Dental caries detection using faster region-based convolutional neural network with residual network

Andre Citro Febriliyan Lanyak[1], Agi Prasetiadi[1], Haris Budi Widodo[2], Muhammad Hisyam Ghani[2], Abiyan Athallah[2]
[1]Department of Informatics, Faculty of Informatics, Telkom Institute of Technology Purwokerto, Purwokerto, Indonesia
[2]Department of Dental Medicine, Faculty of Medicine, University of Jenderal Soedirman, Purwokerto, Indonesia

Article Info

Article history:

Received Jul 31, 2023
Revised Sep 29, 2023
Accepted Nov 7, 2023

Keywords:

Deep learning
Dental caries
Faster R-CNN
Object detection
Residual network

ABSTRACT

Dental caries is the highest prevalent dental disease in the world by 2022. Caries can be stopped by early detection of patients through efficient screening. Previously, there have been several methods used to detect caries such as single shot multibox detector (SSD), faster region-based convolutional neural network (Faster R-CNN) and you only look once (YOLO). This research aims to develop accurate dental caries detection using Faster R-CNN. Using a dataset collected from scraping on the internet, this research is started by creating an original dataset consisting of 81 base images which are then augmented to a total of 486 images and annotated by dental health experts from Jenderal Soedirman University. Transfer learning using pre-trained Faster R-CNN residual network (ResNet)-50 and ResNet-101 model is utilized to detect and localise dental caries. The Faster R-CNN ResNet-50 model trained using the Adam optimizer produces a mean average precision (mAP) of 0.213, and those using the momentum optimizer produce a mAP of 0.177. While the Faster R-CNN ResNet-101 model trained using the Adam optimizer produces a mAP of 0.192, and those using the momentum optimizer produce a mAP of 0.004. The model trained on the dataset showed satisfactory results in detecting dental caries, especially ResNet-50 with Adam optimizer.

Corresponding Author:

Andre Citro Febriliyan Lanyak
Department of Informatics, Faculty of Informatics, Telkom Institute of Technology Purwokerto
Purwokerto, Indonesia
Email: 19102274@ittelkom-pwt.ac.id

1. INTRODUCTION

Based on data from the Global Burden of Disease, it is estimated that as many as 3.5 billion people are affected by dental disease worldwide in 2019 [1]. Dental caries is the most common and prevalent dental diseases. In Indonesia alone approximately 89% of the population having suffered from caries [2]. This high prevalence of dental caries indicates a lack of access to dental care. This is supported by the study done by the basic health research or riset kesehatan dasar (RISKESDAS) of Indonesia with a sample of 300,000 families, and it was noted that the largest proportion of health problems in Indonesia was oral and dental disease with a proportion of 57.6% and only 10.2% received health services [3].

Dental caries is a condition in which damage occurs to the hard tissues of the teeth due to acidic by products of the bacterial fermentation process, which is produced from carbohydrates in food [4]. Dental caries begins with a microbiological shift in the bacterial biofilm (dental plaque) that covers the tooth surface and is influenced by key factors such as saliva flow, fluoride exposure, dietary sugar consumption, and preventive behaviour (tooth cleaning) [5]. Caries can be stopped by accurate early detection of patients through efficient screening. But accuracy of diagnosis of dental caries is still can be a challenging problem for dentists [6]. Deep learning application especially object detection can help practitioner to increase accuracy and reliability.

