Evaluation of Convolutional Neural Network for Automatic Caries Detection in Digital Radiograph Panoramic on Small Dataset

Arna Fariza, Rengga Asmara and Muhammad Oktavian Fajar Rojaby
Department of Informatics and Computer Engineering
Politeknik Elektronika Negeri Surabaya
Surabaya, Indonesia
{arna & rengga}@pens.ac.id

Eha Renwi Astuti and Ramadhan Hardani Putra
Department of Dentistry
Universitas Airlangga
Surabaya, Indonesia
{eha-r-a & ramadhan.hardani}@fkg.unair.ac.id

2022 International Conference on Data and Software Engineering (ICoDSE) | 979-8-3503-9705-5/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICODSE56892.2022.9972183

Abstract - **Dental caries or tooth decay is damage to the hard tissues of the teeth that can occur in the enamel, dentin, and cementum areas. Panoramic radiography is a screening tool for tactile or visual examination of the oral cavity which is useful for further diagnosis and treatment. The process of segmentation of panoramic radiographs is a difficult process because there is no homogeneity between panoramic images with one another. Noise levels, vertebral column images, and low contrast are the main challenges in image processing. This study evaluates CNN to detect caries automatically on panoramic radiographs on a small dataset. The dataset consisted of manually cropped maxillary and mandibular premolars and molars. An augmentation strategy consisting of horizontal flip, vertical flip, and affine transformation is used to produce a wider variety of images. This study compares the architecture of non-pretrained and pretrained models consisting of 3-layer CNN, 3-layer CNN with batch normalization, ResNet18, and ResNeXt50 32×4d. Evaluation was carried out on 400 training data and 76 testing data. Combination of augmentation strategies and pre-trained ResNet18 and ResNeXt50 32×4d achieves high accuracy compared to other models.**

Index Terms* – *caries detection, radiograph panoramic, residual network.

I. Introduction

Dental caries is a process of demineralization of dental tissue due to dental infection which is often exposed to fermentable carbohydrates and is influenced by saliva and fluoride and other elements [1]. Dental caries or tooth decay is damage to the hard tissues of the teeth that can occur in the enamel, dentin, and cementum areas [2]. If dental caries is not treated, it can have further detrimental effects on a person's quality of life, productivity, growth, and general health [3].

A panoramic radiography is an intraoral radiograph that reveals every tooth in the upper and lower jaws. This radiograph plays an important role in detecting caries [1]. A screening tool for tactile or visual evaluation of the mouth cavity is panoramic radiography [4]; this examination is helpful for subsequent diagnosis and therapy. To sharpen or enhance the quality of the information included in the image so that it can be properly comprehended by the human eye, image processing is one example of how information technology is applied in the medical field [5]. The use of computer-aided diagnosis (CAD) to detect and classify pathology in cavities and periodontitis provides a valuable second opinion to dental professionals in an automated manner [6].

Conventional CAD methods require feature extraction which is the most difficult and time-consuming task before training to recognize and classify images [6]. The process of segmentation of panoramic radiographs is a difficult process because there is no homogeneity between panoramic images with one another [7]. Noise levels, vertebral column images, and low contrast are the main challenges in image processing [8]. Deep learning techniques have been successfully applied to handle complex radiology problems in order to solve this issue [9]. Convolutional neural network (CNN) research has advanced quickly, and it is anticipated to result in faster, more accurate diagnoses [10]. CNN promises caries detection results with high accuracy on radiographic data [11-15].

This study evaluates CNN to detect caries automatically on panoramic radiographs on a small dataset. The dataset consisted of manually cropped maxillary and mandibular premolars and molars. An augmentation strategy consisting of horizontal flip, vertical flip, and affine transformation is used to produce a wider variety of images. This study compares the architecture of non-pretrained and pretrained models consisting of 3-layer CNN, 3-layer CNN with batch normalization, ResNet18, and ResNeXt50 32×4d. Evaluation was carried out on 400 training data and 76 testing data.

II. Related Works

Several CNN classifications studies on radiographic data promise better accuracy than conventional feature extraction and classification methods. Transfer learning architecture and modified version are used to get accurate caries detection.

