ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

A unified convolution neural network for dental caries classification

Wattanapong Suttapak[1], Wannakamon Panyarak[2], Dauangporn Jira-apiwattana[3] and Kittichai Wantanajittikul[4]

ABSTRACT

Dental caries is one of the most common chronic diseases in the oral cavity. The early detection of initial dental caries is needed for treatment. It is problematic to diagnose the initial carious lesion, as known as enamel caries, due to the similarity of a tiny hole to human perception error. In this paper, we propose a unified convolution neural network to improve the diagnostic and treatment performance for dentists using classification from bitewing radiographs. We adapt the AlexNet and ResNet models to properly classify the dental caries dataset. The modified ResNet successfully achieves excellent binary-classification performance with accuracy of 86.67%, 87.78% and 82.78% of teeth with all conditions, teeth without dental restoration, and only teeth with dental restorations, respectively. For multilevel classification, our model has good performance with 5-class average accuracy of 80%. Remarkably, our adapted ResNet-18 has good performance with enamel caries and secondary caries with accuracy of 86.67% and 77.78%, respectively. Conversely, our ResNet-50 and ResNet-101 have contradictory low performance with enamel and secondary caries but high performance with sound teeth, dentin caries and teeth with restoration of 90%, 78.89% and 88.89%, respectively. The accuracies of our model are good enough that our model could support dentists to enhance diagnostic performance.

Article information:

Keywords: Dental Caries Classification, Proximal Caries Classification, Secondary Caries Classification, Bitewing Radiograph, Convolution Neural Network

Article history:
Received: October 14, 2021
Revised: December 6, 2021
Accepted: March 5, 2022
Published: June 4, 2022 (Online)

DOI: 10.37936/ecti-cit.2022162.245901

1. INTRODUCTION

Dental caries, commonly named as tooth decay, is one pandemic diseases that has been highly prevalent for several decades [9]. The early detection of initial dental caries could lead to non-invasive dental procedures saving cost in the treatment. Nevertheless, solely clinical examination might be inadequate for some lesions such as proximal caries or carious lesion under restoration. The combination of visual and tactile inspection, together with bitewing radiographs are the gold standard of diagnostic tools [1].

However, there are limitations in the reliability and validity of diagnosis results between dentists and a proportion of false positive or false negative detec-

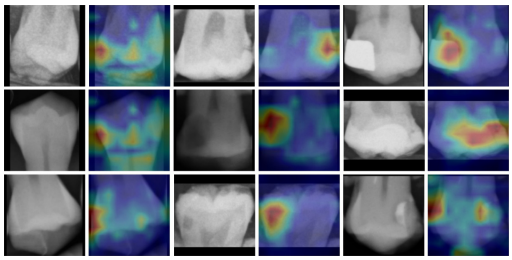


Fig. 1: *Each row is the example of dental caries heatmap. The first, third and fifth column are original bitewing radiograph of enamel caries, dentin caries and secondary caries, respectively. The following second, fourth and sixth columns are blending of heatmap and original one of first, third and fifth column, respectively.*

[1]The author is with Division of Computer Engineering, School of Information and Communication Technology, University of Phayao, Phayao, Thailand 56000, E-mail: wattanapong.su@up.ac.th
[2]The author is with Division of Oral and Maxillofacial Radiology, Faculty of Dentistry, Chiang Mai University, Chiang Mai, Thailand 50200, E-mail: wannakamon.p@cmu.ac.th
[3]The author is with Dental Department, Phayao Hospital, Thailand 56000, E-mail: jira_apiwattana@outlook.com
[4]The author is with Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand 50200, E-mail: kittichai.wan@cmu.ac.th
Corresponding author: kittichai.wan@cmu.ac.th

tions in radiographic interpretation [5, 8, 10], especially for initial carious lesions such as enamel caries [16]. The chance of precisely detecting proximal caries lesion by experienced dentists are approximately four times greater than with less experienced dentists [10]. According to systematic reviews of detecting carious lesions, the false negative detections occured more often in radiographic interpretations for both initial and advanced lesions [12, 26] and secondary caries [2, 14, 22].

Currently, deep neural networks are widely used in biomedical image processing including for dental caries diagnosis [4, 11, 21, 25, 31] Examples, include using near-infrared light transillumination images, panoramic X-ray images, and bitewing radiographs with deep learning neural network. K. Devito et al. [7] implemented to multilayer perceptron neural network for diagnosis of proximal dental caries from bitewing radiographs. The ROC curve result found that the perceptron model was superior to overall examiners as 0.884 and 0.717, respectively. M. Srivastava [28] developed a Computer Aided Diagnosis (CAD) system to assist dentists in detecting dental caries from 3000 bitewing radiographs. They implemented 100 layers of deep convolution neural network to classify the dental carious lesions. The experimental result found that their implemented deep learning model outperformed three dentists. J. Lee et al. [18] used GoogleNet Inception V3 [30] to detect and diagnose dental carious lesions. The diagnostic of both premolar and molar models provided performance by accuracy of 82.0% and the AUC of ROC curve of 0.845. A. Cantu et al. [3] proposed the segmentation model of carious lesions on bitewing radiographs using U-Net [24] and challenged the experts with the deep neural network results. The experimental result of their modified U-Net model was excellent and significantly more accurate than dentists. L. Lian et al. [19] used U-Net [24] to detect the carious lesion panoramic films and DenseNet121 [15] to classify multilevel of carious lesion panoramic radiographs. M. Moran et al. [20] invented a technique using image processing and a deep convolution neural network to classify lesion severity of proximal cavity in bitewing radiographs. The Inception [29] and ResNet [13] were used to compare learning rate. They suggested that their models could assist dentists to diagnose lesion severity in bitewing radiographs.

