ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

Periodontal Disease Diagnosis App using Deep Learning with Radiographic Images

[1] Rachakrit Monyanon, [2] Supanee Rassameemasmaung, [3] Jira Kitisubkanchana, [4] Pee Topothai, [5] Nuanwan Soonthornphisaj*

[1] [5] Department of Computer Science, Faculty of Science, Kasetsart University, Thailand
[2] [4] Department of Oral Medicine and Periodontology, Faculty of Dentistry, Mahidol University, Thailand
[3] Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Mahidol University, Thailand
Corresponding Author Email: [1] rachakrit.m@ku.th, [2] supanee.ras@mahidol.ac.th, [3] yjira@hotmail.com, [4] pee.top@mahidol.ac.th, [5] nuanwan.s@ku.th

Abstract— Periodontitis is a common oral disease in Thailand that can cause tooth loss if left untreated. This study developed a deep learning model and web/mobile application for diagnosing periodontal disease using radiographic images. The model achieved an impressive accuracy of 94.95% on the test set, surpassing the accuracy of expert dentists. The developed web and mobile application provides an accessible and user-friendly tool for dentists to improve the accuracy and efficiency of the diagnostic process for periodontal disease. The involvement of expert dentists from Mahidol University ensured the accuracy and reliability of the dataset used in this study.

Index Terms— deep learning, periodontal disease, radiographic images, web/mobile application.

I. INTRODUCTION

Periodontitis is a prevalent oral disease among Thai adults and the elderly, with a potential for tooth loss if left untreated. A 2017 national oral health survey found that 25.9% of adults and 36.3% of the elderly had periodontitis with a periodontal pocket depth greater than 4 mm. Radiographic images are commonly used for diagnosis, but the process is subjective and dependent on the expertise of dentists. Therefore, this study aims to develop a user-friendly deep learning model that can classify radiographic images of teeth and serve as a diagnostic tool for dentists. The model will be trained and tested using a dataset of 1,978 radiographic images from the Faculty of Dentistry, Mahidol University, to achieve accuracy, sensitivity, and specificity. Ultimately, the model will be integrated into a user-friendly web and mobile application to diagnose periodontal disease using radiographic images.

II. LITERATURE REVIEW

The use of deep learning in dentistry has shown remarkable progress in recent years. Deep learning algorithms have the ability to process and analyze large amounts of data, enabling the development of new diagnostic and treatment methods that can improve patient outcomes. A fundamental concept in deep learning was proposed by Lecun et al. (1998) [1] who introduced a gradient-based learning approach for document recognition.

In dentistry, deep learning has been utilized for various applications, including the detection of dental caries, classification of radiographic images, and diagnosis of periodontal diseases. Lee et al. (2018) [2] developed a deep learning-based convolutional neural network algorithm that can predict and diagnose periodontally compromised teeth. The deep learning algorithm developed by Lee et al. (2018) [2] showed promising results in diagnosing and predicting periodontally compromised teeth (PCT) using periapical radiographic images. The diagnostic accuracy for PCT was found to be 81.0% for premolars and 76.7% for molars, while the accuracy for predicting extraction was 82.8% for premolars and 73.4% for molars. These results suggest that the deep CNN algorithm can be a useful tool for diagnosing and predicting PCT. Further optimization of the PCT dataset and improvements in the algorithm can be expected to make computer-aided detection systems an effective and efficient method for diagnosing and predicting PCT.

In 2020, Bhattacharjee [3] proposed an automated system for detecting dental cavities, which incorporates deep learning and explainable AI techniques. The system utilizes an intraoral camera to capture more than 500 de-identified images from online sources and consenting human participants. Several neural network architectures and training techniques were experimented with, and the ResNet-27 architecture was found to be the most effective for cavity detection, achieving an accuracy of 77.8% and a sensitivity score of 0.69, through transfer learning from an ImageNet1k dataset. To further improve the accuracy and sensitivity score, two-stage curriculum learning was implemented, resulting in an accuracy of 82.8% and a sensitivity score of 1.0. Additionally, LIME was used to generate visual explanations for the cavity diagnoses made by the system, enabling it to explain its diagnosis to end-users in a comprehensible manner. As a result, the system is now capable of detecting the presence of cavities and providing

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

interpretable explanations for its diagnosis.

Additionally, Thanathornwong and Suebnukarn [4] developed an automated approach for identifying periodontally compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks (R-CNN). The model achieved an average precision rate of 0.81, indicating a significant overlap between predicted and actual regions. The average recall rate of 0.80 further demonstrated that the method was able to effectively exclude healthy teeth areas while identifying periodontally compromised teeth regions. With a sensitivity of 0.84, specificity of 0.88, and F-measure of 0.81, the model performed satisfactorily in detecting periodontally compromised teeth, even when trained on a limited amount of labeled imaging data. This approach could potentially reduce diagnostic effort and save assessment time by allowing for automated screening documentation.