In recent year the use of object detection in caries detection has seen an increace [7]–[10], other then caries similar application of object detection in periodontal detection is observed [11], [12]. Juyal *et.al.* [7] published a study on the use of faster region-based convolutional neural network (Faster R-CNN) and you only look once (YOLO) to detect caries, with results showing that Faster R-CNN had higher accuracy. Zhang *et.al.* [8] use single shot multibox detector (SSD) with visual geometry group-16 (VGG-16) model to detect caries showing area under the curve (AUC) of 85.65%. Thanh *et.al.* [9] aim to develop a deep learning model capable of detecting caries using a smartphone camera, and one of the models uses Faster R-CNN with residual network (ResNet)-50. Lee *et.al.* [10] try to evaluate the efficacy of deep learning for detection of dental caries on periapical radiographs. Other than caries, there are some similar applications of deep learning in detecting other dental diseases like periodontal disease. Kim *et.al.* [11] developed a method to detect periodontal bone loss in panoramic dental radiographs. Thanathornwong and Suebnukarn [12] proposed a Faster R-CNN with ResNet-101 model to detect periodontal compromised teeth in digital panoramic radiographs, with result showing average precision (AP) rate of 0.81. The use of transfer learning is also observed, and it shows improvements in training time and model performance [13]. Some of the research mentioned above employed transfer learning in their methodology. For instance, VGG-16 was used in [8], ResNet50 and inception-ResNet-v2 were used in [9], GoogLeNet inception v3 was utilized in [10], and ResNet-101 was applied in [12]. This demonstrates that transfer learning is commonly utilized in object detection tasks.

With the increasing prevalence of object detection for dental diagnostics, particularly with the utilization of transfer learning, the author's interest has been piqued in exploring dental caries detection using an object detection model with transfer learning. Object detection methods can generally be categorized into two types: single-stage and two-stage detectors [14]. The former involves separating the detection process into a region proposal module and a classification module, while the latter integrates the detection into a single feed-forward fully convolutional network that directly provides the bounding boxes and object classification [15].

In this paper, a Faster R-CNN model with ResNet is employed to construct an object detection model for detecting dental caries. The study employs four different schemes with combination of the following, ResNet-50, ResNet-101, momentum optimizer, and Adam optimizer. The selection of Faster R-CNN model with ResNet is based on its demonstrated sweet spot between speed and accuracy [16]. With the above method this research aims to develop an object detection detector that can detect dental caries.

2. METHOD

2.1. Dataset

In this study, the dataset utilized for the methodology comprises 81 images containing dental caries. The initial dataset was obtained by scraping images from various online sources. To ensure accurate labeling and ground truth annotations, each image in the dataset was meticulously labeled by graduates from the Department of Dental Medicine at Jenderal Soedirman University, who possess real-world experience in dental diagnosis and pathology. Dataset then labelled using "LabelImg" a python based graphical image annotation tool. All the images are saved as 'XML' files with PASCAL VOC format. Furthermore, to enhance the dataset's diversity and generalizability, data augmentation techniques were employed. Various filters, such as Gaussian blur, sharpening, noise addition, grayscale conversion, and color channel manipulation (BGR), were applied to the images, resulting in an augmented dataset totaling 486 images. This filter application can be seen in Figure 1. These enhancements aim to simulate different imaging conditions and challenges that may be encountered in real-world dental imaging scenarios. The curated and augmented dataset will serve as a robust foundation for the dental caries detection algorithm, facilitating its training and evaluation with a comprehensive range of dental caries representations [17]. A separate dataset containing 14 images different from the 81 images training dataset is use to evaluate the model after training.

2.2. Faster region-based convolutional neural network

Faster R-CNN is an object detection architecture that consists of two modules. Figure 2 shows the architecture of the model. The first module is a deep fully convolutional network or region proposal network (RPN), this module is in charge of providing a proposal region. The second module is a detector that performs object detection on the proposal region. Feeding into those two modules is a feature mapping outputted by convolutional layer [18]. The RPN in Faster R-CNN generates proposal regions by sliding a small window over the convolutional feature map and creating multiple anchor boxes of different scales and aspect ratios at each spatial location. It predicts the probability of each anchor box containing an object and performs bounding box regression to adjust their coordinates. Non-maximum suppression (NMS) is applied to select the most confident and diverse bounding box candidates. The remaining high-scoring proposals become the final proposal regions used in subsequent stages for object detection, optimizing speed and accuracy compared to exhaustive sliding window approach [18], [19].

The proposed regions are then fed into the region of interest (RoI) pooling layer, which aligns them to a fixed size grid for further processing. The RoI pooling output is split into two branches: one for object classification, which predicts the probability of each region belonging to a specific class, and the other for bounding box regression, which refines the bounding box coordinates. NMS is applied to select the most confident and diverse bounding box predictions, ensuring efficient and accurate object detection [18].