As previously mentioned, we found that other existing dental caries classification methods could enhance the classification ability and identified the proper model for each specific dental carious level including even for teeth with dental restoration. In this paper, we propose a unified convolution neural network for dental caries classification that modifies the fully connected layer at the end of the backbone model to fit with the number of our dental caries classes and train it in an end-to-end fashion. Moreover, we analyze the classification heatmap by implementing GradCam [27] which is able to find the correct location of dental caries as illustrated in Fig.1. Our contributions can be summarized as follows;

2. RELATED WORK

In this section, we describe the background of generalization in deep learning and the application of dental caries classification.

2.1 CNN Feature Extraction

For the generalization of image dataset, AlexNet [17] won the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) [6] in 2012. The network architecture of AlexNet is illustrated in Fig. 2. This model can classify 10,184 categories and 8.9 million images at 67.4% and 40.9% at top-1 and top-5 error rates, respectively.

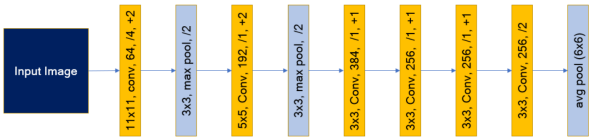


Fig.2: *Network structure of AlexNet.*

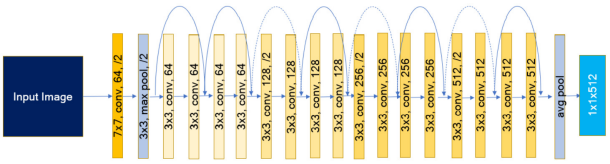


Fig.3: *Network structure of ResNet-18.*

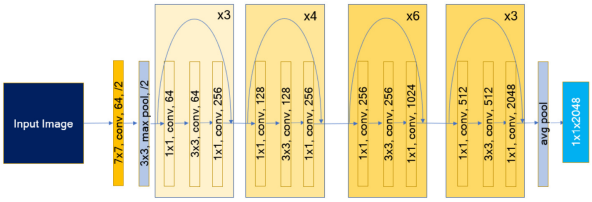


Fig.4: *Network structure of ResNet-50.*

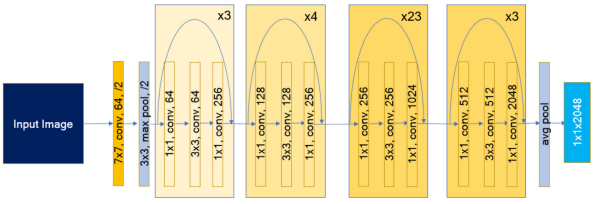


Fig.5: *Network structure of ResNet-101.*

In 2016, ResNet [13] was invented by leveraging the residual-learning block concept. This concept presents the shortcut connection for each layer to

learn the identity mapping which can solve the degradation problem of its shallower counterpart. This deep convolution neural network model was presented as 5 different models: ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. The number at the end of each ResNet model represents the number of convolution layers. The larger model significantly enhances the complicated feature extraction from the image dataset. In this paper, we choose 3 ResNet models for classifying the dental caries. The network architectures of ResNet-18, ResNet-50 and ResNet-101 are illustrated in Fig. 3, 4 and 5, respectively.

2.2 Image Classification

Beforehand, deep convolution neural network model can classify the image dataset. We required the feature extractions which are based on AlexNet [17] and ResNet [13] as explained in Section 2.1. These features are fed into the fully connected layer whose network complexity depends on the aforementioned feature extractor. For example, Alexnet's features are fed into three fully connected layers as shown in Fig. 6. The number of classes is 1000 ImageNet-object classes [6] for both AlexNet and ResNet.

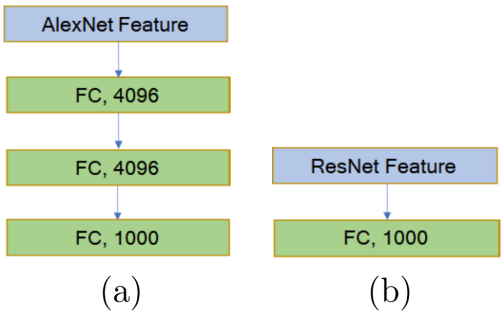


Fig.6: *Fully connected layer of (a) AlexNet and (b) ResNet.*

2.3 Dental Caries Classification

A systematic review found that, based on 117 studies and using data from 13,375 teeth, the detection of both initial and advanced lesions had a mean specificity that ranged between 0.89 and 0.97, while the mean sensitivity ranged between 0.24 and 0.42[26]. Likewise, the detection of secondary caries had a mean specificity ranging between 0.78 and 0.83, while the mean sensitivity ranged between 0.50 and 0.59[2]. According to this information, the false negative detections occur more often in radiographic interpretations. Computer assisted diagnosis should be implemented to support dentists.