To find carious lesions on 800 digital bitewing radiographs, CNN utilized a modified YOLO model. The results of the trial on 200 other radiographs achieved 94.59% accuracy, 72.26% sensitivity, and 98.19% specificity [11]. Lee et al. [12] evaluated CNN GoogleNet Inception v3 in detecting and diagnosing dental caries on periapical radiographs. The trial was conducted on 3000 periapical radiographic images consisting of premolars and molars divided into training and validation datasets. The premolar, molar, and molar-premolar accuracy were 89.0%, 88.0%, and 82.0%, respectively. Estai et al. [13] evaluated CNN for automatic caries detection in 1257 caries and 1211 normal bitewing radiographs. The caries identification module was trained and validated using CNN with Inception-ResNet-v2 architecture based on a ROI of 3297 caries and 5321 normal.

The CNN pretrained model shows superior detection accuracy in small datasets. Classification of dental caries using CNN MobileNet V2 on 400 panoramic images cut in the mandibular and maxillary third molars is performed [14]. The 100 panoramic radiographs were tested, and the results showed 0.87 accuracy, 0.86 sensitivity, and 0.88 specificity. Lin et al. [15] evaluated the performance of CNN on a small dataset for detecting proximal caries of different severity on periapical radiographs. The 800 periapical radiographs randomly divided into training and validation data sets were taught using a pre-trained CNN Cifar-10Net.

In comparison to AlexNet, VGG, Xception, and ResNeXt, the ResNet architecture has a high level of accuracy for dental caries detection [16]. By utilizing earlier layer activations, ResNet (residual network) addresses the primary issue of fading gradients that networks encounter when using shortcut connections. For the rest of the network to learn and explore new feature areas, this connection compression is followed by layer expansion [17]. The analysis of medical pictures has seen widespread use of residual networks and their modifications. ResNeXt adds cardinality to several block paths to reduce validation errors without having to have high depth and width [18]. This study evaluates the 3-layer CNN architecture and pre-trained residual network on a small dataset of panoramic radiographs for caries detection.

III. Material and Method

The automatic detection of dental caries using CNN can be seen in Figure 1. The training dataset consists of the maxillary and mandibular premolars and molars generated manually from a panoramic radiograph. The training dataset is preprocessed which consists of augmentation, resizing and normalization. The training dataset was then split into two portions at random, i.e., training and validation data, using a ratio of 0.8:0.2. The training process uses 4 CNN architectures and hyperparameters to produce the best model. The prediction process uses a randomly selected testing dataset. Before making predictions using the best model, preprocessing is carried out which consists of resizing and normalization. Prediction results produce images of teeth with caries or normal.

A. Data Collection

Panoramic radiograph data of the Indonesian race was acquired from Parahita Diagnostic Center Sidoarjo, Indonesia. These images were manually cropped to a size of 224 × 224 pixel on the maxillary and mandibular premolars and molars. Cropped image showing a full tooth element from crown to tooth root. This is carried out due to the fact that dental caries can affect any tooth portion. Some examples of the results of cutting teeth with caries and normal can be seen in Figure 2. Dental image labeling with caries and normal has been validated by the dentists.

