Development and validation of a visually explainable deep learning model for classification of C-shaped canals of the mandibular second molars in periapical and panoramic dental radiographs.

Sujin Yang, DDS, PhD, Hagyeong Lee, BS student, Byounghan Jang, MS student, Kee-Deog Kim, DDS, MSD, PhD, Professor, Jaeyeon Kim, BS, PhD student, Researcher, Hwiyoung Kim, MSD, PhD, Professor, Wonse Park, DDS, MSD, PhD, Professor

PII: S0099-2399(22)00277-1

DOI: https://doi.org/10.1016/j.joen.2022.04.007

Reference: JOEN 5030

To appear in: *Journal of Endodontics*

Received Date: 6 October 2021

Revised Date: 5 April 2022

Accepted Date: 5 April 2022

Please cite this article as: Yang S, Lee H, Jang B, Kim K-D, Kim J, Kim H, Park W, Development and validation of a visually explainable deep learning model for classification of C-shaped canals of the mandibular second molars in periapical and panoramic dental radiographs., *Journal of Endodontics* (2022), doi: https://doi.org/10.1016/j.joen.2022.04.007.

Development and validation of a visually explainable deep learning model for classification of C-shaped canals of the mandibular second molars in periapical and panoramic dental radiographs.

Sujin Yang*[1], Hagyeong Lee*[2], Byounghan Jang[3], Kee-Deog Kim[4], Jaeyeon Kim[5], Hwiyoung Kim†[6], Wonse Park†[7]

[1] DDS, PhD, Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea prismspects@gmail.com

[2] BS student, Department of Computer Science and Engineering, Ewha Womans University, Seoul, Korea, lhky0708@ewhain.net

[3] MS student, Department of Integrated Medicine, College of Medicine, Yonsei University, Seoul, Korea bhjang@yonsei.ac

[4] DDS, MSD, PhD, Professor; Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea kdkim@yuhs.ac

[5] BS, PhD student, Researcher; Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea jyeon@yuhs.ac

[6] MSD, PhD, Professor; Department of Radiology, Center for Clinical Imaging Data Science, Research Institute of Radiological Sciences, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea hykim82@yuhs.ac

[7] DDS, MSD, PhD, Professor; Department of Advanced General Dentistry, College of Dentistry, Yonsei University, Seoul, Korea wonse@yuhs.ac

*** These authors (Sujin Yang, Hagyeong Lee) contributed equally to this study.**

† **Co – corresponding authors.**

Co - corresponding author: Wonse Park, Department of Advanced General Dentistry, College of Dentistry,

Yonsei University. 50-1, Yonsei-ro, Seodaemun-gu, Seoul, 03722, Rep. of KOREA. Tel: +82-2-2228-8980, Fax: +82-2-2227-8906, E-mail: wonse@yuhs.ac

Co - corresponding author: Hwiyoung Kim, Department of Radiology, Center for Clinical Imaging Data Science, Research Institute of Radiological Sciences, Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea 03722, Rep. of KOREA. Tel: +82-2-2228-7397, E-mail: hykim82@yuhs.ac

Acknowledgement

The authors deny any conflicts of interest related to this study.

Development and validation of a visually explainable deep learning model for classification of C-shaped canals of the mandibular second molars in periapical and panoramic dental radiographs.

Running title: Deep learning classification of the C-shaped canal in mandibular second molars.

ABSTRACT

Introduction

The purpose of this study was to develop and validate a visually explainable deep learning model for the classification of C-shaped canals of the mandibular second molars in dental radiographs.

Methods

The periapical, panoramic images of 1000 mandibular second molars were collected from 372 patients. The diagnostic performance of the deep learning system using periapical and panoramic radiographs was investigated in respect to its ability to determine whether the second mandibular molar showed a C shaped canal configuration. The assessment of the canal configuration of CBCT volumes from 372 patients (740 mandibular second molars) was used as a gold standard.

Results

The deep Convolutional Neural Network (CNN) algorithm model showed high accuracy in predicting the C-shaped canal variation among mandibular second molars in both periapical and panoramic images. The model demonstrated best results when using image patches including only the root portion of the tooth and when using both periapical and panoramic images for training (AUC = 0.99). The model's diagnostic performance using only the root portion of the tooth (AUC; periapical = 0.98, panoramic = 0.95) was similar to a specialist (AUC; periapical = 0.95, panoramic = 0.96) and better than a novice general clinician (AUC; periapical = 0.89, panoramic = 0.91). Both the specialist and general clinician showed better diagnostic performance when reading panoramic radiographs compared to periapical images.

Conclusions

With further optimization of the test data using a larger dataset and improvements made in the model, a deep learning system may be expected to effectively diagnose C-shaped canals and aid clinicians in practice and education.

KEYWORDS

Deep Learning; Machine Learning; Computer vision; Convolutional Neural Networks; C-shaped canal; Mandibular second molar; Dental anatomy; Radiology.

INTRODUCTION

Knowledge of the root shape and root canal configuration is essential for the successful outcome of surgical and non-surgical root canal treatment. (1) (2) However, the anatomy of the root and root canal systems can be difficult to judge from a two-dimensional radiograph, and in many cases reading a periapical or panoramic radiograph requires experience and mastery by the clinician. (3) This is especially important in cases of the mandibular second molars which show a high prevalence of C-shaped root canals. Many studies have reported the prevalence of C-shaped canals to differ among different ethnic groups, with the prevalence among Asians reported to be the highest between 10-45%. (4) (5) (6) (7) The C-shaped canal variation of the mandibular second molar shows a complex canal system compared to the 3 canal system normally found in Caucasians (8) and can lead to challenges in canal debridement, disinfection and obturation. (9)

Since a complex anatomic feature such as the C-shaped canal renders periodontal and endodontic treatment along with maintenance a difficult task, it is important for the clinician to accurately evaluate the root and root canal system morphology. (10) For this matter cone-beam computed tomography (CBCT) is applied as the gold standard for evaluating such features. CBCT produces 3 dimensional (3D) images which are free of problems commonly found in conventional radiographs such as superimposition or image distortion and is reported to show high diagnostic accuracy similar to that obtained from conventional CTs while having lower radiation doses. (11) Nonetheless, since CBCT examinations still result in significantly higher radiation doses compared to conventional panoramic or periapical radiographs, they cannot be applied to all patients unless there is an indication for a requisite CBCT, e.g., an adjacent impacted third molar in need of extraction that shows proximity

to the inferior alveolar nerve.