In 2018, K. Devito et al. [7] proposed using artificial intelligence to improve the radiographic diagnosis of proximal caries. They implemented a multilayer perceptron neural network for diagnosis of proximal dental caries from bitewing radiographs. The comparison of ROC curve area between 25 examiners and the proposed model found that the perceptron model is superior to overall examiners as 0.884 and 0.717, respectively. However, teeth with restorations or occlusal caries were not included in the dataset. J. Lee et al. [18] applied a deep convolution neural network to detect and diagnose dental caries. The model implementation was based on GoogleNet Inception V3 [30]. They selected 3000 periapical radiographic images which consisted of 2400 and 600 images for training and testing, respectively. The diagnostic accuracies of both premolar and molar models were 82.0% while the AUC of ROC curve is 0.845. However, the selected radiographic images did not include occluded noise, haziness, distortion, and shadows. A. Cantu et al. [3] proposed the segmentation of carious lesions on bitewing radiographs. They segmented the radiographic dataset using U-Net [24] to compete with dentists. The experimental result was excellent and significantly more accurate than dentists. Nevertheless, the evaluated metrics were obscure due to general segmentation metrics could not express sensitivity and specificity without mean average precision (∼mAP) and intersection over union (∼IOU).

In this paper, we propose a method for the classification of dental caries which can be separated into binary classification and multilevel lesion classification. The binary classification dataset is comprised of sound teeth and carious lesions while multilevel lesion classification dataset includes sound teeth, enamel carious teeth, dentin carious teeth and teeth with dental restoration with and without carious lesion. Both experiments are explained in the next Section.

3. METHODOLOGY

3.1 Dental Dataset

This study was approved by the ethical review board of the Faculty of Dentistry, Chiang Mai University (approval no.22/2021). The bitewing radiographs were retrieved from the PACs system from January 2015 to December 2019. All radiographic data in this study were taken with intraoral image receptors (Carestream 7600 Phosphor Plate; Carestream, Rochester, NY, USA) using intraoral x-ray units (Haliodent Plus; Dentsply Sirona, Charlotte, NC, USA) in the Oral and Maxillofacial Radiology Clinic. All bitewing radiographs of all permanent teeth were chosen, regardless of patient's age, gender, or clinical information.

An oral and maxillofacial radiologist with 7 years of experience transferred the radiographic data in a Portable Network Graphics format (PNG) to a personal computer without image adjustment. The radiologist cropped each tooth from each radiograph and stored it in the PNG format. Subsequently, the cropped images were categorized as follows: sound tooth, enamel carious tooth, dentin carious tooth, tooth with dental restoration, and carious lesion un-

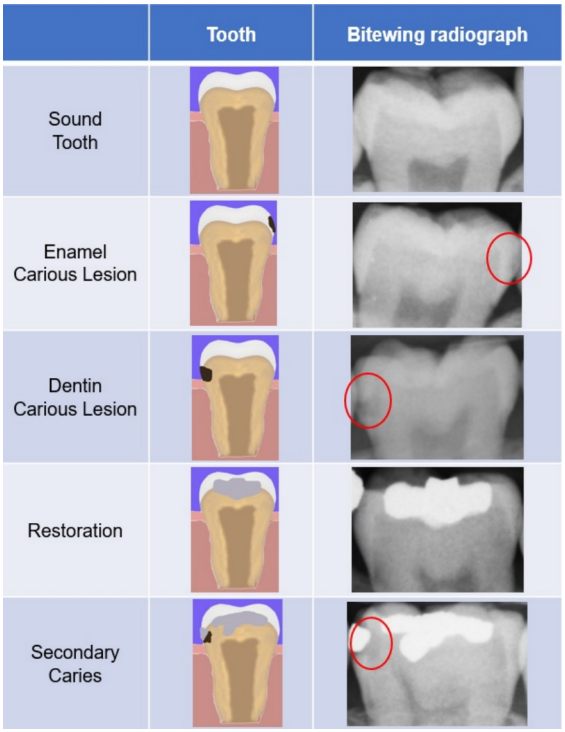
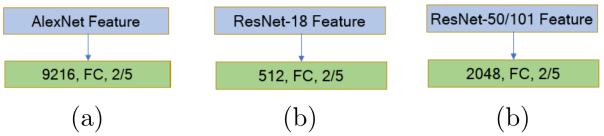


Fig.7: *Dental images and comparative bitewing radiographs, sound tooth, enamel carious lesion, dentin carious lesion, restoration and secondary caries are represented respectively from top to bottom row. Noticeably, the gray color represents the restoration material and dark brown represents the carious lesion.*

der restoration as illustrated in Fig. 7 in which the second column represents tooth model and the third column represents bitewing radiograph. We further add brief definitions of each term for easier understandings as follows;
- Healthy tooth (sound tooth) refers to tooth which appears to be in a good physical condition. Neither tooth decay nor dental fillings are shown in the radiograph.
- Enamel carious lesion refers to tooth decay which radiographically appears as a radiolucent lesion (a small dark area within the opaque region) limited in the enamel layer (the outermost layer of tooth structure).
- Dentin carious lesion refers to tooth decay which radiographically appears as a radiolucent (dark) lesion extending from the outer surface to the dentin layer (the middle layer between enamel layer and the dental pulp of tooth structure).
- Restoration, as known as dental fillings, refers to dental materials used for a treatment of dental caries. Radiographically, it appears to be opaquer than the enamel or dentin and its margin should be definite without a radiolucent (dark) area underneath.
- Secondary caries refers to a carious lesion occurring beneath the restoration. Radiographically, there is an ill-defined radiolucent area under the radiopacity



(a) (b) (b)

Fig.8: *Our modified fully connected layer of (a) AlexNet, (b) ResNet-18, (c) ResNet-50, and ResNet-101.*

of the dental restoration.
- Proximal caries refers to a tooth decay occurring at the area between two adjacent teeth.

