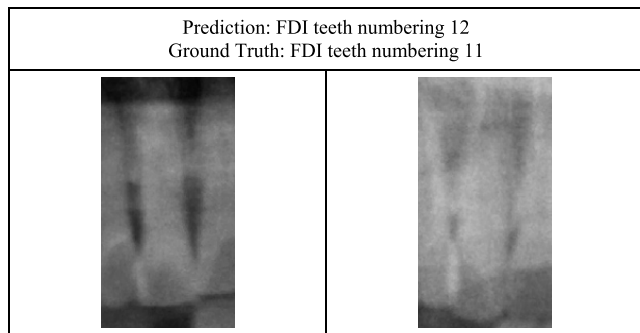
TABLE 12. ResNet with data augmentation in horizontal displacement method for dental position identification data.

FDI Tooth number	Precision	Recall	F1-score
11	0.86	0.60	0.71
12	0.77	1.00	0.87
16	0.91	1.00	0.95
17	1.00	0.90	0.95
18, 13-15	1.00	1.00	1.00
21	1.00	0.90	0.95
22	0.91	1.00	0.95
24	1.00	0.80	0.89
25	0.91	1.00	0.95
27	0.91	1.00	0.95
23, 26, 28	1.00	1.00	1.00
31	1.00	0.90	0.95
32-37	1.00	1.00	1.00
41	0.91	1.00	0.95
42-48	1.00	1.00	1.00

Prediction: FDI teeth numbering 12
Ground Truth: FDI teeth numbering 11

FIGURE 17. Instances of tooth numbers that are easily misclassified. (FDI teeth numbering 11 and 12).

TABLE 13. ResNet with data augmentation in horizontal displacement method for dental condition identification data.

Dental condition	Precision	Recall	F1-score
Orthodontics	1.00	1.00	1.00
Endodontic Therapy	1.00	0.90	0.95
Dental Restoration	0.91	1.00	0.95
Impaction	1.00	1.00	1.00
Implant	1.00	1.00	1.00
Dental Prosthesis	1.00	1.00	1.00

number of predicted classification errors. This study found that the category most likely to be misclassified in the second phase is endodontic therapy, and this category is most often classified as dental restoration. Fig. 19 shows an instance of tooth conditions that are easily misclassified in the second phase.

c: DISCUSSION WITH THE RESULTS OF PREVIOUS STUDIES

In this section, the methods and results of past related studies are compared and discussed with the methods proposed in this study. The first is about the study of Y.-F. Kuo *et al.* [7]. This study classified one-fourth of DPR blocks using the CNN method. The aim was to identify whether a quarter of the DPR blocks had undergone root canal condition with a classification accuracy of 85%. This study was conducted to classify the dental position and dental condition for a single tooth block in the DPR. The identification of the condition can be divided into six categories: orthodontics, endodontic therapy, dental restoration, impaction, implant, and dental prosthesis. The accuracy of the classification of these six conditions can reach up to 98%. Therefore, compared with the study conducted by [7], the present study was superior in terms of the predicted image area, the number of condition items, and the accuracy of the classification. The comparison of this study with Y.-F. Kuo *et al.* and the results are shown in Table 14. In addition, their past work only has studied the identification method of endodontic treatment. Compared with that, the identification accuracy of the method proposed in this study is 95%, which is much better than the past methods. Moreover, our proposed method is used in the identification of more than only endodontic therapy.
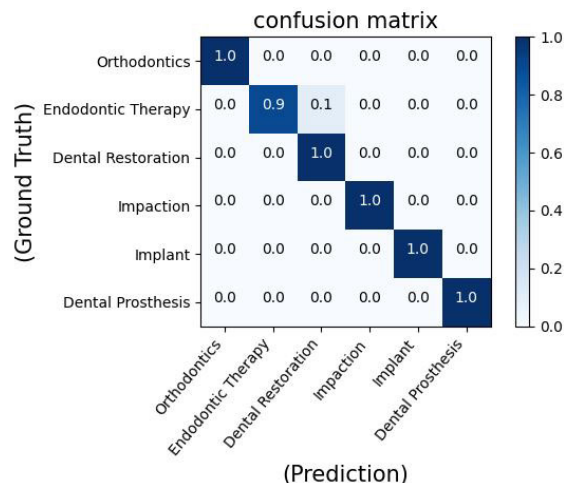

FIGURE 18. Confusion matrix of second phase.