Figure 1. Sample of the dataset

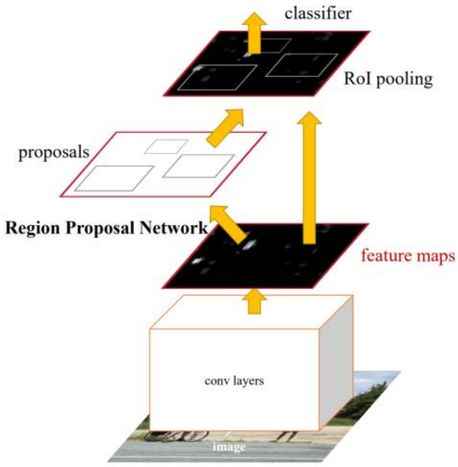


Figure 2. Faster R-CNN object detection architecture. It is a single unified network having two parts: RPN and classifier [18]

2.3. Residual network

ResNet is a deep convolutional neural network architecture that introduces residual block with skip connections to overcome the vanishing gradient problem in very deep networks [20], [21]. Figure 3 shows the skip connection. These skip connections learn the residual between the input and output and add it back to the original input to obtain the final output. By focusing on the residuals, the neural network only needs to learn the deviations or changes required to convert inputs into outputs. The concept of ResNet can be described by: $\mathcal{F}(x) = \mathcal{H}(x) - x$. In ResNet, the goal is to approximate an underlying mapping $\mathcal{H}(x)$ using stacked layers. Instead of directly approximating $\mathcal{H}(x)$, the approach is to approximate the residual function $\mathcal{F}(x) = \mathcal{H}(x) - x$. By doing this, the original function becomes $\mathcal{F}(x) + x$. The hypothesis is that multiple nonlinear layers can asymptotically approximate complicated functions and also asymptotically approximate the residual functions. While both forms can achieve the desired approximation, the ease of learning may vary. Corrected version: the ResNet model can achieve a depth of up to 152 layers, deeper than VGG-19, while maintaining lower complexity [20].

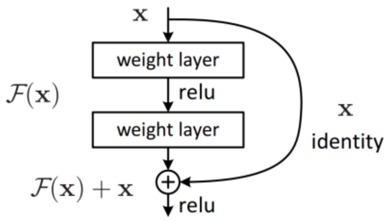
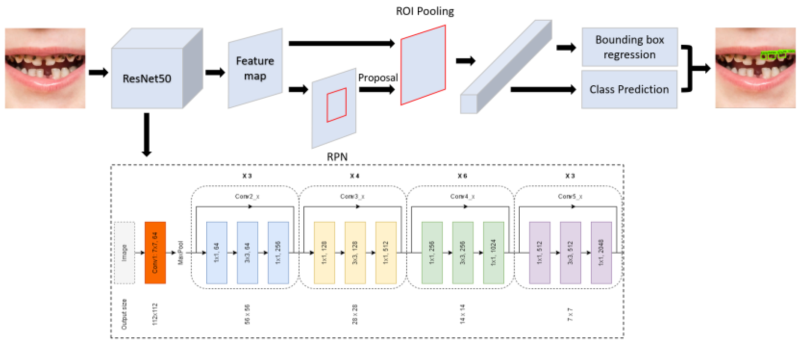


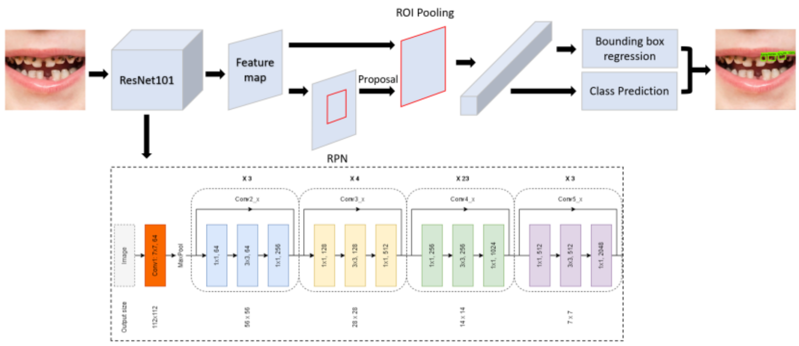
Figure 3. Building block of residual learning [20]

