Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary : Rapid Review : Open Access Journal eISSN: 2656-8632

RESEARCH ARTICLE OPEN ACCESS

Manuscript received August 03, 2022; revised October 4, 2022; accepted October 10, 2022; date of publication October 30, 2022
Digital Object Identifier (**DOI**): https://doi.org/10.35882/jeeemi.v4i4.255

How to cite: Zendi Iklima, Trie Maya Kadarina, and Rinto Priambodo, "Reduction of Feature Extraction for COVID-19 CXR using Depthwise Separable Convolution Network for Convolutional Performance Consistency", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 4, no. 4, pp. 204–209, October. 2022

Reduction of Feature Extraction for COVID-19 CXR using Depthwise Separable Convolution Network for Convolutional Performance Consistency

Zendi Iklima[1], Trie Maya Kadarina[1], and Rinto Priambodo[1]

[1]Universitas Mercu Buana, Jakarta, Indonesia

Corresponding author: Zendi Iklima (e-mail: zendi.iklima@mercubuana.ac.id).

This study was supported by Department and Research Centre of Universitas Mercu Buana.

ABSTRACT A Convolutional Neural Network (CNN) classifier is generally utilized to classify an image tensor according to the mapped labels. The simplification of the classifier causes CNN to be often used to classify images, especially in the biomedical field. Thus, CNN is widely used to classify computer tomography (CT) and chest X-ray (CXR) images against the mapped labels. Several transfer learning models were implemented to classify CXR images for preliminary detection of COVID-19 infection, e.g., ResNet, Inception, Xception, etc. However, a transfer learning model has a maximum and minimum input resolution. Thus, the computational cost tends to be huge and unable to be optimized. Therefore, A custom CNN model can be a solution to reduce computational costs by configuring the feature extraction layers. This study proposed an efficient reduction of feature extraction for COVID-19 CXR namely Depthwise Separable Convolution Network. Furthermore, numerous strategies were adopted to lower the computational cost while retaining accuracy, including customizing the Batch Normalization (BN) layer and replacing the convolution layer with a separable convolution layer. The proposed model successfully reduced the feature extraction represented by the decreases in trainable parameters from 28.640 trainable parameters to 4.640 trainable parameters. The CNN was successfully trains COVID-19 CXR with accuracy 72.96%, loss 12.43%, recall 74.67%, precision 77.67%, and F1-score 75.33%. Meanwhile, the depthwise separable convolution effectively retains the performance accuracy 71.07%, loss 11.55%, recall 70.00%, precision 77.67%, and F1-score 70.00%. The CXR augmentation is also successfully increase the performance accuracy 74.55%, loss 11.37%, recall 77.67%, precision 79.56%, and F1-score 78.33%. Based on these results, the depthwise separable layers simplify the CNN layers by maintaining the performance of the overall architecture. The accuracy improvement can be initiated by enriching the dataset in the image augmentation or synthetic image stage.

INDEX TERMS CNN, COVID-19 CXR, Transfer Learning, Feature Extraction, Depthwise Separable Convolution Network.

I. INTRODUCTION

The Deep Learning technique is an intriguing topic for developing an automated system for diagnosing the presentation of COVID-19 viral using Chest X-Ray images [1], [2]. A Convolutional Neural Network (CNN) is a deep learning classifier that is commonly used to categorize an image tensor based on mapped labels. Because of the classifier's simplicity, CNN is frequently used to categorize images, particularly in the biomedical areas. As a result, CNN is commonly used to categorize computer tomography (CT) and chest X-ray (CXR) iamges against mapped labels. Several transfer learning models, such as ResNet, MobileNet, Xception, and others, were used to identify CXR images for preliminary detection of COVID-19 infection.