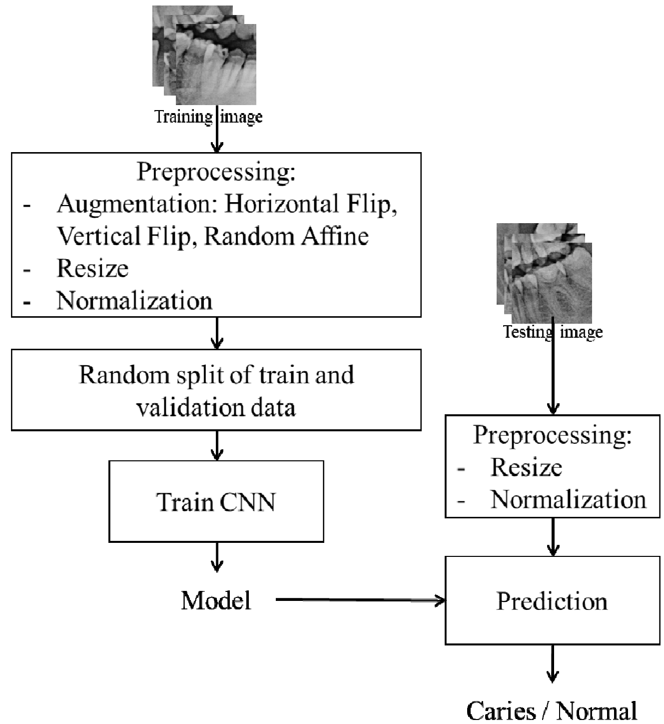


Fig. 1 Design System

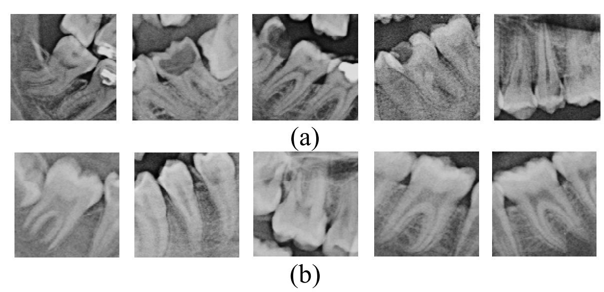


Fig. 2 Some examples of cropped radiography panoramic (a) with caries; (b) normal

The dataset collected was balanced between carious and normal teeth, each with 238 images. The testing dataset was randomly selected from the data consist of 76 images (32%). While the training dataset consists of 400 images.

B. Preprocessing

Augmentation is the method used to increase the volume of training data. The augmentation strategy was chosen to help prevent the model from being overfitted as more variation forms in the data. Training dataset of 400 images was expanded to 4000 images. Augmentation uses random horizontal flip (with probability 0.5), random vertical flip (with probability 0.5), and random affine modification with a degree range of -90 to 90 to maintain the collected information consistent with the original image.

In addition to data augmentation, resizing was carried out to reduce the image size from 224 × 224 pixel to 100 × 100 pixel. The normalization process was carried out to distinguish teeth and background with mean = 0.5 and standard deviation = 0.5. The results of the augmentation, resizing, and normalization processes on some image data can be seen in Figure 3.

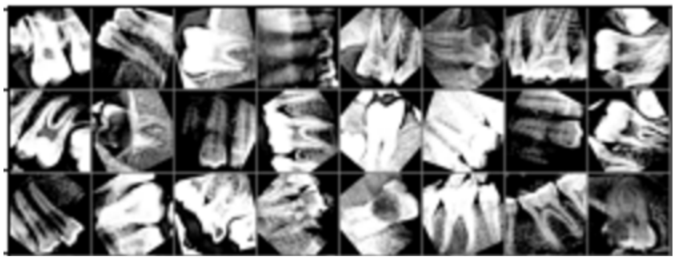


Fig. 3 Some examples of image after augmentation and normalization process.

C. CNN Architecture

This research evaluates the architecture without and with transfer learning models. Characteristics of images that have noise levels, vertebral column images, and low contrast, in addition to small datasets are a challenge. The evaluated architecture consisted of 3-level CNN, 3-level CNN with batch normalization, pre-trained ResNet18, and pre-trained ResNeXt50 32×4d. The input image that goes into the architecture is 100 × 100 pixels and produces 2 labels, namely caries and normal.

Figure 4 is a simple 3-layer CNN architecture. There are 3 times the convolution process with a 3 × 3 filter and ReLU. Before the classification process is carried out using a linear function, max pooling 2 × 2 is added and a dropout of 0.5 is added.

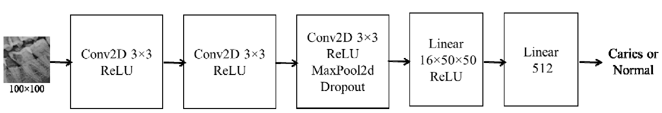


Fig. 4 3-layer CNN architecture

Figure 5 adds batch normalization to the convolution and classification sections. Batch normalization can increase CNN learning speed and avoid overfitting. Before the classification process, an average pooling process of 2 × 2 and a dropout of 0.5 was added.