2.4. Model architecture

The method uses Faster R-CNN as the meta-architecture for object detection and use ResNet as feature extractor. Figure 4(a) shows the architecture with ResNet-50 and Figure 4(b) shows the architecture with ResNet-101. To expedite model training and enhance performance, we leverage transfer learning by using pre-trained Faster R-CNN models. The Faster R-CNN framework forms the foundation of our detection model, facilitating region proposal and object classification. For feature extraction, we integrate ResNet, a competitive deep convolutional neural network capable to capture intricate features from images [22]. Specifically, two model variants are investigated, with one utilizing ResNet-50 and the other employing ResNet-101 as the feature extractor.



(a)



(b)

Figure 4. Faster R-CNN model with (a) ResNet-50 and (b) ResNet-101

2.5. Training scheme and strategy

Table 1 shows the method model scheme. The method involves training two variants of pre-trained Faster R-CNN ResNet model: ResNet-50 and ResNet-101. For each model variant, we employ two different optimizers: Adam optimizer and Momentum optimizer. The goal is to explore the impact of different feature extractors (ResNet-50 and ResNet-101) and optimizer choices on the model's performance. By training the models using both Adam optimizer and Momentum optimizer, we can assess their respective effectiveness in optimizing the model's parameters and improving the convergence speed during training. Each model variant is trained for 15,000 steps to allow the model to learn from the dataset thoroughly. In total there four model that will be trained. The training proceses utilize object detection pipeline form Tensorflow object detection application programming interface (API). The dataset is converted into tfrecord format to facilitate efficient and streamlined training. The training pipeline automatically resizes the images, removing the need for manual image resizing beforehand. The training scheme allows us to comprehensively evaluate the impact of different model architectures (ResNet-50 and ResNet-101) and optimizer choices on the object detection performance. By leveraging the training pipeline and optimizing strategies, we aim to achieve accurate and efficient object detection results for caries detection. Moreover, we also employ the cosine decay learning rate, which prior studies have shown to improve model performance [23].

Table 1. The model scheme

Model	Feature extractor	Optimizer
M1	ResNet-50	Momentum
M2	ResNet-50	Adam
M3	ResNet-101	Momentum
M4	ResNet-101	Adam

2.6. Analysis

To assess the performance of Faster R-CNN in detecting caries, this study employed mean average precision (mAP) as evaluation metrics. The mAP was calculated by averaging the AP of all the classes with (1) as [24]:

$$mAP = \frac{1}{N}\sum_{i=1}^{N} AP_i \tag{1}$$

Meanwhile AP indicating the area under the precision-recall curve. AP is calculated by averaging all the precision from precision-recall interpolated pairs. AP is calculated at several intersection over union (IoU) thresholds. AP can be calculated using (2) [24]:

$$AP = \sum_{k=0}^{K}(R_r(k) - R_r(k+1))Rr_{interp}\,(R_r(k)) \tag{2}$$

Precision is calculated by dividing accumulated true positif detection by all detection. Recall is calculated by deviding true positif detection by all ground truth. Precision and recall are calculated using (3) and (4) [24]:

$$Pr = \frac{\sum_{n=1}^{S} TP_n}{all\ detection} \tag{3}$$

$$Pc = \frac{\sum_{n=1}^{S} TP_n}{all\ ground\ truth} \tag{4}$$

The approach of calculating mAP at several IoU thresholds serves as a reward mechanism for the detector, favoring models with better localization and more accurate detection. By considering multiple IoU thresholds, the evaluation takes into account various degrees of overlap between predicted bounding boxes and ground truth annotations, providing a comprehensive assessment of the model's performance. The use of mAP as evaluation metrics allows for a thorough evaluation of the Faster R-CNN model's ability to precisely and effectively detect caries in the given dataset, enhancing our understanding of its suitability for caries detection tasks [25]. In addition to mAP, the precision-recall curve was also utilized for evaluation. This curve visually depicts the trade-off between precision and recall across different confidence thresholds, offering valuable insights into the model's performance across varying levels of confidence [24]. By employing a combination of evaluation metrics, this study aims to provide a thorough and robust assessment of the Faster R-CNN model's efficacy in detecting caries.