In the recent decade artificial intelligence (AI) has actively been introduced and applied in dentistry. Convolutional Neural Networks (CNNs) have shown excelling performances in computer vision, yielding promising results in terms of detection and classification of certain diseases in the radiological and pathological field. The deep learning-based CNN models report high accuracy and efficiency and prove their potential to be implemented in various clinical situations such as tooth segmentation (12) or detection of diseases : periapical lesions (13), dental caries (14), periodontitis (15), osteoporosis (16), and TMJ osteoarthritis (17). Additionally, CNN models have become explainable through visualizing methods such as the application of GRAD-CAM, which visualizes class activation maps highlighting regions contributed to the model's decision. (18) Applying this deep learning system to the detection of root canal variations using conventional radiographs would prove effective for successful treatment. So far there have been only two studies using deep learning models to detect and classify C-shaped canal anatomies, one on panoramic images (19) and one on CBCT images (20). The studies have shown promising results in performance outcomes, but issues such as model robustness, training data processing, and explainability still need improvement. Accordingly, the purpose of this study was to develop and validate a visually explainable deep learning model for the classification of C-shaped canals of mandibular second molars in periapical and panoramic radiographs.

MATERIALS AND METHODS

1. Data collection

This study was approved by the institutional review board of Yonsei University Dental Hospital. (Approval number: 2-2020-0076). The images were retrospectively selected and collected from a database of patients who visited Yonsei University Dental Hospital, Department of Advanced General Dentistry between October 2018 and October 2020 for extraction of mandibular third molars. The patients had undergone a basic screening of panoramic radiographs and periapical radiographs in the second and third molar areas. An additional dental CBCT examination was taken for diagnosis of the impacted third molars. A total of 401 patients were included, and in the process 31 patients were excluded because of blurred images due to patient movement during radiograph taking, cropped images of teeth not showing the entire tooth and other outliers such as overlapped dental implants, orthodontic appliances, plates, and screws. The patient cohort consisted of 162 males and 208 females, with an

age range of 17-62 years (median 26 years; mean age 28.45 ± 9.24 years). 740 mandibular second molars were identified in total. The prevalence of C-shaped canals was 43.9% (325 out of 740).

CBCT volumes were taken in a standard upright position on RAYSCAN Alpha plus (RAYSCAN Alpha plus; Ray Co., Hwaseong, Korea) or Pax-Zenith 3D (Pax-Zenith; Vatech Co., Hwaseong, Korea) with the following parameters: scanning time, 14 seconds; field of view, 100 × 100 cm; tube voltage, 90 kVP; tube current, 12 mA; and voxel size, 0.18, based on the patient's size on the scanning device.

Panoramic images were obtained using Pax-i plus (Pax-I Plus; Vatech Co., Hwaseong, Korea) with the standard parameters, including a tube voltage of 73 kV, tube current of 9 mA, and acquisition time of 13.5 s. Periapical images were obtained under a tube voltage of 60kV, tube current of 7mA, with acquisition time of 0.125 seconds.

2. Ground truth annotation

The diagnostic performance of the deep learning system using periapical and panoramic radiographs was investigated in respect to its ability to determine whether the second mandibular molar showed a C shaped canal configuration. The results of the canal configuration on CBCT were used as the ground truth.

Image patches were prepared by first anonymizing the information of the patient's name, gender, age, and patient number. Then the image was cropped by drawing a bounding box from either the periapical images or the panoramic images to contain the crown and root of the mandibular second molar area or only the root portion. These cropped images were sorted as "Crown and root" or "Root only" groups. The images were then adjusted to a 300 × 400-pixel size for "Crown and root" or 300 × 300 pixel-size for "Root only", both with a resolution of 96 pixel per mm and saved in BMP (Bitmap) format image file using PhotoscapeX software. (ver.4.1 MOOII Tech.) A total of 1000 periapical and 1000 panoramic image patches were obtained from the 740 mandibular second molars. This was possible because some patients had periapical and panoramic radiographs taken before and after mandibular third molar extraction. 325 mandibular second molars, 440 patches from both periapical and panoramic images showed a C-shaped canal configuration.

The prepared images were then reviewed and classified into two categories: "nonC" referring to non-C shaped or "C" referring to C-shaped, by a specialist with expertise of over 7 years and using the CBCT data as ground truth (GT).

3. Data preparation

Ten-fold cross validation was performed to overcome deviations in the training datasets, mitigate overfitting and to determine hyperparameters (used to control the learning process) to result in lowest test errors. The datasets were randomly split to 10 portions, 9 being the training set and one being used as the testing set. Specifically, in the non C-shaped canal group 504 and 56 image patches per fold were used as training and testing datasets respectively, while in the C-shaped canal group, 396 and 44 image patches per fold were used as training and testing. Care was given so that each image was not mistakenly allocated into both training and testing groups.