The deep convolutional neural network (DCNN) is a complex architecture of CNN. DCNN was customized as a COVID-19

CXR classifier to identify infected and uninfected labels. CovidNet has 1.75 million trainable parameters that are trained with 93.3% accuracy [3]. Parallel-Dilated COVIDNet is a COVID-19 identification method from chest X-ray images based on a parallel-dilated convolutional neural network (PDCOVIDNet). First, the publically accessible chest X-ray collection was fully preloaded and upgraded, and then the suggested approach was used to classify the images. The proof-of-principle for utilizing PDCOVIDNet to extract radiological characteristics for COVID-19 detection is demonstrated by varying the convolution dilation rate in parallel. PDCOVIDNet was trained using 2905 chest X-ray pictures from three different instances (COVID-19, normal, and viral pneumonia), and empirical tests demonstrated that the proposed technique identified more relevant characteristics associated to probable illness more quickly. The experimental findings show that our suggested strategy increases performance measures significantly: accuracy, precision, recall, and F1 scores are 96.5%, 96.58%, 96.59%, and 96.5%, respectively [4]. 498 CXR images of infected labels, 1583 images images of normal labels, and 4273 CXR images of pneumonia labels were used to train a costumize DCNN. This model has 63.15 million trainable parameters and has an accuracy of 87.3%, recall of 84.46%, and precision of 89.67% [5]. A COVIDX-Net system for COVID-19 classification was investigated utilizing 7 deep learning classifiers: VGG19, DenseNet121, ResNetV2, InceptionV3, InceptionResNetV2, Xception, and MobileNetV2. The COVIDX-Net performs in 89% of normal label F1-scores and 91% of COVID-19 label F1-scores [6]. The separable convolution approach is used to reduce the CNN architecture's trainable parameter. This design successfully decreases the number of trainable parameters from 23.57 million to 0.84 million [7]. Xception is a DCNN built with depthwise convolution layers followed by pointwise layers. This model was trained on 180 CXR images and then evaluated on 11302 CXR photos. The accuracy of the COVID-19 identification method is 91.4% [8]. DenseNet103's segmentation-based approach demonstrates consistently better performance than the overall global approach to metrics. In particular, the ResNet18 method as a classifier showed a sensitivity of 92.5% for Covid-19 and viral pneumonia [9]. A simple CNN model was examined with a 224x224 image input consisting of a non-COVID-19 COVID-19 label with 140 images to train the model and 160 images to test the model. The accuracy obtained is 95% in 20 iterations [10].
However, a transfer learning model has an input resolution constraint. As a result, the computational cost is sometimes huge and cannot be reduced. As a consequence, a customized CNN model may be employed to reduce computing costs by adjusting the feature extraction layers. Depthwise Separable Convolution Network was proposed as an efficient feature extraction reduction approach for COVID-19 CXR in this paper. Furthermore, numerous strategies were employed to lower computational cost while retaining accuracy, including adjusting the Batch Normalization (BN) layer and replacing the convolution layer with a separable convolution layer. Additionally, to improve the accuracy, the model will be trained using a synthetic dataset that was produced by our previous study [11]. Therefore, the main contribution of this paper as follows:

1) Feature reduction of the conventional CNN can be implemented by updating the convolution layer within the depth-dependent and point-wise layer.
2) The BN layer is placed to retain the accuracy of the depthwise separable network.
3) This paper used the synthetic dataset to improve the model accuracy while trained using depthwise separable network.

II. MATERIAL AND METHOD

A. Runtime Specification

The training process was performed using Google Colaboratory Pro Instance Intel(R) Xeon(R) 4xvCPU @2.30GHz, 30 GB RAM, and GPU using NVIDIA P100, Peripheral Component Interconnect Express (PCI Express) 16 GB.

B. Research Flow

FIGURE 1 represents the research flow consists of particular stages such as dataset distribution, dataset augmentation, model initialization, model training, model evaluation, model testing, and export trained model.