3. RESULTS AND DISCUSSION

3.1. Result

Table 2 presents the results of all the models. Each model is evaluated using evaluation dataset containing images never seen by the model before. Model M1 achieved a mAP of 0.177 at IoU 0.50:0.95, while model M2 obtained a higher mAP of 0.213. On the other hand, model M3 exhibited a much lower mAP of 0.004, and model M4 achieved a mAP of 0.192. These results were consistent across different IoU thresholds, such as at IoU 0.50 and 0.75, where model M1 obtained mAP scores of 0.465 and 0.093, respectively. Model M2 demonstrated higher accuracy, with mAP scores of 0.527 at IoU 0.50 and 0.158 at IoU 0.75. Meanwhile, model M3 obtained mAP scores of 0.017 and 0.000 at the same IoU thresholds, and model M4 achieved mAP scores of 0.497 and 0.126. The results indicate that models M2 performed better than models M1, M3, and M4; and model M3 show the worse result among all the model trained.

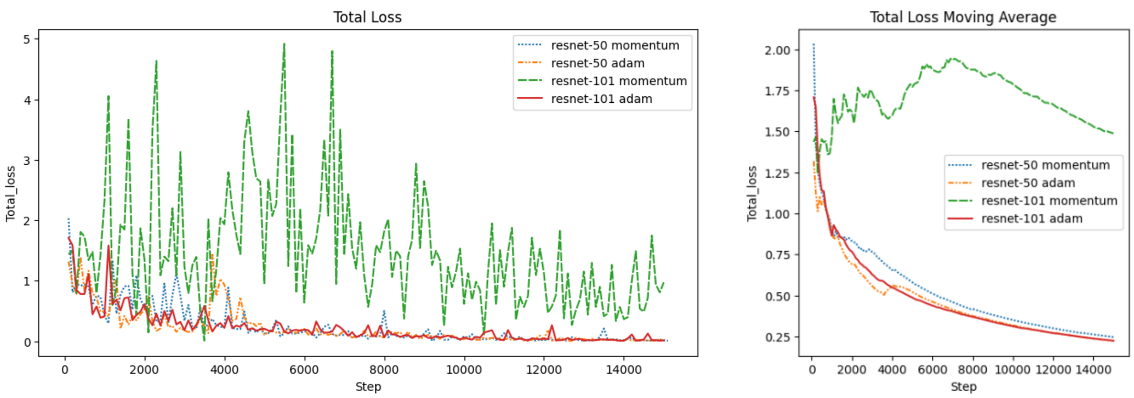
Table 2. Evaluation result of the model

Metric	M1	M2	M3	M4
mAP (IoU 0.50:0.95)	0.177	0.213	0.004	0.192
mAP (IoU 0.50)	0.465	0.527	0.017	0.497
mAP (IoU 0.50:0.95)	0.093	0.158	0.000	0.126

3.2. Discussion

The model was trained using the TensorFlow object detection API training pipeline for 15,000 steps. Evaluating the training results of the four models revealed that three out of the four models showed promising performance. Specifically, models M1, M2, and M4 demonstrated encouraging results, as they managed to consistently reduce the total loss and converge, even though they experienced some fluctuations during the initial steps of training. On the other hand, model M3 exhibited continuous fluctuations throughout the training process, which set it apart from the other three models. Figure 5 illustrates the total loss of all the models during training, where we observe that model M3 consistently experiences fluctuations from the start until the end of the training process, while models M1, M2, and M4 show a more stable and gradual decrease in total loss. The higher fluctuation in model M3 is likely due to the complexity of ResNet-101, which has twice as many layers as ResNet-50. Nonetheless, model M1 with the same optimizer and base learning rate successfully converged and showed good performance on evaluation dataset.