Data augmentation techniques were applied to prevent overfitting owing to the small dataset size. Augmentation was performed with rotation, horizontal flip, or altering the contrast, brightness, and sharpness of the prepared images. The images were resized to (224,224) and a rotation angle from -20° to 20 ° in 0.5 ° increments and a gamma range from 0.7 to 1.3 in increments of 0.3 were used for the random rotation and gamma correction, respectively. Random translation was also added in a 20-pixel range so that the root tip will not be cut. The average and standard deviation of pixel values for the entire data were used for image normalization. The total number of images increased by 10 times.

4. CNN model training and evaluation

EfficientNet was utilized as the CNN model for this study due to its efficiency and high-performance for image classification. EfficientNet has been demonstrated to show high performance because the compound scaling method is used to scale up the accuracy, which uniformly scales all dimensions of depth, width, and resolution using a compound coefficient. (21) This allows for the reduction of training time and the number of images required to create a suitable classifier.

The model was optimized by a stochastic gradient descent method with the weight decay and momentum of 0.0001 and 0.9, respectively. Categorical cross-entropy was used as the loss function for training, and RMSprop was used as the optimizer. The training was performed for 400 epochs. The learning rate and the batch size was set to 0.00002 and the batch size was set to 32, respectively.

This model was implemented on a system with CPU Xeon(R) E5-2690 v4 @ 2.60GHz, RAM 252GB (Intel,

Santa Clara, California), and GPU Tesla V100 (NVIDIA, Santa Clara, California). Pytorch version 1.7.0 with CUDA 10.1 on Ubuntu 18.04 was used.

The accuracy, sensitivity (recall), specificity, positive predictive value (PPV; or precision), negative predictive value, and F1-score of the deep learning system were calculated in respect to the gold standard CBCT results. The area under the curve (AUC) was also evaluated from the receiver operating characteristic (ROC) analysis. Averages of the diagnostic performances from each of the 10 cross-validation folds were calculated, and these were regarded as the overall predictive performance of the model.

5. Comparison between CNN model and clinicians.

The diagnostic performance of the CNN model in classifying the C-shaped canal anatomy in mandibular second molars was compared with the performance of a specialist of over seven years of expertise and a general clinician with one year of experience. To simulate the clinical environment the original uncropped periapical and panoramic images were randomly chosen, and for each of the chosen periapical and panoramic images 100 teeth with C-shaped canals and 100 teeth with non-C-shaped canals were prepared for evaluation. The specialist and general clinician both evaluated the presence of a C-shaped canal from 200 periapical images (100 C-shaped, 100 non-C-shaped teeth) and 200 teeth in panoramic images (100 C-shaped, 100 non-C-shaped teeth). The accuracy, sensitivity (recall), specificity, positive predictive value (PPV; or precision), negative predictive value, F1-score and AUC of the observers were calculated and compared with those of the CNN model.

6. Visualization of CNN Model Decisions

To validate the decision-making process of the model and visualize which features are most important for the model to classify C-shaped canals, gradient-weighted class activation mapping technique (Grad-CAM) and Guided Grad-CAM were employed. The most significant regions for screening C-shaped canals in the image patches were highlighted as heatmaps (Grad-CAM) or fine-grained details (Guided Grad-CAM).

RESULTS

1.Diagnostic performance of EfficientNet

The results of the diagnostic performance of the CNN model on six subgroups were obtained respectively and compared. The results were primarily divided into two groups; by the image used containing both the coronal and root portion (“Crown and root group”) or only the root portion (“Root only group”). Then each group was divided by the source of the image data used; the source being from only the periapical images (“Periapical group”), only from the panoramic images (“Panoramic group”), or both images (“All images group”).

The AUC of the “Crown and root group” was 0.95 in the “Periapical group” 0.93 in the “Panoramic group” and 0.95 in the “All images group”. The AUC of the “Root only group” was 0.98 in the “Periapical group” 0.95 in the “Panoramic group” and 0.99 in the “All images group”. The sensitivity of the “Crown and root group” was 0.93 (Periapical), 0.72 (Panoramic), 0.90 (All images) and the “Root only group” was 0.88 (Periapical), 0.85 (Panoramic), 0.97 (All images). The specificity of the “Crown and root group” was 0.87 (Periapical), 0.93 (Panoramic), 0.91 (All images) and the “Root only group” was 0.95 (Periapical), 0.88 (Panoramic), 0.93 (All images).

The accuracy, sensitivity (recall), specificity, positive predictive value (PPV; or precision), negative predictive value, F1-score, and time for training of the model for each group are shown in Table 1. The classification confusion matrices are shown in Figure 1. The ROC curves and AUC of each group are shown in Figure 2.

2. Comparison of diagnostic performance between a specialist, general clinician and the AI.

The results of the comparison in performance between the AI and clinicians (specialist and general clinician) are also shown in Table 1. When reading periapical radiographs the specialist and general clinician showed differences in performance: specifically, accuracy (0.95 : 0.89) , sensitivity (0.95 : 0.91) , specificity (0.94 : 0.87), positive predictive value (0.94 : 0.86), negative predictive value (0.95 : 0.92) F1 score (0.94:0.89) and AUC (0.95 : 0.89). Differences in all performance metrics were also present between the specialist and general clinician when reading panoramic images: specifically, accuracy (0.96 : 0.91) , sensitivity (0.97 : 0.93) , specificity (0.95 : 0.89), positive predictive value (0.95 : 0.89), negative predictive value (0.97 : 0.93) F1 score (0.96:0.91) and AUC (0.96 : 0.91). Both clinicians showed higher false positives (predicting the canal as “C-shaped” whereas it is not) than false negatives (predicting the canal as “non C-shaped” whereas it is not). The clinicians’ overall diagnostic performance was higher when reading panoramic radiographs, which is a difference compared to the AI’s performance. When comparing the AUC calculated from the ROC curves, the specialist’s performance was similar

to the CNN model while the general dentist showed inferior scores compared to the specialist and the AI. However, when the CNN model used image patches containing only the root portion from both periapical and panoramic radiographs for training, better performance outcomes were identified in several performance metrics.