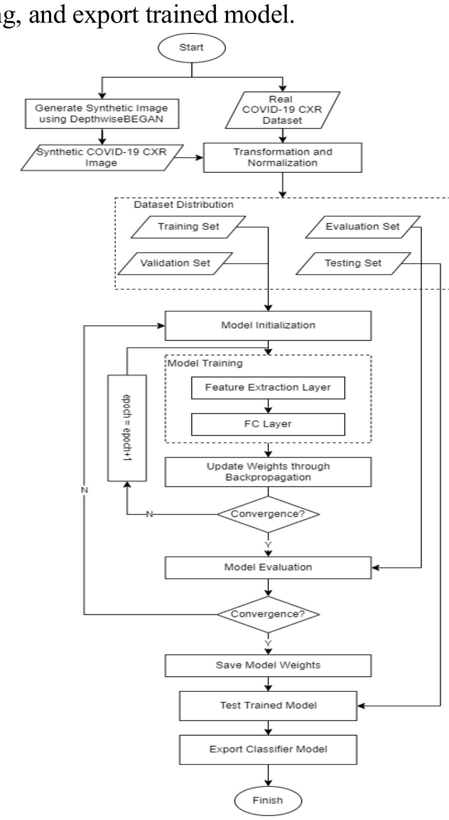


FIGURE 1. **Research Flow**

To retraining and improving the model accuracy, our previous study was produces synthetic dataset of COVID-19 CXR. The image transformation process is carried out to get the same image resolution. The distribution of the dataset is divided into 4 sets, each containing a training set of 63%, a validation set of 18%, an evaluation set of 9%, and a testing set of 10%. The training process is repeated until the iteration limit is reached. If the model's assessment findings are adequate, there is no need to modify the CNN architecture.

C. COVID-19 CXR Dataset

The real COVID-19 CXR dataset was provided on Kaggle. The distribution of the dataset was split into 4 sets. The training set contains 363 images of COVID-19 labels (shown in FIGURE 2 (a)), 997 images of normal labels (shown in FIGURE 2 (b)), and 2692 images of pneumonia labels (shown in FIGURE 2 (c)). The validation set contains 104 images of COVID-19 labels, 285 images of normal labels, and 769 images of pneumonia labels. The evaluation set contains 52 images of COVID-19 labels, 143 images of normal labels, and 385 images of pneumonia labels. The testing set contains 30 images of COVID-19 labels, 30 images of normal labels, and 30 images of pneumonia labels [12].

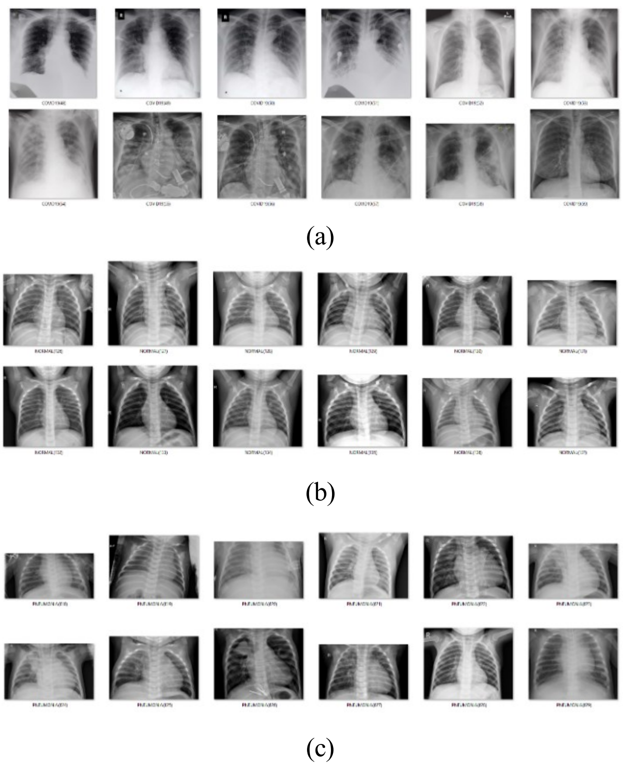


FIGURE 2. **COVID-19 CXR dataset (a) Infected Label, (b) Normal Label, and (c) Pneumonia Label**

D. Convolutional Neural Network

Generally, convolutional neural networks (CNN) consist of a feature extraction layer and a fully connected layer. Dimension reduction of the input tensor occurs in the feature extraction layer, which consists of a convolution layer, an activation layer (rectifier linier unit), and a pooling layer. To transform a 2D representation into an input of a fully connected layer, a flattening layer is utilized as a bridge between the convolution layer and the fully connected layer [13].