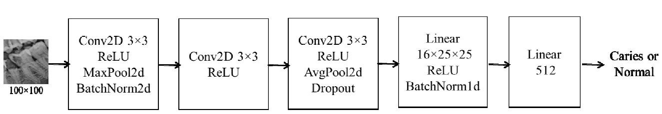
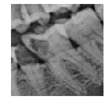


Fig. 5 3-layer CNN with batch normalization architecture

This study also evaluated the pre-trained ResNet18 model and ResNeXt50 32×4d to detect dental caries. ResNet18 consists of 6 convolution layers and average pooling. ResNet18 architecture can be seen in Figure 6. While Figure 7 is the ResNeXt50 32×4d architecture. The difference between ResNeXt architecture and ResNet is the cardinality factor that divides the path and convolution operations; therefore, the architecture does not get deeper, and the number of parameters is the same as the ResNet model.



100×100

Layer Name	Output Size	ResNet-18
conv1	112 × 112 × 64	7 × 7, 64, stride 2
conv2_x	56 × 56 × 64	3 × 3 max pool, stride 2
		$\begin{bmatrix} 3\times3, 64 \\ 3\times3, 64 \end{bmatrix} \times 2$
conv3_x	28 × 28 × 128	$\begin{bmatrix} 3\times3, 128 \\ 3\times3, 128 \end{bmatrix} \times 2$
conv4_x	14 × 14 × 256	$\begin{bmatrix} 3\times3, 256 \\ 3\times3, 256 \end{bmatrix} \times 2$
conv5_x	7 × 7 × 512	$\begin{bmatrix} 3\times3, 512 \\ 3\times3, 512 \end{bmatrix} \times 2$
average pool	1 × 1 × 512	7 × 7 average pool
fully connected	1000	512 × 1000 fully connections
softmax	1000	

Caries or Normal

Fig. 6 Resnet18 architecture

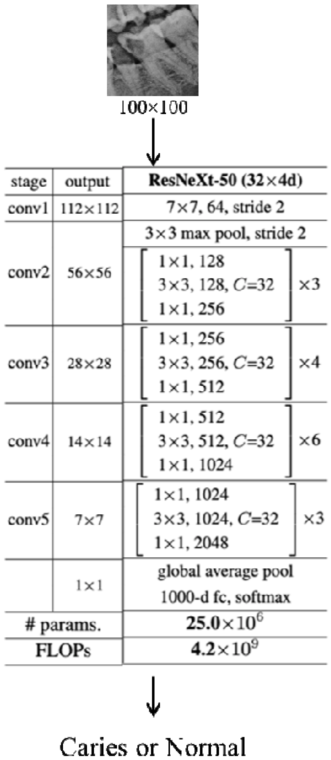


Fig. 7 ResNeXt50 32×4d

D. Training

The training process involves training and validation data which are randomly selected from a total of 4000 images with a ratio of 0.8:0.2. Therefore, the amount of training data is 3200 and validation data is 800. The hyperparameters used in the training process can be seen in Table I. The Adam optimizer and the loss cross entropy function are used in the training, and the learning rate is set to 0.0001.

TABLE I
HYPERPARAMETER

Parameter	Value
Loss function	Cross entropy
Optimizer	Adam
Learning Rate	0.0001

IV. RESULT AND DISCUSSION

Evaluation of automatic caries detection performance on small datasets includes evaluation of the training process and prediction results. The training data amounted to 4000 images after the augmentation process which was divided into training datasets and validation datasets at random as many as 3200 and 800. While the data testing performed predictions as many as 76 data selected randomly.