In this paper, we categorize our dataset by cross sharing between binary classification and multilevel classification. Binary classification of sound teeth and dental caries is comprised of three sub-experiments: overall tooth conditions, without dental restoration, and only dental restoration. Each class of the binary classification of overall tooth conditions consists of 1020 images for training and 180 images for testing. Both sub-experiments of teeth without dental restoration and only teeth with dental restorations are fairly shared from dental images in the first sub-experiment. There are equally 510 sound teeth and dental caries images for training and 90 sound teeth and dental caries images for testing. In the multiple classification experiment, we selected 2550 training images and 450 testing images. We equally separated 510 training images and 90 testing images for each class.

3.2 The Classification Models

As explaned in Section 2.2 , the default AlexNet and ResNet models were created to classify 1000 ImageNet-object classes whereas our dental-caries dataset only has 2 and 5 classes in binary classification and multilevel lesion classification, respectively. We had to modify the classification layer of both AlexNet and ResNet from 1000 to 2 and 5. The AlexNet fully connected layer is reduced to only one layer and the ResNet is shortened likewise. We removed the excessive layers which eliminated the overtraining problem for small classes. The fully connected layer of Alexnet has input size as 9216 dimensions while the input size of ResNet-18 is 512 dimensions. The input size of ResNet-50 and ResNet-101 is 2048 dimensions, and the last block of them has the same structure. Their network structures are illustrated in Fig. 8.

3.3 Implementation

Our training dataset was transformed with multiple transformations; random vertical and horizontal flips, random rotations from 15 to -15 degrees, random rotations from -10 to 10 degrees. The preprocessing training images are fixed to 300×300 pixels. They are scaled and padded under the condi-

tion of maintaining proportional size. The AlexNet and ResNet models are modified, according to explanation in Section 3.2. We implemented these models using Pytorch [23] machine learning framework. Stochastic gradient descent was chosen to be the optimizer. The momentum and weight decay is 0.9 and 0.0001. The learning rate is separated into two states. The warm up state is started from 1^{st} to 10^{th} epoch with logarithm increasing from 0.001 to 0.005. The latter state is logarithm decreasing from 0.005 at 11^{th} to 0.001 at the end of training epochs. We trained 400 epochs for every model and fed 16 images for 1 batch size. Our experimental hardware is NVIDIA RTX3070Ti and Intel 11900KF.

4. EXPERIMENTS

4.1 Evaluation Metrics

In this paper, we report the classification of our models giving accuracy, sensitivity, specificity and F1 score. The accuracy is a general classification metric when every class is balanced. Sensitivity and specificity can measure the classification performance for both true positive rate and true negative rate with respect to actual class. The F1 score is the harmonic mean between precision and recall (∼sensitivity) as shown in equation (1). This metric can measure balancing score between precision and recall. For example, the highest F1 score is 1 when precision and recall is 1. Whenever precision or recall is close to 0, the F1 score is likewise close to 0, regardless of what another one is. Moreover, we visualize the receiver operating characteristic(∼ROC) graph for measuring the optimal threshold.

$$F_1 = 2 * \frac{1}{\frac{1}{precision} + \frac{1}{recall}} \quad (1)$$

4.2 Binary Classification

In this experiment, we compared 4 machine learning models for caries classification: AlexNet, ResNet-18, ResNet-50 and ResNet-101. We separated the bitewing radiographs into 3 sub-experiments: all types of tooth conditions, teeth with restorations, and teeth without dental restorations. This experiment tests the hypothesis that classification of dental restoration is more complicated than for normal teeth. For each sub-experiment, the bitewing radiographs for testing is 20 percent of training.

4.2.1 Binary classification of overall tooth conditions

In this part, there are 2400 bitewing radiographs split into training and testing dataset as 2040 and 360 bitewing radiographs, respectively. The number of bitewing radiographs is equally divided for each class with 1020 images for training and 180 images for testing. The experiment result is illustrated in Table 1. All benchmark results of AlexNet are inferior to ResNet results. The Accuracy of ResNet-50 achieves 86.67%. Noticeably, sensitivity and specificity of ResNet-101 achieve equally 85%. The highest F1 score is 87% of ResNet-50 which is higher than lowest score of AlexNet by 7%.

Table 1: *Benchmark results between AlexNet, ResNet-18, ResNet-50, and ResNet-101 models on 2-class dental caries classification of overall tooth conditions.*

	AlexNet	ResNet18	ResNet50	ResNet101
Accuracy	80.28	85.00	86.67	85.28
Sensitivity	0.77	0.85	0.85	0.85
Specificity	0.75	0.84	0.84	0.85
F1 score	0.80	0.85	0.87	0.85

4.2.2 Binary classification of teeth without dental restorations

In this sub-experiment, our training dataset was equally separated into 510 sound teeth and dental carious lesions. The training and testing dataset totals are 1020 and 180 bitewing radiographs. As illustrated in Table 2, the best benchmark result of this experiment performs better than the experimental result of dental dataset with restoration for every metric.