Furthermore, the study by Thanh et al. (2022) [5] aimed to diagnose the stages of smooth surface caries via smartphone images using a deep learning algorithm. The study used a training dataset consisting of 1,902 photos of the smooth surface of teeth taken with an iPhone 7 from 695 people. The results showed that YOLOv3 and Faster R-CNN models had the highest sensitivity levels for detecting cavitated caries lesions, while the specificity of all four models was above 71% for visually non-cavitated caries. The study concluded that the clinical application of YOLOv3 and Faster R-CNN models for diagnosing dental caries via smartphone images was promising, providing a preliminary insight into the potential translation of AI from the laboratory to clinical practice.

These studies demonstrate the potential of deep learning in various aspects of dentistry and provide valuable insights for the development of the current study.

III. MATERIALS AND METHODS

A. Materials

The dataset used for this study was collected from the Faculty of Dentistry, Mahidol University. The dataset includes a total of 1,978 radiographic images of teeth, out of which 989 images were diagnosed by two out of three expert dentists from the Faculty of Dentistry, Mahidol University, as periodontitis-affected, and 989 images were healthy radiographic images. The images were saved in .bmp and .jpg file formats, and the size of the images was 1,024 pixels × 1,024 pixels.

The dataset used in this study comprises 1,978 radiographic images of teeth, with 1,582 patients in the training set, 198 in the validation set, and 198 in the test set. The training set includes 662 males (41.85%) and 920 females (58.15%). The majority of patients in the training set fall within the 40 to 59 years age group, with 538 (34.01%) patients. The training set includes 791 (50.00%) images diagnosed with periodontal diseases and 791 (50.00%) images that are normal. The validation and test sets are evenly distributed between the two classifications.

Fig.1 Example of radiographic images of teeth affected by periodontitis.

Fig.2 Example of radiographic images of healthy teeth

Table I. Data distribution of the dataset

Characteristics	Training Set	Validation Set	Test Set
Patients	1,582 (100.00%)	198 (100.00%)	198 (100.00%)
Sex			
Male	662 (41.85%)	82 (41.41%)	78 (39.39%)
Female	920 (58.15%)	116 (58.59%)	120 (60.61%)
Age group (yrs.)			
<20	***54*** (3.41%)	11 (5.56%)	3 (1.52%)
20-29	128 (8.09%)	14 (7.07%)	9 (4.55%)
30-39	226 (14.29%)	26 (13.13%)	31 (15.66%)
40-49	***341*** (21.55%)	40 (20.20%)	54 (27.27%)
50-59	***538*** (34.01%)	65 (32.83%)	55 (27.78%)
60-69	***216*** (13.65%)	30 (15.15%)	33 (16.67%)
>=70	***79*** (4.99%)	12 (6.06%)	13 (6.57%)
Diagnosis			
PCT	791 (50.00%)	99 (50.00%)	99 (50.00%)
Healthy	791 (50.00%)	99 (50.00%)	99 (50.00%)

B. Methods

Three models will be developed for diagnosing periodontitis from radiographic images. The first two models will utilize Transfer Learning [6] from pre-trained VGG16 and VGG19 models, respectively, which is a successful method for medical image classification by retraining pre-existing models with radiographic images to solve similar problems. The third model will be a newly-created CNN model from scratch, without using Transfer Learning

IFERP
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 10, Issue 7, July 2023

from other models. Developing a CNN model from scratch requires expertise and may take longer to train than Transfer Learning, but it has the potential to provide better results. The models will be named PCT_VGG16, PCT_VGG19, and PCT_CNN

In order to enhance the performance of the models and prevent overfitting, Image Augmentation techniques [7], [8] will be employed on the dataset, which includes rotation, zooming, and shear. The models will be trained on a portion of the dataset, while the rest will be utilized for validation and testing purposes. Performance evaluation of the models will be conducted using metrics such as accuracy, precision, recall, and F1 score. The best-performing model will be optimized for deployment on a web/mobile application and subjected to rigorous testing to guarantee its accuracy and dependability in diagnosing periodontitis from radiographic images.

The machine used for training the models has the following specifications: an Intel Core i7-12700 2.1GHz CPU, an NVIDIA GEFORCE RTX 3080 graphics card, 32 GB of RAM, and 1 TB of HDD storage.

To speed up the training of the models, the TensorFlow program is set to use CUDA to utilize the GPU for faster training. CUDA is NVIDIA's computing platform that supports GPU acceleration and helps TensorFlow work faster, making the training process faster and more time efficient.

The performance of the best-performing model in diagnosing periodontitis will be compared to that of six professionals, including three general practitioners (GPs) and three specialists (SPs) from the Faculty of Dentistry at Mahidol University. The model with the highest performance will be selected for further development, which will involve creating a web/mobile application in the next phase.