Figure 5(a) show total loss of all the model trained and Figure 5(b) shows the cumulative moving average (CMA) of total loss. Comparing models M1, M2, and M4 also reveals that models M2 and M4 converge almost at the same time, but model M4 has a slight advantage. In contrast, model M1 converges slower and exhibits less stability compared to models M2 and M4. The relatively faster and more stable convergence of models M2 and M4 indicates their potential for accurate and efficient object detection, while model M1 may require further optimization to improve its convergence and overall performance. Further analysis and fine-tuning of model M3 may be required to enhance its stability and convergence for effective object detection. Overall model M4 had the shown the lowest total loss after 15000 steps, followed by models M2, M1, and M3. With total loss values of 0.008699 for model M4, 0.012953 for model M2, 0.016728 for model M1, and 0.973653 for model M3, respectively. These findings highlight the importance of selecting an appropriate architecture and optimizer for object detection tasks. The differences in convergence behavior among these models imply that the choice of architecture and optimizer can significantly impact the training process and overall performance of the model.



(a) (b)

Figure 5. Total loss during training (a) total loss and (b) CMA of total loss

Looking at the total model loss alone is not enough to determine the best model; hence, another metric is needed as a benchmark. One of the standard metrics commonly used for evaluating object detection models is mAP. Precision recall curve also is use to evaluate the models trained shown in Figure 6. mAP is calculated as the average of the AP scores for each class in multi-class scenarios. However, since this research involves only one class, mAP and AP are equivalent. For this study, mAP based on common objects in context (COCO) is utilized, as opposed to mAP based on PASCAL VOC. Notably, mAP COCO measures the AP at 10 IoU points, ranging from 0.50 to 0.95 with 0.05 increments, providing a more comprehensive evaluation of the model's performance. Conversely, PASCAL VOC only calculates mAP at IoU 0.50. Additionally, IoU values at 0.50 and 0.75 are also considered in the analysis.

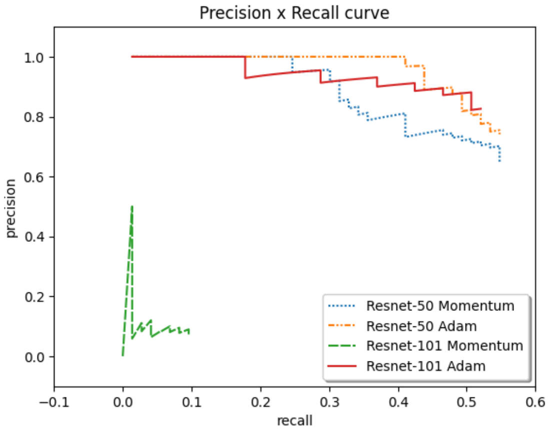


Figure 6. Precision-recall curve of all the model

Based on Table 2 and the result above its can be concluded the best model is model M2 follow by M4, M1, and M3. Although the M2 model has the best performance based on metrics, this advantage is only 2-7% compared to other models. In addition, the precision-recall curve also supports the above conclusion. According to the precision-recall curve shown in Figure 6, model M2 exhibits the best performance, followed by models M4, M1, and M3. Notably, the precision of model M2 remains high up to a recall of 0.4, in contrast to models M1 and M4, which experience a decrease in precision at around 0.2 recall. On the other hand, model M3 aligns with the aforementioned conclusions and displays the poorest results among all the models evaluated. The precision of model M2 stays consistently high as the recall increases, indicating that this model can provide accurate results for detecting caries. Figure 7 shows the detection results from model M2 on an image from the evaluation dataset.

Figure 7. Detection result of model M2, red box indicate ground truth box, and green boxes indicate the prediction boxes

4. CONCLUSION

Based on the results and discussions, we can draw several conclusions regarding the dental caries detection models. Firstly, three out of the four trained models (M1, M2, and M4) exhibit the capability to detect dental caries effectively. Secondly, comparing the optimizers used in the models, it becomes evident that employing the Adam optimizer (M2 and M4) leads to the development of more accurate models in contrast to the Momentum optimizer (M1 and M3). Additionally, concerning model performance, the ResNet-50 architecture stands out as the better performer for this particular case study. With detail result of mAP as follow, ResNet-50 with momentum optimizer (M1) achieved a mAP of 0.177, while ResNet-50 with Adam optimizer (M2) achieved a mAP of 0.213. On the other hand, ResNet-101 with momentum optimizer (M3) obtained a mAP of 0.004, and ResNet-101 with Adam optimizer (M4) achieved a mAP of 0.192.