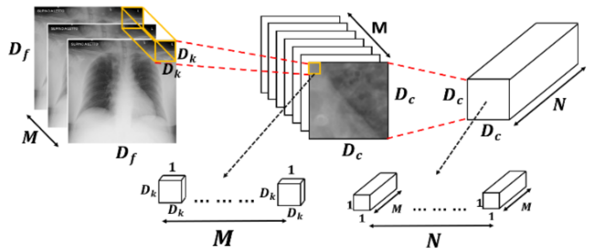


FIGURE 3. **Feature Extraction in Convolution**

Suppose, $D_f \times D_f$ denotes as an input matrix per M channel, $D_f \times D_f \times M$ denotes as filter size, and $D_c \times D_c \times N$ denotes as the output of the convolution operators. Thus, the total trainable parameter of standard convolution formulized as:

$$D_k^2 \times D_c^2 \times M \times N \tag{1}$$

E. Depthwise Separable Convolution

Due to the amount of convolution parameters in the kernel that must be computed in a matrix calculation, CNN models with high quality pictures demand greater memory allocation. As a result, a number of CNN models may be reduced in order to minimize convolution parameters.

Depthwise Separable Convolution (DSC) is one model that reduces the total number of parameters, reduces matrix calculations, and retrains accuracy. The standard CNN standard employs a convolution kernel with the same input channels, allowing matrix calculations to be performed channel by channel. DSC is made up of two convolution processes: Depthwise and Pointwise. Each NN input channel is filtered independently by a distinct filter in depthwise convolution, i.e., spatial convolution. The characteristics between channels are then projected into the new features using pointwise convolution, which is classic convolution with 1×1 windows. Thus, the total trainable parameter of depthwise separable convolution formulized as [14]–[16]:

$$D_k^2 \times D_c^2 \times M \tag{2}$$

$$D_c^2 \times M \times N \tag{3}$$

$$(D_c^2 \times M)(D_k^2 + N) \tag{4}$$

The total number of depthwise convolution parameter denoted by Equation (2). The total number of pointwise convolution parameter denoted by Equation (3). Therefore, the total number of DSC parameter denoted by Equation (4).

Based on Equation (1) and Equation (4), the trainable parameter comparison of standard CNN and DSC denotes as [15]–[17]:

$$\frac{pDSC}{pCNN} = \frac{(D_k^2+N)}{D_k^2 \times N} = \frac{1}{N} + \frac{1}{D_k^2} \tag{4}$$

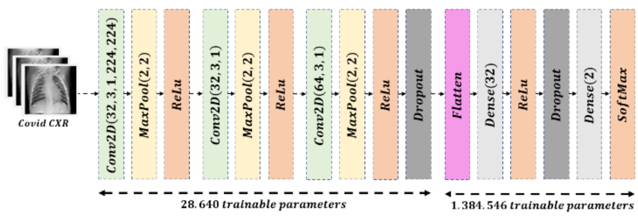
It is proved that the DSC parameters and computations are only $\frac{1}{N}+\frac{1}{D_k^2}$ times of the ordinary convolution. This significantly decreases the model's parameters and computation expenses.