IV. DEVELOPING DEEP LEARNING MODEL

The PCT_VGG16 and PCT_VGG19 models are created through transfer learning from pre-trained VGG16 and VGG19 models, which involves loading the models using the Keras library, removing the fully connected layer, freezing the layers, adding a new fully connected layer with 32 units and ReLU activation, adding an output layer with 2 units and SoftMax activation, creating the new models, and compiling them with the RMSprop optimizer [9], [10] and categorical cross-entropy loss function. The input layer is set to the desired size of 512 × 512 × 1, and the models are designed to classify between two classes considered positive (P) and negative (N).

The PCT_CNN model used in this case is a Convolutional Neural Network (CNN), which is a model created from scratch and not using transfer learning. The PCT_CNN model is created with the following details: an input layer that takes in a 512 × 512 × 1 input image, two Conv2D layers with same padding using 3 × 3 filters and 64 channels, a MaxPooling2D layer with 2 × 2 pool size, a Conv2D layer with same padding using 3 × 3 filters and 128 channels, another MaxPooling2D layer, followed by two more Conv2D layers with same padding using 3 × 3 filters and 64 and 32 channels respectively, and another MaxPooling2D layer. The data is then flattened into a single vector, followed by a Dense layer with 32 nodes and ReLU activation, a Dropout layer with a dropout rate of 0.2 to reduce overfitting, and a final Dense layer with 2 nodes and SoftMax activation to classify between two classes for diagnosing Periodontal Disease.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 512, 512, 1)	0
conv2d_5 (Conv2D)	(None, 512, 512, 64)	640
conv2d_6 (Conv2D)	(None, 512, 512, 64)	36928
max_pooling2d_5 (MaxPooling2	(None, 256, 256, 64)	0
conv2d_7 (Conv2D)	(None, 254, 254, 128)	73856
max_pooling2d_6 (MaxPooling2	(None, 127, 127, 128)	0
conv2d_8 (Conv2D)	(None, 125, 125, 64)	73792
max_pooling2d_7 (MaxPooling2	(None, 62, 62, 64)	0
conv2d_9 (Conv2D)	(None, 60, 60, 32)	18464
max_pooling2d_8 (MaxPooling2	(None, 30, 30, 32)	0
conv2d_10 (Conv2D)	(None, 28, 28, 16)	4624
max_pooling2d_9 (MaxPooling2	(None, 14, 14, 16)	0
flatten_1 (Flatten)	(None, 3136)	0
dense_2 (Dense)	(None, 32)	100384
dropout_1 (Dropout)	(None, 32)	0
dense_3 (Dense)	(None, 2)	66

Total params: 309,754
Trainable params: 100,450
Non-trainable params: 209,304

Fig.3 The architecture of PCT_CNN

Table II. Summarize the characteristics of all three models.

Model	Total Params	Trainable Params	Non-trainable Params	Number of Layers
PCT_VGG16	17,926,786	4,194,402	13,732,384	21
PCT_VGG19	23,119,554	4,194,402	18,925,152	25
PCT_CNN	309,754	100,450	209,304	16

To train the three models, a dataset of 1,582 radiographic images will be used and each model will be trained in eight different ways, incorporating various image augmentation techniques outlined in Table III. Before input to the model,

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

the radiographic image size will be reduced to 512 pixels × 512 pixels and converted to grayscale during training. The hyperparameters will be set to a train batch size of 16, a validation batch size of 16, and an epoch of 100, with the RMSprop optimizer function being used.

Table III. Details of the 8 image augmentation techniques used for training each model.

Augment Type	Description
A	no augment
B	zoom 120%
C	rotate 45
D	shear
E	zoom 120%, rotate 45
F	zoom 120%, shear
G	rotate 45, shear
H	zoom 120%, rotate 45, shear

Image augmentation can be performed using ImageDataGenerator from Keras. The rotation range parameter can be set to 45, which randomly rotates the images between -45 and 45 degrees. The shear range can be set to apply a shear transformation with a range of 0.1. A zoom_range of 0.2 allows for zooming in or out of the images by up to 20%.