3. Explainability using Grad-CAM and Guided Grad-CAM

To compare the model performance based on clinical knowledge and add explainability to the model, Grad-CAM and Guided Grad-CAM was applied to show visualized class activation map (CAM)s upon the crown or root within the radiographic image. Most CAMs highlighted the root portion area, and especially the apical third of the root. However, in some cases the highlights appeared in Cemento-Enamel Junction (CEJ) areas and such misinterpretations were more frequently found in the crown and root group. The examples of Grad-CAM and guided Grad-CAM of the "C" and "nonC" from "Crown and root" or "Root only" images obtained from periapical or panoramic radiographs are shown in Figure 3.

DISCUSSION

Contrary to the fact that deep learning systems are actively applied and implemented to various medical fields, they have not yet taken place in routine dental practice. This is a surprising result when acknowledging that dentistry is suited to apply AI tasks due to its abundance in image data of a particular anatomic region from different time points, along with the prevalence of dental diseases. (22) When deep learning is applied to the endodontic field it can likewise prove as a powerful tool to diagnosis and can aid in endodontic treatment and follow-up management. Specifically, if early detection of the root canal morphology is possible before endodontic treatment, dentists can preemptively identify those patients who require meticulous canal debridement and sealing, ask for an additional CBCT taking to evaluate the specific root canal morphology (since it is not possible to perform CBCTs on all patients undergoing root canal treatment), and warn the patient of possible outcomes due to the complexity of the root and canal anatomy.

In this study, the diagnostic performance from using only the periapical radiographs was higher than that from the panoramic radiographs while using both images resulted in the best performance. Periapical radiographs generally have better image quality and are frequently taken in the clinical background to focus on a particular tooth number. The periapical radiographs used in this study were not taken in a fixed angle because they were taken before or after mandibular third molar extraction where the third molar is put upon the center of the

radiograph rather than the second molar. However, this did not appear to hinder the diagnostic performance of the model. Panoramic images also show uncertainties due to image magnification or ghost images. The results of this study imply that the accumulated image distortion was greater in the panoramic radiographs. In some clinical situations, panoramic images can be more helpful to the clinician or radiologist when detecting the C shaped canal of the mandibular second molar because there is less rotation of the tooth when the image is taken, alongside the fact that the opposite side tooth can be compared. This coincides with the results shown from the clinicians where better diagnostic results were obtained from panoramic images. However, this did not act upon the model's performance since the image was cropped in advance. Nevertheless, this should be regarded in further studies where the mandibular second molar is first detected in the panoramic radiograph and then classified as whether it has a C-shaped canal variation. Best results were found when images from both panoramic and periapical radiographs were used in training and testing. We believe that this is the outcome of regularization from the intensity and resolution difference of the periapical and panoramic images. The periapical images have higher resolution and therefore show better results compared to panoramic images, but panoramic images have a magnifying effect compared to periapical images. Therefore, when all images are used it will act as if augmentation in views, intensity and resolution was made resulting in a more robust model and better performances. (23) Also, the performance was higher when only the root portion of the tooth was used for training. This is presumably the aftermath of overlapping of various prostheses in the crown portion such as inlays or crowns in image patches where the crown and root portions are both present. Without the crown portion we could minimize the chances in which the anatomy or prosthesis of the crown portion interferes with the reading of the root portion or chances the CNN model mistakenly focuses on the coronal portion of the tooth for root canal classification. (24) In further studies, by adding an object detection model that can detect and crop the area of the second mandibular molar root to the classification algorithm, we may be able to build a more robust and efficient model fit for this task.

In this study both Grad-CAM and guided Grad-CAM were used to add explainabilty to the model. Grad-CAM is made by using the feature maps produced by the last convolutional layer of the CNN. Grad CAM visualizations are class-discriminative and presented as generally known heatmaps localizing relevant image regions, but do not highlight the fine-grained pixel importance. Guided Grad-CAM is known to combine Grad-CAM and Guided Backpropagation via element-wise multiplication, which visualizes gradients with respect to the image where

negative gradients set to zero to highlight import pixel in the image when backpropagating through ReLU layers. Thus, the results of Guided Grad-CAM results are presented as class discriminative highlighted outlines with high resolution, while Grad-CAM shows broad heatmaps. Most Grad-CAM and Guided Grad-CAM results of this study highlighted the apical third of the root where the canals showed convergence as in most C-shaped canals, which is consistent with the clinicians' interpretation. However, some highlighted other areas such as the cemento-enamel junction or pulp chamber area. This may be because in most C-shaped canal cases the canal orifices tend to be closer to each other, giving in some cases a distinct pulp shape in the cervical third area compared to non-C shaped canal cases. There were no heat maps located outside of the tooth area, allowing us to confirm that the algorithm has used safe and meaningful prediction strategies. (24) Heat maps focused upon areas apart from the apical third of the root were more frequently shown in the radiographic type II C shaped canal following the classification of Fan et al. (25) This is because the type II C shaped canal shows that the mesial and distal canals assume their own individual course to the apex, thus looking similar to a non-C-shaped case in the periapical or panoramic radiographs. These misclassified heat maps also imply overfitting of the model. More data is needed to ensure increased accuracy and transparency. On the other hand, while the results of Guided Grad-CAM of this study showed more detailed highlights, there are suggestions that guided backpropagation performs partial image recovery and acts like an edge detector, rather than providing insight into a trained model. (26) Also, there are suggestions that the saliency map derived from guided grad-CAM is not related to the model's parameters or the training data and therefore does not clearly build trust to deep learning models. (27) Nonetheless, Guided Grad-CAM was adopted as a visualization method in this study to obtain fine-grained highlights rather than broad heatmaps and check the detailed position on where the model was focusing. As more visually explainable methods are being suggested future DL models will need to implement these methods to build trust and ensure its adoption.