REFERENCES

[1] B. A. Dye, "The global burden of oral disease: research and public health significance," *Journal of Dental Research*, vol. 96, no. 4, pp. 361–363, Apr. 2017, doi: 10.1177/0022034517693567.

[2] C. M. A. Santoso, T. Bramantoro, M. C. Nguyen, Z. Bagoly, and A. Nagy, "Factors affecting dental service utilisation in Indonesia: a population-based multilevel analysis," *International Journal of Environmental Research and Public Health*, vol. 17, no. 15, pp. 1-11, Jul. 2020, doi: 10.3390/ijerph17155282.

[3] S. Susilawati *et al.*, *Indonesian Oral Health Survey Implementation-National Basic Health Research (RISKESDAS) 2018*. Jakarta: Monograph Press, 2020, doi: 10.32793/monograph.v1i1.605.

[4] R. H. Selwitz, A. I. Ismail, and N. B. Pitts, "Dental caries," *The Lancet*, vol. 369, no. 9555, pp. 51–59, Jan. 2007, doi: 10.1016/S0140-6736(07)60031-2.

[5] I. Strużycka, "The oral microbiome in dental caries," *Polish Journal of Microbiology*, vol. 63, no. 2, pp. 127-135, 2014.

[6] H. Mohammad-Rahimi *et al.*, "Deep learning for caries detection: a systematic review," *Journal of Dentistry*, vol. 122, Jul. 2022, doi: 10.1016/j.jdent.2022.104115.

[7] A. Juyal, H. Tiwari, U. K. Singh, N. Kumar, and S. Kumar, "Dental caries detection using Faster R-CNN and YOLO V3," *ITM Web of Conferences*, vol. 53, pp. 1-16, Jun. 2023, doi: 10.1051/itmconf/20235302005.

[8] X. Zhang *et al.*, "Development and evaluation of deep learning for screening dental caries from oral photographs," *Oral Diseases*, vol. 28, no. 1, pp. 173–181, Jan. 2022, doi: 10.1111/odi.13735.

[9] M. T. G. Thanh, N. Van Toan, V. T. N. Ngoc, N. T. Tra, C. N. Giap, and D. M. Nguyen, "Deep learning application in dental caries detection using intraoral photos taken by smartphones," *Applied Sciences*, vol. 12, no. 11, pp. 1-10, May 2022, doi: 10.3390/app12115504.

[10] 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," *Journal of Dentistry*, vol. 77, pp. 106–111, Oct. 2018, doi: 10.1016/j.jdent.2018.07.015.

[11] J. Kim, H.-S. Lee, I.-S. Song, and K.-H. Jung, "DeNTNet: deep neural transfer network for the detection of periodontal bone loss using panoramic dental radiographs," *Scientific Reports*, vol. 9, no. 1, pp. 1-9, Nov. 2019, doi: 10.1038/s41598-019-53758-2.

[12] B. Thanathornwong and S. Suebnukarn, "Automatic detection of periodontal compromised teeth in digital panoramic radiographs using faster regional convolutional neural networks," *Imaging Science in Dentistry*, vol. 50, no. 2, pp. 169–174, 2020, doi: 10.5624/isd.2020.50.2.169.

[13] J. Wang, X. He, S. Faming, G. Lu, H. Cong, and Q. Jiang, "A real-time bridge crack detection method based on an improved inception-resnet-v2 structure," *IEEE Access*, vol. 9, pp. 93209–93223, 2021, doi: 10.1109/ACCESS.2021.3093210.

[14] Z. Zou, K. Chen, Z. Shi, Y. Guo, and J. Ye, "Object detection in 20 years: a survey," *Proceedings of the IEEE*, vol. 111, no. 3, pp. 257-276, Mar. 2023, doi: 10.1109/JPROC.2023.3238524.

[15] M. Carranza-García, J. Torres-Mateo, P. Lara-Benítez, and J. García-Gutiérrez, "On the performance of one-stage and two-stage object detectors in autonomous vehicles using camera data," *Remote Sensing*, vol. 13, no. 1, pp. 1-23, Dec. 2020, doi: 10.3390/rs13010089.