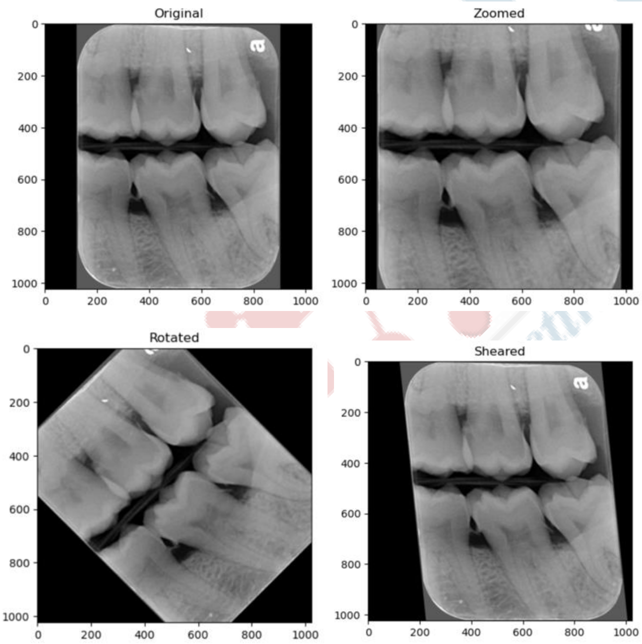

Fig.4 Example images with image augmentation using rotation, zoom, and shear techniques

After training the models, saving them in the widely used Hierarchical Data Format version 5 (HDF5) will enable easy integration into web and mobile applications. HDF5 is a popular file format for storing and exchanging scientific data, particularly in the field of deep learning, and it enables efficient storage and retrieval of complex datasets, including trained machine learning models. As a result, there will be a total of 24 models saved in HDF5 files.

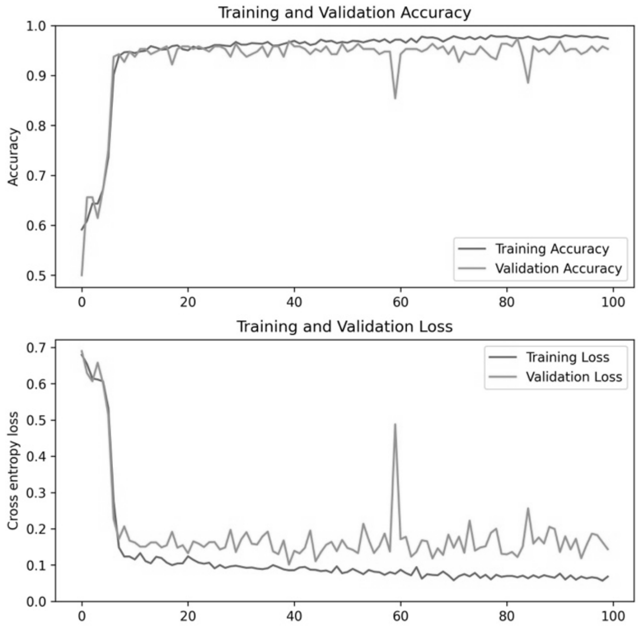

Fig.5 Example results of training the PCT_CNN model with augment type C

V. DEEP LEARNING MODEL RESULTS

To evaluate the performance of the 24 trained models, various metrics such as the confusion matrix, accuracy, sensitivity, precision, and F1 Score will be employed. The confusion matrix is a table that provides the number of true positive, true negative, false positive, and false negative predictions. Accuracy is computed by dividing the number of correctly classified samples by the total number of samples. Sensitivity, or true positive rate, measures the proportion of actual positive cases that are correctly identified, while precision measures the proportion of true positive cases out of all positive predictions made. The F1 Score is a harmonic mean of precision and recall and provides a balanced measure of the model's accuracy. The following equations will be used to compute these metrics:

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + TN + FP + FN)}$$

$$\text{Sensitivity (Recall)} = \frac{TP}{(TP + FN)}$$

$$\text{Precision} = \frac{TP}{(TP + FP)}$$

$$\text{F1 Score} = 2 \times \frac{(\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})}$$

where TP represents true positive, TN true negative, FP false positive, and FN false negative.

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

The performance of the 24 trained models was tested on a total of 198 images in the test set, which consisted of 99 positive and 99 negative cases with ground truth labels.

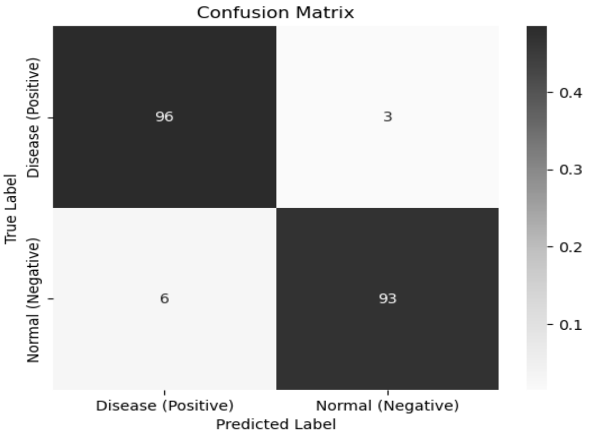

Fig.6 Example of the confusion matrix for the PCT_CNN model using augment type C

The table IV. presents performance metrics of 24 trained models for disease diagnosis. The models include PCT_VGG16_A through H, PCT_VGG19_A through H, and PCT_CNN_A through H. The evaluation metrics, including accuracy, precision, recall, and F1 score, help to understand the strengths and weaknesses of each model.