There are notable limitations in this study. First, the sample size used in this study was relatively small compared to other related studies. This was because the radiographic images were obtained from patients who had visited the dental clinic for mandibular third molar extraction and thus the patients enrolled for data collection were limited. The number of images prepared in this study was multiplied after data augmentation; nonetheless, the quantity of training dataset used in this study may still be considered insufficient. Although it is known that deep CNN algorithms such as EfficientNet are highly useful in cases where available training sets are limited, small datasets can be a bottleneck to further advancement of computer aided detection. (28) Second, because periapical and panoramic images obtained from patients before and in some, after mandibular third molar extraction, the

issue of duplicate data remains. Duplicates act as non-random sampling, and may bias the fitted model. More data must be added to solve this problem. Thirdly, data homogeneity was present because the data was collected from one institution. Although pre-trained deep CNN models are effective for general image classification owing to their pre-trained weights, the possibility of overfitting cannot be excluded with a small dataset with a singular character coming from a limited data source. External validation is needed, and multicenter studies are necessary for increased accuracy of DL classification and diagnosis. Fourthly, because we handpicked EfficientNet among many available deep learning algorithms, it is important to keep up with the development of other CNN algorithms. A different model using the same dataset shows different diagnostic outcomes because of variations in the CNN structure. New CNN algorithms differing in depth, width, or modified stratification methods are continually being developed, and it is important to consider these progressions in future studies. Lastly, the root figures were classified into only two categories: C-shaped or non-C shaped. However, the root or root canal morphology of the mandibular second molar can be subdivided to many categories. The non-C shaped canals can have either 2 or 3 roots, and although not commonly found, even C shaped canals can be present in a two rooted second molar, or a single rooted second molar can have a conical or O-shaped canal instead of the prevalent C-shaped canal. (29) Such variables, along with age, sex, or population-related morphological variations may possibly be implemented for advances in accurate diagnosis and classification.

In this study EfficientNet could accurately classify a mandibular second molar having a C-shaped canal morphology. Within the limitations of this study, we have developed and validated of a visually explainable deep learning CNN model for classification of C-shaped canals of the mandibular second molars in dental radiographs. Further studies are needed, but the results of this study may play an important role in C-shape canal prediction in the clinical and educational field of dentistry.

ACKNOWLEDGEMENTS

The authors declare no potential conflicts of interest with respect to the authorship and/or publication of this article.

DATA VAILABILITY STATEMENT

The authors confirm that the data supporting the findings of this study are available within the article [and/or] its

supplementary materials.

REFERENCES

1. Cooke HG, 3rd, Cox FL. C-shaped canal configurations in mandibular molars. J Am Dent Assoc. 1979;99:836-9.
2. Vertucci FJ. Root canal anatomy of the human permanent teeth. Oral Surg Oral Med Oral Pathol. 1984;58:589-99.
3. Endres MG, Hillen F, Salloumis M, Sedaghat AR, Niehues SM, Quatela O, et al. Development of a Deep Learning Algorithm for Periapical Disease Detection in Dental Radiographs. Diagnostics (Basel). 2020;10.
4. Zhang R, Wang H, Tian YY, Yu X, Hu T, Dummer PM. Use of cone-beam computed tomography to evaluate root and canal morphology of mandibular molars in Chinese individuals. Int Endod J. 2011;44:990-9.
5. Gulabivala K, Aung TH, Alavi A, Ng YL. Root and canal morphology of Burmese mandibular molars. Int Endod J. 2001;34:359-70.
6. Haddad GY, Nehme WB, Ounsi HF. Diagnosis, classification, and frequency of C-shaped canals in mandibular second molars in the Lebanese population. J Endod. 1999;25:268-71.
7. Helvacioglu-Yigit D, Sinanoglu A. Use of cone-beam computed tomography to evaluate C-shaped root canal systems in mandibular second molars in a Turkish subpopulation: a retrospective study. Int Endod J. 2013;46:1032-8.
8. Plotino G, Tocci L, Grande NM, Testarelli L, Messineo D, Ciotti M, et al. Symmetry of root and root canal morphology of maxillary and mandibular molars in a white population: a cone-beam computed tomography study in vivo. J Endod. 2013;39:1545-8.
9. Manning SA. Root canal anatomy of mandibular second molars. Part II. C-shaped canals. Int Endod J. 1990;23:40-5.
10. Torabinejad M, Kutsenko D, Machnick TK, Ismail A, Newton CW. Levels of evidence for the outcome of nonsurgical endodontic treatment. J Endod. 2005;31:637-46.
11. Scarfe WC, Farman AG. What is cone-beam CT and how does it work? Dent Clin North Am. 2008;52:707-30, v.
12. Lahoud P, EzEldeen M, Beznik T, Willems H, Leite A, Van Gerven A, et al. Artificial Intelligence for Fast and Accurate 3-Dimensional Tooth Segmentation on Cone-beam Computed Tomography. J Endod. 2021;47:827-35.
13. Setzer FC, Shi KJ, Zhang Z, Yan H, Yoon H, Mupparapu M, et al. Artificial Intelligence for the Computer-aided Detection of Periapical Lesions in Cone-beam Computed Tomographic Images. J Endod. 2020;46:987-93.
14. Lee JH, Kim DH, Jeong SN, Choi SH. Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm. J Dent. 2018;77:106-11.