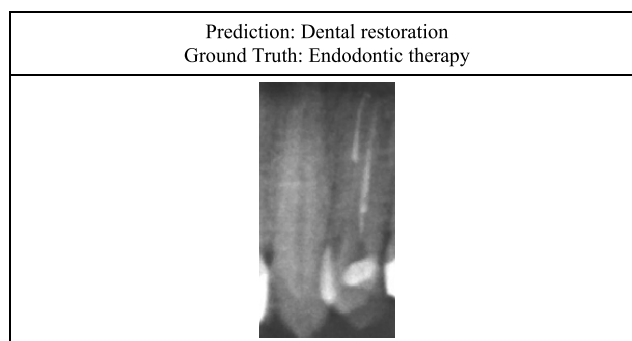Prediction: Dental restoration
Ground Truth: Endodontic therapy

FIGURE 19. Instance of tooth conditions that are easily misclassified. (dental restoration and endodontic therapy).

The next study to be compared is the research of N.-H. Lin *et al.* [6]. In the identification of tooth position numbering, the past study [6] achieved an accuracy of more than 90% in 25 out of 32 teeth and up to 96% in the identification of dental condition. Compared to their study, which used traditional CNN and image processing methods to identify tooth position numbering and seven dental conditions, this study used deep learning models such as AlexNet, VGGNet, GoogLeNet, Xception, and ResNet with deeper model layers, to identify tooth position numbering and six dental conditions. In this study, the accuracy of tooth position numbering classification was achieved in 29 out of 32 teeth with 95% accuracy. This study achieved a 98% accuracy in the identification of 6 dental conditions. Therefore, compared with the study of [6], the present study is more advantageous in terms of both the method used and the results and conditions identified. Table 15 shows the items and results of this study compared with the study of N.-H. Lin *et al.*

D. EXPERIMENTAL SUMMARY

The experiment conducted in this study can be divided into two phases. The first phase is the tooth position numbering classification experiment. The second phase is the dental treatment classification experiment. The two-phase

TABLE 14. Comparison with the past study [7].

	Number of dental conditions	Accuracy	Accuracy on endodontic therapy
Y.-F. Kuo et al.	1	85%	85%
This study	6	98%	95%

TABLE 15. Comparison with the past study [6].

	Tooth position recognition results	Number of dental conditions / Accuracy of dental condition classification
N.-H. Lin et al.	The accuracy of identifying the position of 25 out of 32 teeth was 90%.	7 / 95%
This study	Accuracy of identification of 29 out of 32 tooth positions was 95%	6 / 98%

experiment is preceded by a pre-treatment of the dental panoramic radiograph (DPR) in this study. Each DPR is cut into a single tooth image and divided into a training dataset and a testing data set. The processed training dataset and the testing dataset are then converted into matrices. The last step of pre-processing is to process the training dataset and the testing dataset separately according to different dataset process methods. The training dataset is then fed into the deep learning model for training. Once the training of the model was completed, the testing dataset processed in the above steps was input into the model for evaluation, and the results of the model evaluation and discussion are as follows.

(1) After experimental testing, it was found that the deep learning model is suitable for the automatic classification of DPR.

(2) This study has confirmed that the experimental setup using higher level deep learning models is suitable for solving the image classification problem. It can obtain higher classification accuracy. The use of a deeper layer model architecture allows the model to absorb a wider variety of image features, and therefore the performance of the model becomes better and better.