The accuracy of the models ranges from 0.5 to 0.9545. The highest accuracy is achieved by the PCT_CNN_C.h5 model, with a score of 0.9545. The precision scores range from 0.5196 to 0.9490. The recall scores range from 0 to 0.9697, and the F1 score ranges from 0 to 0.9552.

The results indicate that the PCT_CNN models perform better than the PCT_VGG16 and PCT_VGG19 models. Additionally, some models have a good precision score but a low recall score (e.g., PCT_VGG19_E.h5), while others have a good recall score but a low precision score (e.g., PCT_VGG16_E.h5). It is important to note that some models have a NaN value in the precision or F1 score column, which indicates that the value cannot be computed.

Table IV. Performance summary of the 24 trained models

No.	Model	Performance on Test Data			
		Accuracy	Precision	Recall	F1 Score
1	PCT_VGG16_A.h5	0.6970	0.7600	0.5758	0.6552
2	PCT_VGG16_B.h5	0.5960	0.5725	0.7576	0.6522
3	PCT_VGG16_C.h5	0.6970	0.6822	0.7374	0.7080
4	PCT_VGG16_D.h5	0.7172	0.6903	0.7879	0.7358
5	PCT_VGG16_E.h5	0.5354	0.5196	0.9394	0.6691
6	PCT_VGG16_F.h5	0.5909	0.5600	0.8485	0.6747
7	PCT_VGG16_G.h5	0.6970	0.8095	0.5152	0.6296
8	PCT_VGG16_H.h5	0.6111	0.7200	0.3636	0.4832
9	PCT_VGG19_A.h5	0.7172	0.7416	0.6667	0.7021
10	PCT_VGG19_B.h5	0.5960	0.5960	0.5960	0.5960
11	PCT_VGG19_C.h5	0.6717	0.6308	0.8283	0.7162
12	PCT_VGG19_D.h5	0.7172	0.6903	0.7879	0.7358
13	PCT_VGG19_E.h5	0.5051	0.6667	0.0202	0.0392
14	PCT_VGG19_F.h5	0.5000	NaN	0	NaN
15	PCT_VGG19_G.h5	0.7222	0.7444	0.6768	0.7090
16	PCT_VGG19_H.h5	0.5000	0.5000	1.0000	0.6667
17	PCT_CNN_A.h5	0.9293	0.9126	0.9495	0.9307
18	PCT_CNN _B.h5	0.9293	0.9381	0.9192	0.9286
19	***PCT_CNN_C.h5***	***0.9545***	***0.9412***	***0.9697***	***0.9552***
20	***PCT_CNN_D.h5***	***0.9444***	***0.9490***	***0.9394***	***0.9442***
21	PCT_CNN_E.h5	0.9293	0.9474	0.9091	0.9278
22	PCT_CNN_F.h5	0.9343	0.9479	0.9192	0.9333
23	***PCT_CNN_G.h5***	***0.9343***	***0.9388***	***0.9293***	***0.9340***
24	PCT_CNN_H.h5	0.9293	0.9126	0.9495	0.9307

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

The performances of the top three models (PCT_CNN_C, PCT_CNN_D, and PCT_CNN_G) in diagnosing periodontitis will be compared to those of six professionals, including three general practitioners (GPs) and three specialists (SPs) from the Faculty of Dentistry at Mahidol University.

"GP Vote" to be considered, at least two out of three General Practitioners must agree on a diagnosis or have a concurrence score of two or more. This level of consensus plays a role in the overall decision-making process.

"SP Vote," a consensus must be reached among at least two out of three Specialists, either through providing the same diagnosis or achieving an agreement score of two or more. This score is a factor in the overall decision-making process.

"CNN Vote" to be considered, at least two out of three models that use a convolutional neural network (CNN) must agree on a prediction or achieve an agreement score of two or more. This score is a factor in the overall decision-making process.

Table V. Summary of periodontitis diagnosis results by General Practitioners (GP), Specialists (SP), and selected CNN models

Specialist/ Model	Performance on Test Data			
	Accuracy	Precision	Recall	F1 Score
GP1	0.7828	0.7000	0.9899	0.8201
GP2	0.8283	0.7982	0.8788	0.8365
GP3	0.8030	0.7381	0.9394	0.8267
GP Vote	***0.8434***	***0.7698***	***0.9798***	***0.8622***
SP1	0.8586	0.8515	0.8687	0.8600
SP2	0.5960	0.5531	1.0000	0.7122
SP3	0.7273	0.6642	0.9192	0.7712
SP Vote	***0.7273***	***0.6596***	***0.9394***	***0.7750***
PCT_CNN_C	0.9545	0.9412	0.9697	0.9552
PCT_CNN_D	0.9444	0.9490	0.9394	0.9442
PCT_CNN_G	0.9343	0.9388	0.9293	0.9340
CNN Vote	***0.9495***	***0.9406***	***0.9596***	***0.9500***

The table V. presents performance metrics for three models (PCT_CNN_C, PCT_CNN_D, and PCT_CNN_G) and six dental professionals (three general practitioners and three specialists) in diagnosing periodontitis. The models and professionals were evaluated based on accuracy, precision, recall, and F1 score.