[16] J. Huang *et al.*, "Speed/accuracy trade-offs for modern convolutional object detectors," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 3296-3297, doi: 10.1109/CVPR.2017.351.

[17] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019, doi: 10.1186/s40537-019-0197-0.

[18] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, vol. 28, pp. 1-9, Jun. 2015.

[19] M. Moniruzzaman, S. M. S. Islam, P. Lavery, and M. Bennamoun, "Faster R-CNN based deep learning for seagrass detection from underwater digital images," in *2019 Digital Image Computing: Techniques and Applications (DICTA)*, Dec. 2019, pp. 1–7, doi: 10.1109/DICTA47822.2019.8946048.

[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[21] E. Rezende, G. Ruppert, T. Carvalho, F. Ramos, and P. de Geus, "Malicious software classification using transfer learning of resnet-50 deep neural network," in *2017 16th IEEE International Conference on Machine Learning and Applications (ICMLA)*, Dec. 2017, pp. 1011–1014, doi: 10.1109/ICMLA.2017.00-19.

[22] A. Fuentes, S. Yoon, S. Kim, and D. Park, "A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition," *Sensors*, vol. 17, no. 9, pp. 1-21, Sep. 2017, doi: 10.3390/s17092022.

[23] T. He, Z. Zhang, H. Zhang, Z. Zhang, J. Xie, and M. Li, "Bag of tricks for image classification with convolutional neural networks," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2019, pp. 558–567, doi: 10.1109/CVPR.2019.00065.

[24] R. Padilla, W. L. Passos, T. L. B. Dias, S. L. Netto, and E. A. B. da Silva, "A comparative analysis of object detection metrics with a companion open-source toolkit," *Electronics*, vol. 10, no. 3, pp. 1-28, Jan. 2021, doi: 10.3390/electronics10030279.

[25] R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Jul. 2020, pp. 237–242, doi: 10.1109/IWSSIP48289.2020.9145130.

BIOGRAPHIES OF AUTHORS

Andre Citro Febriliyan Lanyak from Makassar, currently studying for a Bachelor's Degree in Informatics Engineering at the Telkom Institute of Technology, Purwokerto. Active student in campus participating in multiple events and extracurricular activity. Former chair of Astralic English Club in ITTP and Chess Club. His research areas include data science engineer, machine learning, deep learning, computer vision, and object detection. He can be contacted at email: 19102274@ittelkom-pwt.ac.id.

Agi Prasetiadi holds a Bachelor of Electronics degree in General Electrical Engineering, VLSI and Computer Vision at Bandung Institute of Technology. He holds a Master of Engineering (M.Eng.) in Realtime Systems, Bio-Inspired Technology at Kumoh National Institute of Technology. He currently teaches at Telkom Institute of Technology Purwokerto as a lecturer in Informatics. His research areas include computer vision, machine learning, and deep learning. He can be contacted at email: agi@ittelkom-pwt.ac.id.

Haris Budi Widodo is an academic staff at Department of Public Health, Jenderal Soedirman University, Indonesia. He received his Ph.D. from Airlangga University Indonesia. He also received his M.S. degree and bachelor degree from Gadjah Mada University Indonesia. His research interest including public health, e-health, and telemedicine. He can be contacted at email: harisbudiwidodo@gmail.com.

Muhammad Hisyam Ghani graduated from Bachelor of Dentistry (S.KG.), Universitas Jenderal Soedirman, Banyumas in 2021. Recently conducted research on the antibacterial activity of begonia multangular blume hand extract against fusobacterium nucleatum biofilm. Currently undergoing Professional Dentistry Education at Universitas Jenderal Soedirman. He can be contacted at email: hisyam.ghani@mhs.unsoed.ac.id.

Abiyan Athallah graduated with a Bachelor of Dentistry degree from Jenderal Soedirman University, Purwokerto, Central Java. His latest research conducted on the effect of rice husk nanocellulose on the shear strength of composite resins. Currently, he is studying Professional Dentistry Education at Universitas Jenderal Soedirman. He can be contacted at email: abiyan.a@mhs.unsoed.ac.id.