(3) Sobel image pre-processing can be counter-productive in some cases or for some identifications. Generally, Sobel is most commonly used to remove redundant image background features and detect contours in images. Although in this study, the performance of most models cannot be improved by using Sobel. But from the experimental results, we can still find that in the first phase, the model architecture that can improve performance after using Sobel is VGGNet. After using Sobel in the second phase, the model architecture that can improve performance is AlexNet.

(4) In the comparison between Xception and GoogLeNet, this study found that in the first phase, Xception is more suitable than GoogLeNet. In the second phase,

GoogLeNet's experimental results are slightly better than Xception. However, ResNet was the best model architecture in the two phases.

The horizontal displacement data augmentation method is suitable for DPR classification. This data augmentation method can bring a good growth rate to the model with a deep learning model, for classification of both tooth position numbering and dental condition. If the suitable values are selected are chosen, the model can improve up to 10% with the data augmentation method.

V. CONCLUSION

The DPR provides dental care personnel with additional information on the patient and plays an important role in clinical care. However, in order to understand the patient's dental condition, it is usually necessary to spend a large amount of professional human resource cost and time, which may be prone to human errors. In this study, we proposed a novel two-phase dental panoramic radiograph classification method based on deep learning. The goal of this study is to design and construct a DPR classification system based on CNNs, image pre-processing and data augmentation for tooth position numbering and condition classification. Moreover, this method could classify the dental diagnosis condition, thus providing useful information to assist the professional dentists and medical staff. This medical diagnostic decision support system can be used by dentists to do the follow-up dental treatments for patients. The experiment was divided into two phases: tooth position numbering and condition classification. The major contributions of this study are as follows:

(1) This study can automatically classify where a single tooth image is located in the DPR and determine what condition the tooth currently has. This will reduce the workload of the dental staff and increase the efficiency of DPR processing so that the dental staff can know the patient's dental status and follow up treatment in a short time.

(2) Compared with the past study of N.-H. Lin et al. [6] that used image processing for dental position classification in DPR, the accuracy of dental position numbering classification was higher than 90% in 25 out of 32 teeth. In this study, the number of tooth positions with an accuracy of more than 95% was 29 teeth, and the accuracy of 21 teeth was 100%. The model and the data processing method proposed in this study proved to be helpful in improving the accuracy of tooth position numbering recognition. In addition, in the dental condition classification, our study has 98% accuracy, which is higher than the past study of N.-H. Lin et al. (95%).

(3) Compared with the past study of Y.-F. Kuo et al. [7] that used basic CNN model to recognize the root canal treatment, the accuracy of the past study was 85%. Our study used a newer model with data augmentation to obtain 98% accuracy of all dental

condition identification. This study demonstrates a significant improvement in dental condition identification compared to previous studies.

(4) The data augmentation method proposed in this study showed a good increase in both tooth position numbering classification and dental condition classification. The accuracy of the deep learning model with the data augmentation method can be increased up to 5.94% in dental position numbering recognition. The accuracy of the deep learning model with data augmentation method increased up to 10% in dental condition identification. The average growth rate of the model with data augmentation was about 5% in the first or second phase. This demonstrates that the data augmentation method used in this study can help improve the accuracy of the two phases.

Nevertheless, if we can cut out a single tooth image more accurately and completely through some methods, it may indeed reduce the noise and recognition error rate. That is, the image semantic segmentation method is used to obtain each independent dental tooth image and label them. This is also our future research direction. In this study, the data was processed on the original dataset using image pre-processing methods and data augmentation methods. After training the data with a deep learning model, the best combination of data processing methods and models was obtained. The training time of this model and the time required to test one image are presented for the reference of dental clinics and subsequent researchers.

ACKNOWLEDGMENT

REFERENCES

[1] 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, doi: 10.1145/3065386.

[2] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, pp. 3371–3408, Dec. 2010.