The models achieved accuracy scores ranging from 0.9343 to 0.9545, with PCT_CNN_C achieving the highest accuracy score of 0.9545, followed by the CNN Vote model at 0.9495. Precision scores for the models ranged from 0.9388 to 0.9412, while recall scores ranged from 0.9293 to 0.9697. F1 scores ranged from 0.934 to 0.9552.

In contrast, dental professionals (GPs and SPs) achieved an accuracy range of 0.596 to 0.8586, with SP1 achieving the highest accuracy at 0.8586. Precision scores for dental professionals ranged from 0.5531 to 0.8515, while recall scores ranged from 0.8687 to 1. F1 scores ranged from 0.7122 to 0.86.

The results indicate that the PCT_CNN models outperformed the dental professionals in terms of accuracy, precision, recall, and F1 score. The CNN Vote model achieved the highest F1 Score of 0.95, while the SP Vote achieved the highest recall score of 0.9394. Notably, although SPs generally performed better than GPs, their overall performance was still lower than that of the PCT_CNN models.

VI. DEVELOPING WEB AND MOBILE APPLICATION

The three best performing models, PCT_CNN_C, PCT_CNN_D, and PCT_CNN_G, were chosen to be developed into a web and mobile application to assist dentists in diagnosing periodontitis. The development process was divided into two parts:

The Server (Web/Web API) was developed using Python and the Flask Framework, and then converted into a Docker Image to be deployed as a container on the AZURE App Service.

The Client (Mobile Application) was developed using Ionic for both the iOS and Android operating systems. The mobile application communicates with the server through the HTTP protocol to call the Web API and predict the results of radiographic images.

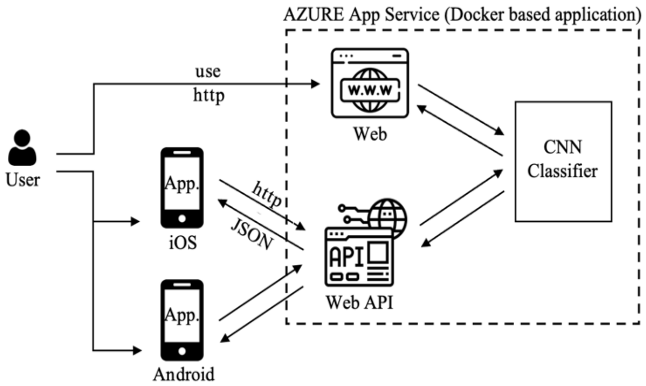

Fig.7 System architecture of the application

A. Server (Web/Web API)

The development process of a Flask Framework-based web application for predicting whether radiographic images are positive or negative for periodontitis is as follows: Firstly, the installation of Flask Framework and necessary modules such as TensorFlow, Keras, NumPy, and OpenCV. Secondly, the creation of a web page for users to upload images with HTML, CSS, and JavaScript. Thirdly, defining a Flask route for receiving the uploaded image through http post and storing it in the Temp Folder. Fourthly, loading three CNN models: PCT_CNN_C.h5, PCT_CNN_D.h5, and

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 10, Issue 7, July 2023

PCT_CNN_G.h5, from file storage. Fifthly, developing a predict function to process the image with the three CNN models and return the result. Sixthly, creating a vote function to combine the results of all three CNN models and return the final outcome as positive or negative. Finally, using Flask route to display the result on the web page with HTML, CSS, and JavaScript. After completion, users can upload radiographic images through the web page, and the system will predict whether they are positive or negative for periodontitis using the voting result from all three CNN models.

Deploying a Flask-based web application on Azure Web App using Docker for predicting radiographic images involves creating a Dockerfile, uploading the Docker Image to Azure Container Registry using docker push, and creating an Azure Web App for Containers configured with the Docker Image. After setting up the necessary environment variables, including the Flask secret key, the app is started, making the Flask app accessible on Azure Web App for predicting radiographic images.

Fig.8 Web page screenshot for periodontitis diagnosis

B. Client (Mobile Application)

The operation of the mobile application involves capturing a photo using the mobile device camera, converting the photo into a Base64 string, and transmitting it to the server via http post. After receiving the result from the server indicating whether the photo is Positive or Negative, the App displays the outcome on the mobile screen and saves it to the local storage for offline viewing of historical data.