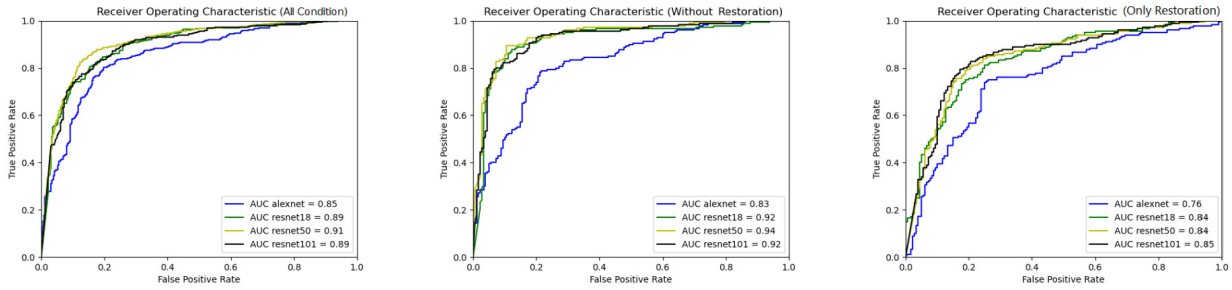
Table 2: *Benchmark results between AlexNet, ResNet-18, ResNet-50, and ResNet-101 models on 2-class dental caries classification of teeth without dental restoration.*

	AlexNet	ResNet18	ResNet50	ResNet101
Accuracy	78.33	80.33	87.78	86.67
Sensitivity	0.80	0.88	0.86	0.88
Specificity	0.77	0.89	0.90	0.86
F1 score	0.78	0.88	0.88	0.87

4.2.3 Binary classification of teeth with dental restorations

In this part, the training dataset was equally separated into 510 sound teeth with dental restorations and dental carious lesions with dental restorations. The training and testing dataset sizes are equivalently set to the previous part as 1020 and 180 dental images. As illustrated in Table 3, this experimental result has the worst performance compared to the previous binary-classification sub-experiments.

Fig. 9 displays the comparison ROC-curve graph of binary classification results between radiographs of all type of tooth conditions, radiographs of teeth without and with dental restorations. Fig. 9(a) has highest AUC of ROC while Fig. 9(c) has the lowest AUC of ROC. These AUC of ROC metrics



(a) (b) (c)

Fig.9: *ROC curve of 2-class dental caries classification results.*

Table 3: *Benchmark results between AlexNet, ResNet-18, ResNet-50, and ResNet-101 models on 2-class dental caries classification of teeth with dental restoration.*

	AlexNet	ResNet18	ResNet50	ResNet101
Accuracy	74.44	82.22	82.22	82.78
Sensitivity	0.72	0.80	0.81	0.80
Specificity	0.68	0.78	0.80	0.78
F1 score	0.75	0.82	0.82	0.83

are consistent with the evaluation metrics in Table 2 which shows that the radiograph of teeth with dental restoration is rather difficult to classify.

Table 4: *Comparison of multiclass classification in terms of accuracy with different ResNet models.*

	ResNet18	ResNet50	ResNet101
healthy	78.89	90.00	82.22
enamel	86.67	72.22	76.67
dentin	73.33	77.78	78.89
restoration	76.67	83.33	88.89
2nd caries	77.78	73.33	73.33
average	78.67	79.33	80.00

4.3 Multilevel Lesion Classification

We separate training and testing datasets into 5 classes: sound teeth, enamel caries, dentin caries, teeth under dental restoration, and dental carious lesion under dental restoration. There are 510 training images and 90 testing images for each class. This experiment aimed to inspect the classification performance on different caries levels. We removed the AlexNet model on model comparison due to unremarkable performance as already reported in Section 4.2 These results prove that the complexity of Alexnet is not sufficient and does not fit with our dataset whereas it is cross-sharing between Experiment 4.2 and Experiment 4.3.

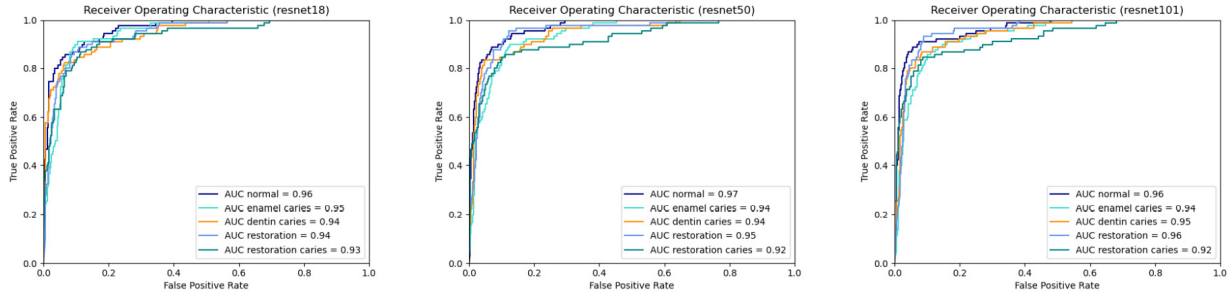
In Table 4, the average accuracy of 5 classes is comparable for each ResNet model. Anyhow, the accuracy for each class is totally different. ResNet-18 of enamel caries and secondary caries radiographs are outstanding which are higher than ResNet-50 by 14.45%, 4.45% and ResNet-101 by 10%, 4.45%, respectively. ResNet-50 is outstanding with sound tooth radiographs which are higher than ResNet-18 by 11.12% and ResNet-101 by 7.78%. While Resnet-101 is outstanding with dentin caries and dental restoration radiographs which are higher than ResNet-18 by 5.56%, 12.22% and ResNet-50 by 1.11%, 5.56%, respectively. In Table 5, the sensitivity of ResNet-50 outperforms ResNet-18 by 11% and ResNet-101 by 8% on healthy tooth radiographs. ResNet-101 outperforms ResNet-18 and ResNet-50 by 6%, 1% on dentin caries radiographs and 12%, 6% on dental restoration radiographs, respectively. The interesting result is ResNet-18 which outperform the ResNet-50 and ResNet-101 on enamel caries by 15%, 10% and secondary caries radiographs by 5% equally. The sensitivity result is entirely consistent with accuracy result and shows that the classification of enamel caries and secondary caries apparently excel on ResNet-18.