[3] L. Deng and J. C. Platt, "Ensemble deep learning for speech recognition," in *Proc. 15th Annu. Conf. Int. Speech Commun. Assoc.*, Sep. 2014, p. 5.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Computer Vision—ECCV 2014*, D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars, Eds. Cham, Switzerland: Springer, 2014, pp. 346–361, doi: 10.1007/978-3-319-10578-9_23.

[5] B. Chandra and R. K. Sharma, "Fast learning in deep neural networks," *Neurocomputing*, vol. 171, pp. 1205–1215, Jan. 2016, doi: 10.1016/j.neucom.2015.07.093.

[6] N.-H. Lin, T.-L. Lin, X. Wang, W.-T. Kao, H.-W. Tseng, S.-L. Chen, Y.-S. Chiou, S.-Y. Lin, J. F. Villaverde, and Y.-F. Kuo, "Teeth detection algorithm and teeth condition classification based on convolutional neural networks for dental panoramic radiographs," *J. Med. Imag. Health Informat.*, vol. 8, no. 3, pp. 507–515, Mar. 2018, doi: 10.1166/jmihi.2018.2354.

[7] Y.-F. Kuo, S.-Y. Lin, C. H. Wu, S.-L. Chen, T.-L. Lin, N.-H. Lin, C.-H. Mai, and J. F. Villaverde, "A convolutional neural network approach for dental panoramic radiographs classification," *J. Med. Imag. Health Informat.*, vol. 7, no. 8, pp. 1693–1704, Dec. 2017, doi: 10.1166/jmihi.2017.2257.

[8] H. Chen, K. Zhang, P. Lyu, H. Li, L. Zhang, J. Wu, and C.-H. Lee, "A deep learning approach to automatic teeth detection and numbering based on object detection in dental periapical films," *Sci. Rep.*, vol. 9, no. 1, p. 3840, Dec. 2019, doi: 10.1038/s41598-019-40414-y.

[9] C.-H. Wu, W.-H. Tsai, Y.-H. Chen, J.-K. Liu, and Y.-N. Sun, "Model-based orthodontic assessments for dental panoramic radiographs," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 2, pp. 545–551, Mar. 2018, doi: 10.1109/JBHI.2017.2660527.

[10] Y. Zhao, P. Li, C. Gao, Y. Liu, Q. Chen, F. Yang, and D. Meng, "TSAS-Net: Tooth segmentation on dental panoramic X-ray images by two-stage attention segmentation network," *Knowl.-Based Syst.*, vol. 206, Oct. 2020, Art. no. 106338, doi: 10.1016/j.knosys.2020.106338.

[11] 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, doi: 10.1016/j.artmed.2020.101996.

[12] I. Aliaga, V. Vera, M. Vera, E. García, M. Pedrera, and G. Pajares, "Automatic computation of mandibular indices in dental panoramic radiographs for early osteoporosis detection," *Artif. Intell. Med.*, vol. 103, Mar. 2020, Art. no. 101816, doi: 10.1016/j.artmed.2020.101816.

[13] K. Motoki, F. P. Mahdi, N. Yagi, M. Nii, and S. Kobashi, "Automatic teeth recognition method from dental panoramic images using faster R-CNN and prior knowledge model," in *Proc. Joint 11th Int. Conf. Soft Comput. Intell. Syst. 21st Int. Symp. Adv. Intell. Syst. (SCIS-ISIS)*, Dec. 2020, pp. 1–5, doi: 10.1109/SCISISIS50064.2020.9322685.

[14] C. Kim, D. Kim, H. Jeong, S.-J. Yoon, and S. Youm, "Automatic tooth detection and numbering using a combination of a CNN and heuristic algorithm," *Appl. Sci.*, vol. 10, no. 16, p. 5624, Aug. 2020, doi: 10.3390/app10165624.

[15] D. V. Tuzoff, L. N. Tuzova, M. M. Bornstein, A. S. Krasnov, M. A. Kharchenko, S. I. Nikolenko, M. M. Sveshnikov, and G. B. Bednenko, "Tooth detection and numbering in panoramic radiographs using convolutional neural networks," *Dentomaxillofacial Radiol.*, vol. 48, no. 4, May 2019, Art. no. 20180051, doi: 10.1259/dmfr.20180051.