Fig.9 Mobile app for diagnosing periodontal disease on Apple (left) and Android (right) smartphones

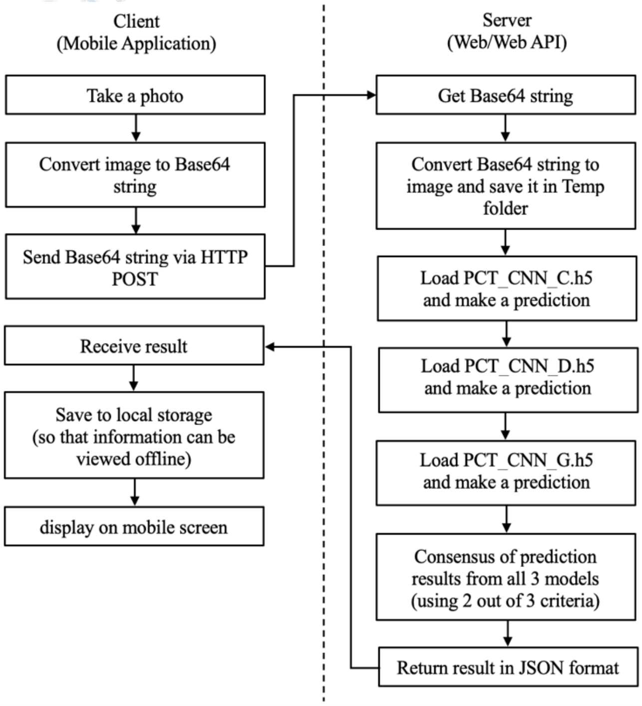

Fig.10 Flowchart illustrating the working process of the mobile application

The process of developing a mobile application for both iOS and Android using the Ionic Framework involves creating an Ionic project with Angular Framework as the core and adding the Camera Plugin to enable the camera feature. The user interface is developed using HTML, CSS, and JavaScript, respectively, with the .ts file utilizing Angular to control the application's functions. The app can be installed on iOS by opening the Ionic project with Xcode and building it as a .ipa file, or for Android, by opening the project with Android Studio and building it as a .apk file for installation.

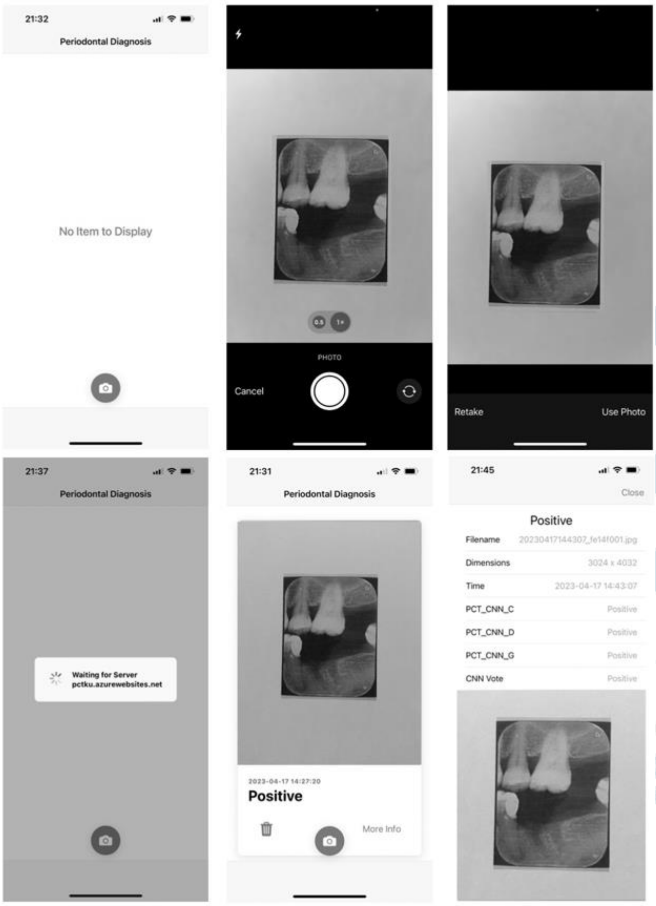

Fig.11 Mobile app screenshot for periodontitis diagnosis

The mobile application was tested on a test set of 198 radiographic images of teeth using the camera on a mobile device, with the following results: an accuracy of 0.7172, precision of 0.6807, recall of 0.8182, and F1 score of 0.7431.