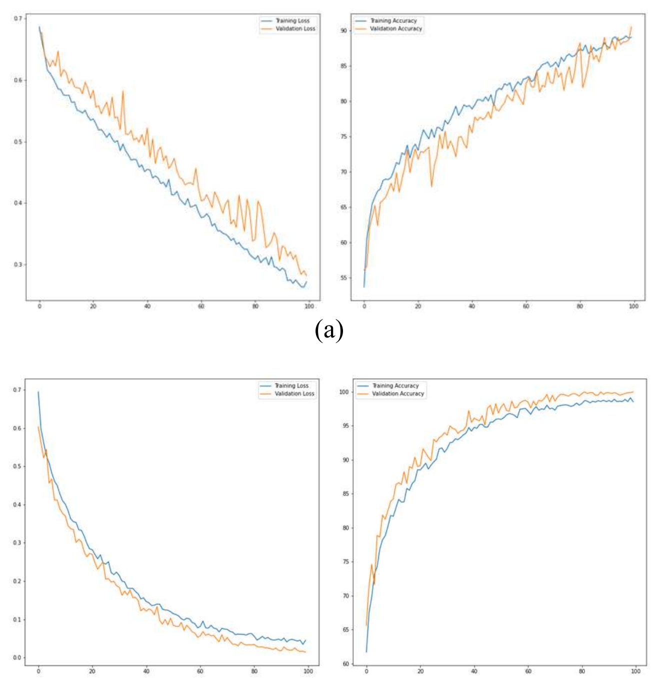
A. Training Process Evaluation

Performance evaluation of the training process can be seen in Table II. The 3-layer CNN and 3-layer CNN architecture training with batch normalization is iterated for 100 epochs, while the pre-trained ResNet18 and ResNeXt50 32×4d are 50 epochs. The ResNeXt50 32×4d model produces the smallest training loss and validation loss, almost equal to the ResNet18 model. The highest training accuracy and validation accuracy were achieved with the ResNet18 model.

Due to the close values of the distances between training loss and validation loss as well as training accuracy and validation accuracy, the four models exhibit good training performance. This is also shown in the graph plot of loss and accuracy in Figure 8. Pre-trained models ResNet18 and ResNeXt50 32×4d perform better training processes than 3-layer CNN and 3-layer CNN with batch normalization.

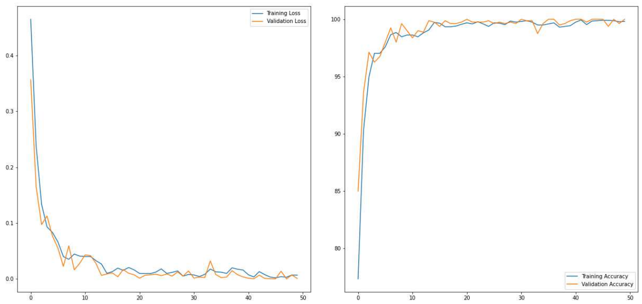
TABLE II
TRAINING EVALUATION

CNN Architecture	Performance			
	3-layer CNN	3-layer CNN with batch normalization	ResNet18	ResNeXt50 32×4d
Number of epochs	100	100	50	50
Training loss	0.272	0.045	0.007	0.001
Validation loss	0.282	0.014	0.001	0.000
Training accuracy (%)	89.03	98.53	99.81	100.00
Validation accuracy (%)	90.50	100.00	100.00	100.00

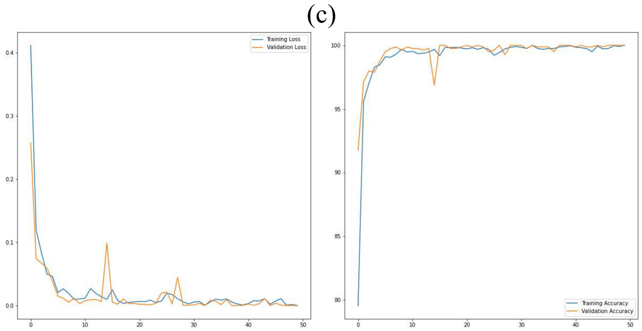


(a)

(b)



(c)



(d)

Fig. 8 Trainig loss anda validation result (a) 3-layer CNN; (b) 3-layer CNN with batch normalization; (c) Resnet18; (d) ResNeXt50 32×4d