[16] R. Kaur, R. S. Sandhu, A. Gera, and T. Kaur, "Edge detection in digital panoramic dental radiograph using improved morphological gradient and MATLAB," in *Proc. Int. Conf. Smart Technol. Smart Nation (SmartTechCon)*, Aug. 2017, pp. 793–797, doi: 10.1109/SmartTechCon.2017.8358481.

[17] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.

[18] I. Kandel and M. Castelli, "Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review," *Appl. Sci.*, vol. 10, no. 6, 2021, 2021.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[20] R. A. Welikala, P. Remagnino, J. H. Lim, C. S. Chan, S. Rajendran, T. G. Kallarakkal, R. B. Zain, R. D. Jayasinghe, J. Rimal, A. R. Kerr, R. Amtha, K. Patil, W. M. Tilakaratne, J. Gibson, S. C. Cheong, and S. A. Barman, "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer," *IEEE Access*, vol. 8, pp. 132677–132693, 2020, doi: 10.1109/ACCESS.2020.3010180.

[21] D. C. Ciresan, U. Meier, J. Masci, L. M. Gambardella, and J. Schmidhuber, "Flexible, high performance convolutional neural networks for image classification," in *Proc. 22nd Int. Joint Conf. Artif. Intell.*, Barcelona, Spain, 2011, pp. 1237–1242.

[22] V. Nair and G. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Madison, WI, USA, 2010, pp. 807–814.

[23] D. Warde-Farley, I. J. Goodfellow, A. Courville, and Y. Bengio, "An empirical analysis of dropout in piecewise linear networks," 2013, *arXiv:1312.6197*.

[24] A. Z. Arifin, S. Adam, A. M. Mohammad, F. Anggris, R. Indraswari, and D. Navastara, "Detection of overlapping teeth on dental panoramic radiograph," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 6, pp. 71–80, Dec. 2019, doi: 10.22266/ijies2019.1231.07.

[25] O. Vincent and O. Folorunso, "A descriptive algorithm for Sobel image edge detection," in *Proc. InSITE Conf.*, 2009, pp. 97–107, doi: 10.28945/3351.

[26] S. Khan and A. Mukati, "Dataset augmentation for machine learning applications of dental radiography," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 2, pp. 453–456, 2020, doi: 10.14569/IJACSA.2020.0110258.

[27] Y. Miki, C. Muramatsu, T. Hayashi, X. Zhou, T. Hara, A. Katsumata, and H. Fujita, "Classification of teeth in cone-beam CT using deep convolutional neural network," *Comput. Biol. Med.*, vol. 80, pp. 24–29, Jan. 2017, doi: 10.1016/j.compbiomed.2016.11.003.

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

HAO-YUN CHANG was born in Taoyuan, Taiwan. He received the B.S. degree in computer science and information engineering from the National Ilan University, Taiwan, in 2021. His research interests include image processing, machine learning, and deep learning.

• • •
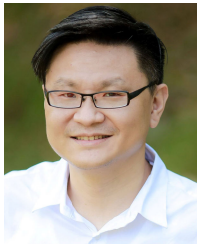
SZU-YIN LIN (Member, IEEE) received the Ph.D. degree from the National Chiao-Tung University, Taiwan, in 2012. He was a Visiting Scholar with Coventry University, U.K., in 2012. From 2013 to 2017, he was an Assistant Professor with the Department of Information Management, Chung Yuan Christian University, Taiwan, where he was an Associate Professor, from 2017 to 2019. Since 2019, he has been an Associate Professor with the Department of Computer Science and Information Engineering, National Ilan University, Taiwan. His research interests include the areas of intelligent data analysis, artificial intelligence, machine learning, deep learning, management information systems, service-oriented computing, and multi-agent distributed computing.