VII. DISCUSSION

The objective of this study was to develop a CNN-based deep learning model and a web and mobile application to aid in the diagnosis of periodontitis using radiographic images. The performance of the model was evaluated using a dataset of 1,978 radiographic images from the Faculty of Dentistry at Mahidol University. The CNN-based model achieved an accuracy of up to 94.95% in diagnosing periodontitis, outperforming dentists' diagnosis accuracy of 72.73%. The CNN architecture demonstrated superior performance compared to the transfer learning method from VGG16 and VGG19. The developed CNN model was used to build a web and mobile application to assist dentists in the diagnosis of periodontitis using radiographic images. However, the use of the camera feature on mobile devices resulted in a decrease in model performance, with accuracy decreasing to 71.72% when capturing images using the camera on a mobile device.

When using a mobile phone camera, the accuracy of the diagnostic model may decrease due to various factors, such as the appearance of shadows in the captured image, brightness levels, sharpness, and other distortions that can affect the model's prediction accuracy. The study demonstrated that image quality and distractions can significantly impact the predictive accuracy of the CNN model developed by the researchers. This is because the model was trained solely on original image files and only used three image augmentation techniques: zoom, rotate, and shear. To improve the accuracy and effectiveness of the mobile application, additional image augmentation techniques, including brightness, contrast, sharpness, and noise, should be implemented to enhance the model's ability to diagnose periodontitis accurately using images captured from mobile phone cameras.

VIII. FUTURE WORK

A. Model Development

In this study, the goal was to develop a classifier model capable of detecting periodontitis in radiographic images. However, to provide more detailed information, future models could use Natural Language Processing (NLP) technology to generate captions or sentences from the radiographic images. By creating clear and accurate captions, dentists could make better decisions about treatment and patient care. The process of generating image captions using NLP involves two primary steps: encoding, which converts the input image into a vector, and decoding, which generates a description of the image using techniques such as RNN, LSTM, Attention Mechanism, or Transformer. The feature extraction section of the model can be used as an Encoder, while the classifier part can be removed, and the Decoder section can be added to allow the model to generate image captions. However, the dataset used for training must be analyzed by experts and annotated with image captions, and the dataset size should be much larger than what was used in this study.

B. Mobile Application Development

This study utilized the client/server concept to predict data through a mobile application that sends images to a server for prediction, requiring internet access. The CNN model used in this study was small in size, prompting the suggestion of developing a standalone mobile application that can predict results without sending data to a server and function offline, with modern smartphones' high processing speeds and GPU capabilities. The Keras CNN model used in this study is stored in an .h5 file and can be converted for iOS and Android using Core ML, integrated using Swift or Objective-C on Xcode for iOS, and TensorFlow Lite Android Support Library using Kotlin or Java for Android. Core ML is a suitable tool for developing an offline mobile application for dentists to diagnose periodontitis.

IX. CONCLUSION

In conclusion, our study successfully developed a deep learning model to diagnose periodontal disease using radiographic images, which was then integrated into both a web and mobile application. The model and application have the potential to improve the accuracy and efficiency of the diagnostic process, leading to early detection and treatment of periodontal disease and improved oral health outcomes. We emphasize that the dataset used in this study was analyzed and annotated by expert dentists from the Faculty of Dentistry, Mahidol University, ensuring the reliability and accuracy of the results.

REFERENCES

[1] Lecun, Y., L. Bottou, Y. Bengio and P. Haffner (1998). "Gradient-based learning applied to document recognition." Proceedings of the IEEE 86(11): 2278-2324.
[2] Lee, J.-H., D.-h. Kim, S.-N. Jeong and S.-H. Choi (2018). “Diagnosis and prediction of periodontally compromised teeth using a deep learning-based convolutional neural network algorithm.”
[3] Bhattacharjee, N. (2020). "Automated Dental Cavity Detection System Using Deep Learning and Explainable AI."
[4] Thanathornwong, B. and S. Suebnukarn (2020). “Automatic detection of periodontal compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks.” Imaging Sci Dent 50(2): 169-174.
[5] Thanh, M. T. G., N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap and D. M. Nguyen (2022). “Deep Learning Application in Dental Caries Detection Using Intraoral Photos Taken by Smartphones.” Applied Sciences 12(11).
[6] Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). “How transferable are features in deep neural networks?.” In Advances in neural information processing systems (pp. 3320-3328).
[7] Shorten, C., & Khoshgoftaar, T. M. (2019). “A survey on image data augmentation for deep learning.” Journal of Big Data, 6(1), 1-48.
[8] Perez, L., Wang, J., & Wang, J. (2017). “The effectiveness of data augmentation in image classification using deep learning.” arXiv preprint arXiv:1712.04621.
[9] Tieleman, T., & Hinton, G. (2012). “RMSProp: Divide the gradient by a running average of its recent magnitude.” COURSERA: Neural Networks for Machine Learning, 4(2), 26-31.
[10] Ruder, S. (2016). “An overview of gradient descent optimization algorithms.” arXiv preprint arXiv:1609.04747.