Table 5: *Comparison of multiclass classification in terms of sensitivity with different ResNet models.*

	ResNet18	ResNet50	ResNet101
healthy	0.79	0.90	0.82
enamel	0.87	0.72	0.77
dentin	0.73	0.78	0.79
restoration	0.77	0.83	0.89
2nd caries	0.78	0.73	0.73
average	0.79	0.79	0.80

The specificity of multiclass classification is excellent, but insignificantly different from every ResNet model as illustrated in Table 6 and likewise by the F1 score in Table 7.

Fig. 10 shows the ROC-curve comparison between 5 classes. The AUC of ROC-curve metrics are consis-



(a) (b) (c)

Fig.10: *ROC curve of 5-class dental caries classification results.*

Table 6: *Comparison of multiclass classification in terms of specificity with different ResNet models.*

	ResNet18	ResNet50	ResNet101
healthy	0.96	0.92	0.96
enamel	0.91	0.95	0.94
dentin	0.96	0.98	0.96
restoration	0.96	0.94	0.93
2nd caries	0.94	0.96	0.96
average	0.95	0.95	0.95

Table 7: *Comparison of multiclass classification in terms of F1-score with different ResNet models.*

	ResNet18	ResNet50	ResNet101
healthy	0.82	0.81	0.84
enamel	0.78	0.75	0.76
dentin	0.78	0.83	0.80
restoration	0.80	0.81	0.82
2nd caries	0.77	0.77	0.78
average	0.79	0.79	0.80

tent across the entire class. They are excellent with a range of 0.92 to 0.97 for all ResNet models.

4.4 Comparison with other dental caries classification models

As mentioned earlier, in other dental caries classification models, we compare our model and other proposed models as illustrated in Fig. 11. The accuracy of our models is 87.78% on exclusion of dental restoration and 86.67% on overall tooth conditions. It is obvious that our models outperform other dental caries classification models [18, 20], even though Jae-Hong Lee et al. [18] exclude the dental restoration.

4.5 Ablation Study

In this section, we analyzed the misclassification result of ResNet-101 which provided the comparative average-class performance with ResNet-18. In Table 8, the highest misclassification occurs in 2 groups. First, enamel carious lesions are misclassified as healthy teeth and dentin carious lesions. Second, secondary carious lesions are misclassified as teeth with restorations. We implemented our custom gradient class activation map (∼GradCAM) [27] as illustrated in Fig. 12 and 13.

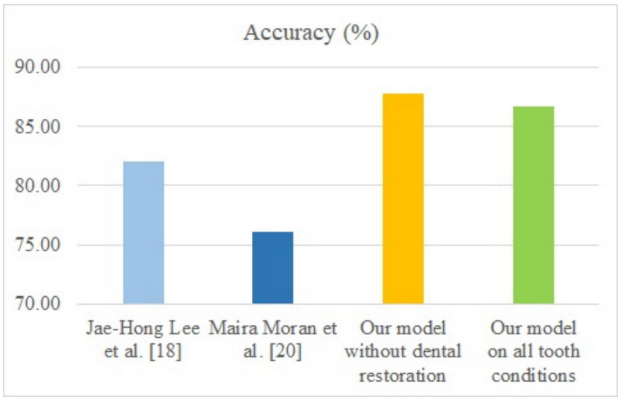


Fig.11: *Comparison our model with other dental caries classification models.*

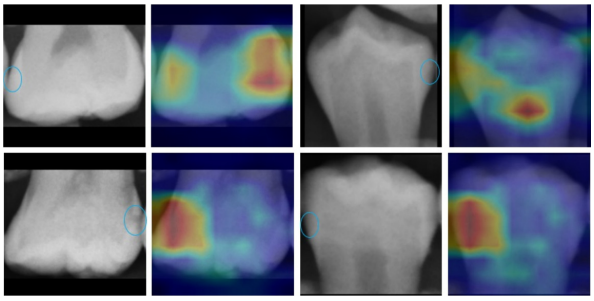


Fig.12: *Heatmaps of enamel carious lesions that are misclassified as healthy teeth on first row and dentin carious lesions on second row. The trivial heatmaps are generated by GradCAM and the blue circles are carious lesion mark points found by experts.*

Table 8: *Confusion matrix of ResNet-101 models.*

		Predict Class				
		healthy	**enamel**	**dentin**	**restoration**	**2nd caries**
Actual Class	**healthy**	74	11	3	2	0
	enamel	8	69	10	1	2
	dentin	4	12	71	1	2
	restoration	1	0	0	80	9
	2nd caries	0	0	3	21	66

4.5.1 Failure Case of Enamel Caries Classification

Fig. 12 shows the enamel caries misclassifications. The top row is misclassification of enamel caries as sound teeth. The heatmaps are generated by Grad-CAM such that they are considered as trivial points by the decision of fully connected layer. The second row is misclassification of enamel caries as dentin caries. The heatmaps are generated in the wrong position due to the similarity of dental carious hole to dentin-enamel junction. This effect appears in the x-ray film that leads to misclassifying healthy teeth to dental caries, or vice versa.