15. Krois J, Ekert T, Meinhold L, Golla T, Kharbot B, Wittemeier A, et al. Deep Learning for the Radiographic Detection of Periodontal Bone Loss. Sci Rep. 2019;9:8495.

16. Lee KS, Jung SK, Ryu JJ, Shin SW, Choi J. Evaluation of Transfer Learning with Deep Convolutional Neural Networks for Screening Osteoporosis in Dental Panoramic Radiographs. J Clin Med. 2020;9.

17. Lee KS, Kwak HJ, Oh JM, Jha N, Kim YJ, Kim W, et al. Automated Detection of TMJ Osteoarthritis Based on Artificial Intelligence. J Dent Res. 2020;99:1363-7.

18. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization2016 October 01, 2016:[arXiv:1610.02391 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2016arXiv161002391S.

19. Jeon SJ, Yun JP, Yeom HG, Shin WS, Lee JH, Jeong SH, et al. Deep-learning for predicting C-shaped canals in mandibular second molars on panoramic radiographs. Dentomaxillofac Radiol. 2021;50:20200513.

20. Sherwood AA, Sherwood AI, Setzer FC, K SD, Shamili JV, John C, et al. A Deep Learning Approach to Segment and Classify C-Shaped Canal Morphologies in Mandibular Second Molars Using Cone-beam Computed Tomography. J Endod. 2021;47:1907-16.

21. Tan M, Le QV. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks2019 May 01, 2019:[arXiv:1905.11946 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2019arXiv190511946T.

22. Schwendicke F, Samek W, Krois J. Artificial Intelligence in Dentistry: Chances and Challenges. J Dent Res. 2020;99:769-74.

23. Chlap P, Min H, Vandenberg N, Dowling J, Holloway L, Haworth A. A review of medical image data augmentation techniques for deep learning applications. J Med Imaging Radiat Oncol. 2021;65:545-63.

24. Lapuschkin S, Wäldchen S, Binder A, Montavon G, Samek W, Müller KR. Unmasking Clever Hans predictors and assessing what machines really learn. Nat Commun. 2019;10(1):1096.

25. Fan B, Cheung GS, Fan M, Gutmann JL, Fan W. C-shaped canal system in mandibular second molars: Part II--Radiographic features. J Endod. 2004;30:904-8.

26. Adebayo J, Gilmer J, Muelly M, Goodfellow I, Hardt M, Kim B. Sanity Checks for Saliency Maps2018 October 01, 2018:[arXiv:1810.03292 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2018arXiv181003292A.

27. Nie W, Zhang Y, Patel A. A Theoretical Explanation for Perplexing Behaviors of Backpropagation-based Visualizations2018 May 01, 2018:[arXiv:1805.07039 p.]. Available from: https://ui.adsabs.harvard.edu/abs/2018arXiv180507039N.

28. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. IEEE Trans Med Imaging. 2016;35:1285-98.

29. Kim SY, Kim BS, Kim Y. Mandibular second molar root canal morphology and variants in a Korean subpopulation. Int Endod J. 2016;49:136-44.

TABLES

	CNN model performance						Specialist performance		General clinician performance	
	Crown and root group			Root only group						
	Periapical (N = 740 teeth)	Panoramic (N = 740 teeth)	All images (N = 740 teeth)	Periapical (N = 740 teeth)	Panoramic (N = 740 teeth)	All images (N = 740 teeth)	Periapical (N = 200 teeth)	Panoramic (N =200 teeth)	Periapical (N =200 teeth)	Panoramic (N = 200 teeth)
Accuracy	0.90	0.85	0.91	0.91	0.86	0.95	0.95	0.96	0.89	0.91
Sensitivity (Recall)	0.93	0.72	0.90	0.88	0.85	0.97	0.95	0.97	0.91	0.93
Specificity	0.87	0.93	0.91	0.95	0.88	0.93	0.94	0.95	0.87	0.89
Positive Predictive Value (Precision)	0.90	0.87	0.93	0.94	0.89	0.94	0.94	0.95	0.86	0.89
Negative Predictive Value	0.91	0.84	0.88	0.89	0.83	0.97	0.95	0.97	0.92	0.93
F1-score	0.91	0.79	0.92	0.91	0.87	0.95	0.94	0.96	0.89	0.91
AUC	0.95	0.93	0.95	0.98	0.95	0.99	0.95	0.96	0.89	0.91

Table1. Prediction performance of the deep CNN model (EfficientNet) on periapical, panoramic, or both radiographs of the Crown and root group or Root only group; and the performance of a specialist of over seven years of expertise and a general clinician of one year experience.

FIGURE LEGENDS

Figure1. Multiclass classification confusion matrixes drawn from the results. The diagonal elements are the number of points where the predicted label matches the true label, while the non-diagonal elements show the wrong detections made by the classifier. The higher the diagonal value and the darker the shade of peach, the more accurate the diagnosis.

Figure2. ROC curves for each of the groups obtained using periapical, panoramic, or both radiographs from the crown and root group or root only group.

Figure3. Radiograph examples and activation maps from the periapical or panoramic radiograph, from the crown and root group or root only group. The radiographic images were normalized, then Class Activation Map (CAM)s were drawn by applying Grad-CAM, and guided Grad-CAM. Heatmaps are focused in the apical one third portion of the tooth; however, in some cases the additional heatmaps are positioned near the pulp chamber area. (Note the heatmap focusing on the coronal one third on the panoramic radiograph example of the C shaped canal from the "Crown and root group".)