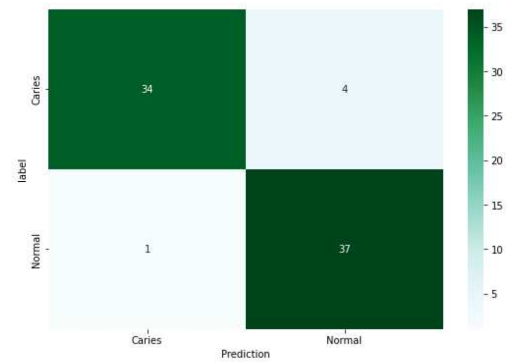
B. Prediction Result

Accuracy, precision, recall, and F1-score are the components of the confusion matrix, which provide information on how the model acts. The percentage of accurate forecasts among all guesses is called accuracy. The accuracy of the model's classification of the sample as positive is measured by precision. The model's capacity to identify positive samples is measured by recall. A statistic that considers recall and precision is the F1-score.

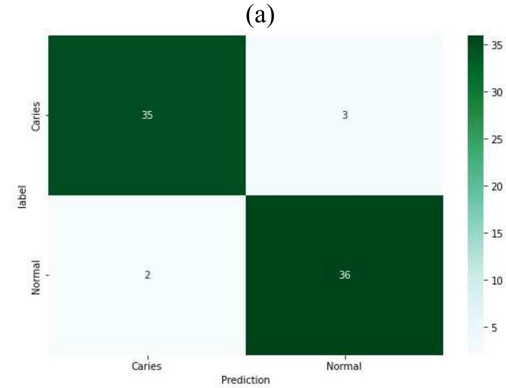
The evaluation of the prediction results for the 4 CNN models can be seen in Table III and the visualization of the confusion matrix in Figure 9. Pre-trained models ResNet18 and ResNeXt50 32×4d achieve better accuracy, precision, recall, and F1-Score values than 3-layer CNN and 3-layer CNN with batch normalization. Pre-trained models ResNet18 and ResNeXt50 32×4d have the same accuracy, but ResNeXt50 32×4d is more precise, ResNet18 is more sensitive, and both have almost the same F1-Score value. Figure 9 shows that ResNeXt50 32×4d can correctly detect caries teeth better than ResNet18 and vice versa. In general, the pre-trained residual network model produces better predictions of dental caries detection than 3-layer CNN and 3-layer CNN with batch normalization for small databases. This shows that the architectural model of 3-layer CNN and 3-layer CNN with batch normalization is still overfitting even though it has been augmented.

TABLE III
PREDICTION EVALUATION

CNN Architecture	Performance (%)			
	Accuracy	Precision	Recall	F1-Score
3-layer CNN	78.95	76.19	84.21	80.00
3-layer CNN with batch normalization	72.37	79.31	60.53	68.66
ResNet18	93.42	90.24	97.37	93.67
ResNeXt50 32×4d	93.42	92.31	94.74	93.51



(a)



(b)

Fig. 9 Confusion matrix (a) Resnet18; (b) ResNeXt50 32×4d

IV. CONCLUSION

This study evaluates the comparison of 3-layer CNN, 3-layer CNN with batch normalization, ResNet18, and ResNeXt50 32×4d models for automatic caries detection in small datasets. Combination of augmentation strategies and pre-trained ResNet18 and ResNeXt50 32×4d achieves high accuracy compared to other models. This shows that the pre-trained model with augmented dataset strategy is effective in recognizing tooth patterns from digital panoramic radiographs which have heterogeneous, noisy, and low contrast characteristics.

This research is the beginning for the creation of a CAD system that is able to detect several oral and dental abnormal automatically.

REFERENCES

[1] D.A. Maharani, A. Rahardjo, “Mothers' dental health behaviors and mother-child's dental caries experiences: study of a suburb area in Indonesia”, *Makara Journal of Health Research*, pp.72-76, 2013.