4.5.2 Failure Case of Secondary Caries Classification

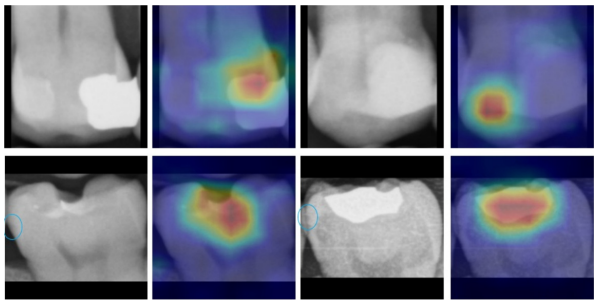


Fig.13: *Heatmaps of secondary carious lesions that are misclassified as teeth with restorations. The blue circles are carious lesion mark points found by experts.*

In Fig. 13, the heatmap of secondary caries misclassifications is shown. The top row is healthy teeth which are misclassified as secondary caries. The second row is secondary caries but it is misclassified as healthy teeth. This result is consistent with Mua at al.[22] that radiolucent halos under dental restoration correlate to the remaining demineralized tissue (RDT) or excessive unfilled resin bonding agent which leads to a false positive. These effects mislead our model to classify them as dental caries.

5. CONCLUSIONS

In this paper, we propose two bitewing radiograph classification models: dental caries classification and multilevel dental caries classification. Our model successfully achieved accuracy of 87.78% on dental caries classification and 80% on multilevel dental caries classification. Noticeably, the ResNet-18 is outstanding compared to the ResNet-50 and ResNet-101 with enamel caries and secondary caries classification, but it has trade-off with an accuracy drop for sound teeth, dentin caries and teeth with restoration. The performance of ResNet-50 and ResNet-101 should provide the higher accuracy than ResNet-18. ResNet-50 and ResNet-101 require large datasets and learn from immense variation but our dental caries radiographs are insufficient for training. Moreover, our model could be practical in real clinical conditions. We successfully implemented the custom GradCam to identify the carious lesion in bitewing radiographs that possibly can support the discussion between dentists to enhance diagnostic and treatment quality.

References

[1] V. Baelum, J. Heidmann, and B. Nyvad, "Dental caries paradigms in diagnosis and diagnostic research," *Eur. J. Oral Sci.*, vol. 114, no. 4, pp. 263–277, 2006.

[2] F. Brouwer, H. Askar, S. Paris, and F. Schwendicke, "Detecting secondary caries lesions: a systematic review and meta-analysis," *J. Dent. Res.*, vol. 95, no. 2, pp. 143–151, 2016

[3] A. G. Cantu et al., "Detecting caries lesions of different radiographic extension on bitewings using deep learning," *J. Dent.*, vol. 100, no. July 2020, p. 103425, 2020, doi: 10.1016/j.jdent.2020.103425.

[4] F. Casalegno et al., "Caries Detection with Near-Infrared Transillumination Using Deep Learning," *J. Dent. Res.*, vol. 98, no. 11, pp. 1227–1233, 2019, doi: 10.1177/0022034519871884.

[5] M. J. Chong, W. K. Seow, D. M. Purdie, E. Cheng, and V. Wan, "Visual-tactile examination compared with conventional radiography, digital radiography, and Diagnodent in the diagnosis of occlusal occult caries in extracted premolars," *Pediatr. Dent.*, vol. 25, no. 4, pp. 341–349, 2003.

[6] P. R. da Silva, M. Martins Marques, W. Steagall Jr, F. Medeiros Mendes, and C. A. Lascala, "Accuracy of direct digital radiography for detecting occlusal caries in primary teeth compared with conventional radiography and visual inspection:

an in vitro study," *Dentomaxillofacial Radiol.*, vol. 39, no. 6, pp. 362–367, 2010.

[7] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009, pp. 248–255.

[8] K. L. Devito, F. de Souza Barbosa, and W. N. F. Filho, "An artificial multilayer perceptron neural network for diagnosis of proximal dental caries," *Oral Surgery, Oral Med. Oral Pathol. Oral Radiol. Endodontology*, vol. 106, no. 6, pp. 879–884, 2008, doi: 10.1016/j.tripleo.2008.03.002.

[9] B. L. Edelstein, "The dental caries pandemic and disparities problem," in *BMC oral health*, 2006, vol. 6, no. 1, pp. 1–5.

[10] M.-A. Geibel, S. Carstens, U. Braisch, A. Rahman, M. Herz, and A. Jablonski-Momeni, "Radiographic diagnosis of proximal caries—influence of experience and gender of the dental staff," *Clin. Oral Investig.*, vol. 21, no. 9, pp. 2761–2770, 2017.

[11] A. Haghanifar, M. M. Majdabadi, and S.-B. Ko, "PaXNet: Dental Caries Detection in Panoramic X-ray using Ensemble Transfer Learning and Capsule Classifier," pp. 1–14, 2020, [Online]. Available: http://arxiv.org/abs/2012.13666

[12] H. Hajizadeh, M. Akbari, S. Zarch, A. Nemati-Karimooy, and A. Izadjou, "Evaluation of the interpretation of bitewing radiographs in treating interproximal caries," *Eur. J. Gen. Dent.*, vol. 8, no. 1, p. 13, 2019.

[13] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[14] E. R. Hewlett, K. A. Atchison, S. C. White, and V. Flack, "Radiographic secondary caries prevalence in teeth with clinically defective restorations," *J. Dent. Res.*, vol. 72, no. 12, pp. 1604–1608, 1993.