Figure 1

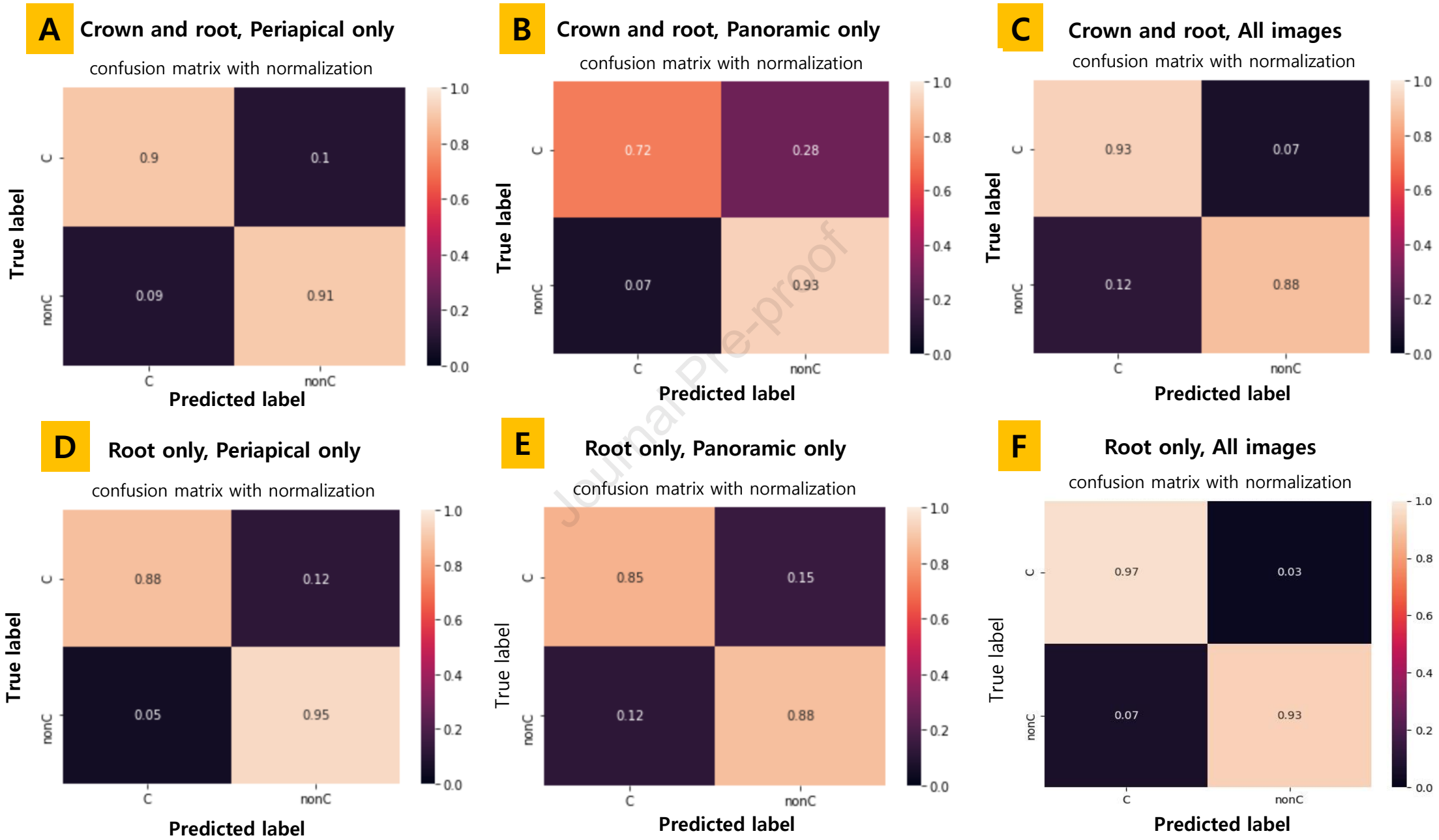
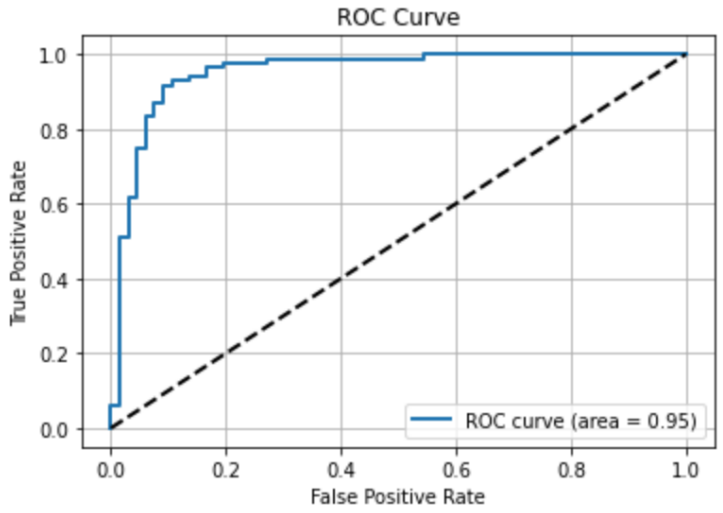
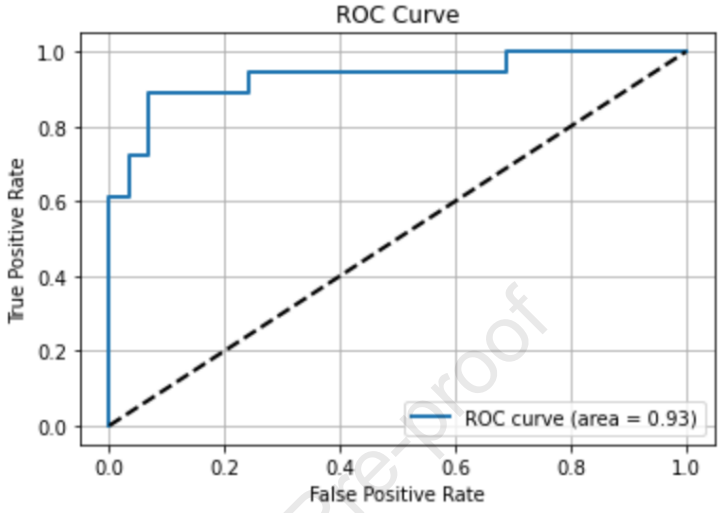