[2] S. V. Tikhe, A. M. Naik, S. D. Bhide, T. Saravanan, K. P. Kaliyamurthie, “Algorithm to identify enamel caries and interproximal caries using

dental digital radiographs", In 2016 IEEE 6th International Conference on Advanced Computing (IACC) (pp. 225-228). IEEE, 2016.

[3] I. Mota-Veloso, M. E. C. Soares, B. M. Alencar, L. S. Marques, M. L. Ramos-Jorge, J. Ramos-Jorge, "Impact of untreated dental caries and its clinical consequences on the oral health-related quality of life of schoolchildren aged 8–10 years", *Quality of Life Research*, Vol. 25 No. 1, pp. 193-199, 2016.

[4] R. Ahmad, N. F. Affandi, N. A. F. M. Ayub, N. A. Mustafa, M. Y. P. M. Yusof, T. N. M. Dom, "The value of panoramic radiograph as a screening tool prior to complete denture construction: a restrospective study", *Journal of Dentistry Indonesia*, Vol. 26 No. 2, pp. 65-69, 2019.

[5] J. Naam, "Accuracy of Panoramic Dental X-Ray Imaging in Detection of Proximal Caries with Multiple Morpological Gradient (mMG) Method", *JOIV: International Journal on Informatics Visualization*, Vol. 1 No. 1, pp. 5-11, 2017.

[6] O. Kwon, T. H. Yong, S. R. Kang, J. E. Kim, K. H. Huh, M. S. Heo, W. J. Yi, "Automatic diagnosis for cysts and tumors of both jaws on panoramic radiographs using a deep convolution neural network". *Dentomaxillofacial Radiology*, Vol. 49 No. 8, 20200185, 2020.

[7] N. Kanuri, A. Z. Abdelkarim, S. A. Rathore, "Trainable WEKA (Waikato Environment for Knowledge Analysis) Segmentation Tool: Machine-Learning-Enabled Segmentation on Features of Panoramic Radiographs", *Cureus*, Vol. 14 No. 1, 2022.

[8] G. Silva, L. Oliveira, M. Pithon, "Automatic segmenting teeth in X-ray images: trends, a novel data set, benchmarking and future perspectives", *Expert Syst Appl*, 2018.

[9] J. Schmidhuber, "Deep learning in neural networks: an overview", *Neural Network*, Vol. 61, pp. 85–117, 2015.

[10] C. W. Wang, C. T. Huang, J. H. Lee, C. H. Li, S. W. Chang, M. J. Siao, C. Lindner, "A benchmark for comparison of dental radiography analysis algorithms", *Medical image analysis*, Vol. 31, pp. 63-76, 2016.

[11] Y. Bayraktar, E. Ayan, "Diagnosis of interproximal caries lesions with deep convolutional neural network in digital bitewing radiographs", *Clinical oral investigations*, Vol. 26 No. 1, pp. 623-632, 2022.

[12] J. H. Lee, D. H. Kim, S. N. Jeong, S. H.Choi, "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm", Journal of dentistry, Vol. 77, pp. 106-111, 2018.

[13] M. Estai, M. Tennant, D. Gebauer, A. Brostek, J. Vignarajan, M. Mehdizadeh, S. Saha, "Evaluation of a deep learning system for automatic detection of proximal surface dental caries on bitewing radiographs. Oral Surgery", *Oral Medicine, Oral Pathology and Oral Radiology*, 2022.

[14] S. Vinayahalingam, S. Kempers, L. Limon, D. Deibel, T. Maal, M. Hanisch, T. Xi, "Classification of caries in third molars on panoramic radiographs using deep learning", *Scientific Reports*, Vol. 11 No. 1, pp. 1-7, 2021.

[15] X., Lin, D. Hong, D. Zhang, M. Huang, H. Yu, "Detecting Proximal Caries on Periapical Radiographs Using Convolutional Neural Networks with Different Training Strategies on Small Datasets", *Diagnostics*, Vol. 12 No. 5, pp. 1047, 2022.

[16] Y. P. Huang, S. Y. Lee, "Deep learning for caries detection using optical coherence tomography", medRxiv, 2021.

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

[18] S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, "Aggregated residual transformations for deep neural networks", In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1492-1500), 2017.