[15] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, "Densenet: Implementing efficient convnet descriptor pyramids," *arXiv Prepr. arXiv1404.1869*, 2014.

[16] J. R. Keenan and A. V. Keenan, "Accuracy of dental radiographs for caries detection," *Evid. Based. Dent.*, vol. 17, no. 2, p. 43, 2016.

[17] A. Krizhevsky, Ii. Sulskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst.*, vol. 60, no. 6, pp. 84–90, 2012, doi: 10.1145/3065386.

[18] J. H. Lee, D. H. Kim, S. N. Jeong, and S. H. Choi, "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm," *J. Dent.*, vol. 77, no. June, pp. 106–111, 2018, doi: 10.1016/j.jdent.2018.07.015.

[19] L. Lian, T. Zhu, F. Zhu, and H. Zhu, "Deep learning for caries detection and classification," *Diagnostics*, vol. 11, no. 9, 2021, doi: 10.3390/DIAGNOSTICS11091672.

[20] M. Moran, M. Faria, G. Giraldi, L. Bastos, L. Oliveira, and A. Conci, "Classification of approximal caries in bitewing radiographs using convolutional neural networks," *Sensors*, vol. 21, no. 15, pp. 1–12, 2021, doi: 10.3390/s21155192.

[21] K. Moutselos, E. Berdouses, C. Oulis, and I. Maglogiannis, "Recognizing Occlusal Caries in Dental Intraoral Images Using Deep Learning," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 1617–1620, 2019, doi: 10.1109/EMBC.2019.8856553.

[22] B. Mua, V. R. Fontanella, F. C. Giongo, M. Maltz, and others, "Radiolucent halos beneath composite restorations do not justify restoration replacement.," *Am. J. Dent.*, vol. 28, no. 4, pp. 209–213, 2015.

[23] A. Paszke et al., "PyTorch: An Imperative Style, High-Performance Deep Learning Library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, Eds. Curran Associates, Inc., 2019, pp. 8024–8035. [Online]. Available: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

[24] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*, 2015, pp. 234–241.

[25] F. Schwendicke, K. Elhennawy, S. Paris, P. Friebertshäuser, and J. Krois, "Deep learning for caries lesion detection in near-infrared light transillumination images: A pilot study," *J. Dent.*, vol. 92, no. December 2019, p. 103260, 2020, doi: 10.1016/j.jdent.2019.103260.

[26] F. Schwendicke, M. Tzschoppe, and S. Paris, "Radiographic caries detection: a systematic review and meta-analysis," *J. Dent.*, vol. 43, no. 8, pp. 924–933, 2015.

[27] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 618–626.

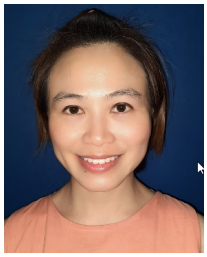
[28] M. M. Srivastava, P. Kumar, L. Pradhan, and S. Varadarajan, "Detection of Tooth caries in Bitewing Radiographs using Deep Learning," no. Nips 2017, 2017, [Online]. Available: http://arxiv.org/abs/1711.07312

[29] C. Szegedy et al., "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.

[30] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 2818–2826, 2016, doi: 10.1109/CVPR.2016.308

[31] A. Wirtz, S. G. Mirashi, and S. Wesarg, "Automatic teeth segmentation in panoramic X-ray images using a coupled shape model in combination with a neural network," in *International conference on medical image computing and computer-assisted intervention*, 2018, pp. 712–719.

Wannakamon Panyarak received a bachelor's degree in Dental Surgery from the Faculty of Dentistry, Chiang Mai University in 2013. She was a PhD student at the Graduate School of Dental Science, Kyushu University and achieved a PhD degree in Dental Science (Oral and Maxillofacial Radiology) in 2020. Her thesis mainly involved MRI assessment of differential diagnosis of orofacial lesions. Due to limited facility accessibility, her research interest shifts to deep learning-based study using available data at the institution. She was introduced to a biomedical engineer and a software engineer who facilitated and provided valuable suggestions regarding deep learning and algorithmic research. With these supportive colleagues, she is hoping to continuously publish articles and conduct research in deep learning in the dentistry field.



Dauangporn Jira-apiwattana received the Doctor of Dental Surgery degree in 2008 and higher graduate diploma in clinical science (Endodontics) in 2012 from Chiang Mai University, Thailand. Furthermore, she received master's degree in Public Health from University of Phayao, Thailand in 2021. She has been working as a dentist at Phayao Hospital and a special clinical instructor in oral radiology and oral diagnosis department, School of Dentistry, University of Phayao, Thailand.

Wattanapong Suttapak received M.E. degree in Computer Engineering from Chiang Mai University, Thailand in 2010 and B.Sc. degree in Physics from Chiang Mai University Thailand in 2005. He has been working as lecturer in Computer Engineering, School of Information Communication and Technology at University of Phayao, Thailand. His current research interests cover deep learning, visual tracking, adversarial learning, and biomedical engineering.

Kittichai Wantanajittikul received his bachelor's and master's degrees in Electrical Engineering. Next, he applied for the Ph.D. student in the Biomedical Engineering Program at Chiang Mai University and achieved a Ph.D. degree in 2016. His thesis involved the automatic cardiac T2* relaxation time calculation from MRI using image processing and machine learning techniques for estimating the iron-overloaded in the myocardium. He has been working at the Department of Radiologic Technology, Faculty of Associated Medical Sciences, Chiang Mai University. His research focused on the use of image processing and machine learning technologies in healthcare.