Figure 2

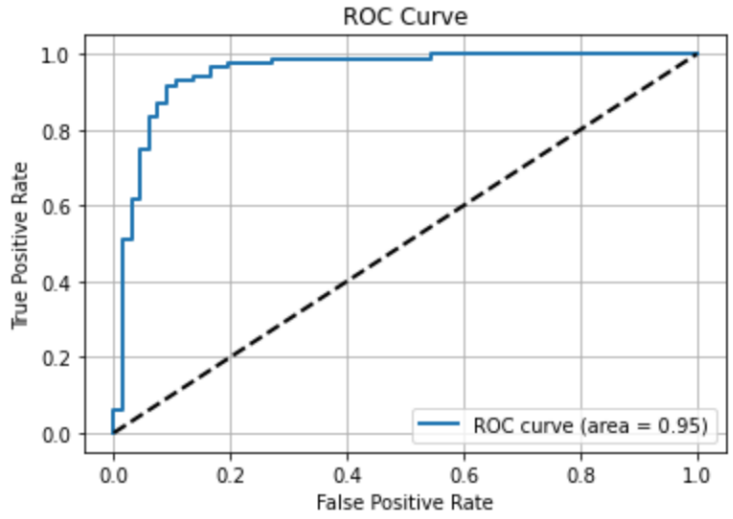
A Crown and root, Periapical only



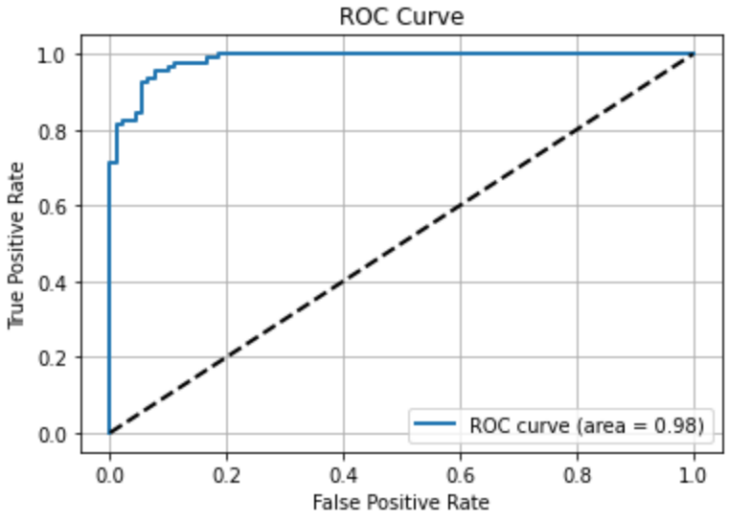
B Crown and root, Panoramic only



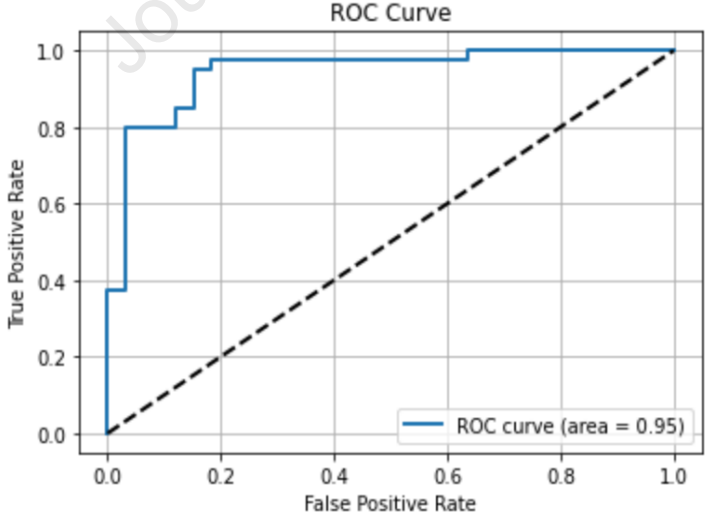
C Crown and root, All images



D Root only, Periapical only



E Root only, Panoramic only



F Root only, All images

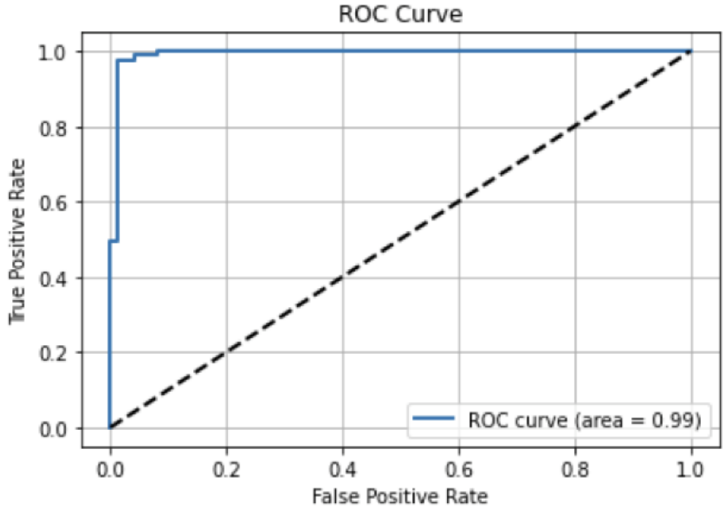


Figure 3