III. RESULT AND DISCUSSION

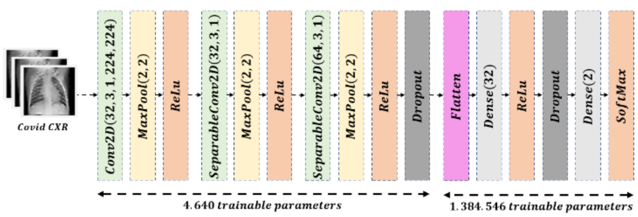
A. PROPOSED ARCHITECTURE

The depthwise separable convolution network proposed in this paper was trained on a synthetic CXR dataset. The image size was set to 224x224 resolution, 50 iterations, and 0.5 dropout as an initial hyperparameter of the training process. To avoid overfitting models, the augmentation setup was configured. The augmentation setup includes a rescale of 1/255, a rotation range of 2 degrees, a zoom range of 0.1, a horizontal flip, and a vertical flip. This training procedure makes use of the 5216 CXR dataset [18] which containing 1341 normal label, 2530 bacteria label, and 1345 virus label.

The proposed architecture includes three convolutional layers with a total of 32, 32, and 32 filters, respectively. This layer employs three kernels and the ReLu activation function. The max-pooling layer was built with a pool size of 2x2 and placed in front of the convolution layer. The fully connected layer was created in two layers, with the most recent layer representing the desired image labels. The first proposed conventional CNN architecture is depicted in Figure 4(a), while the DSC architecture is depicted in Figure 4(b).



(a)



(b)

FIGURE 4. Proposed Architecture (a) Conventional CNN and (b) DSC

According to Figure 4, the DSC was reduced to 83.79% of the proposed architecture's trainable parameters. Each proposed architecture's fully connected layer has trainable parameters that are similar. The proposed models were then compared to those with architectures, as shown in Table 1.

TABLE 1
Trainable Parameters

Description	Parameters
Learning Rate	1e-4
Batch Size	32
Optimizer	Adam
Loss Function	Categorical Crossentropy
Epoch	50
GPU Worker	4
Flipping	True
Rotation Range	1°
Rescaling	1/255
Zoom Range	0.05

B. MODEL PERFORMANCE

The training process was performed using Google Colaboratory Instance Intel(R) Xeon(R) 4xvCPU @2.30GHz, 30 GB RAM, and GPU using NVIDIA P100, Peripheral Component Interconnect Express (PCI Express) 16 GB.

TABLE 2
Model Performance

Models	Accuracy / Loss (%)	Precision / Recall (%)	F1-Score (%)
Model-1	72.96 / 12.43	77.67 / 74.67	75.33
Model-2	71.07 / 11.55	77.67 / 70.00	70.00
Model-1 Aug.	74.55 / 11.37	79.56 / 77.67	78.33
Model-2 Aug.	73.97 / 10.54	79.56 / 75.67	78.33
ResNet50-v2 [8]	63.74 / 21.88	68.67/63.00	62.33
Inception-v3 [19]	54.32 / 35.44	57.00/54.00	47.33
ResNet50-v2 Aug.	66.64 / 22.77	70.33/65.00	66.67
Inception-v3 Aug.	56.41 / 33.41	63.00/57.00	53.00

Table 2 shows the model performance with and without augmentation. The augmentation successfully increased the model performance of the CNN models. To demonstrate that it can classify well given the distribution of different datasets, the trained model required to be validated. The proposed model compared to pre-trained models contains ResNet50-v2 [8] and Inception-v3 [19]. Figure 5 shows the performance comparison of the CNN models.

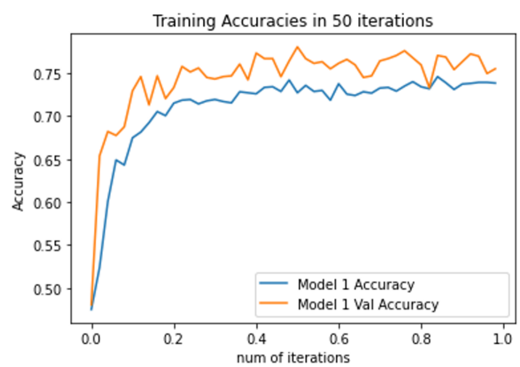


FIGURE 5. Model Performance

Based on Figure 5, the proposed model shows a great performance according to the training indicators. In order to analyze the models' great performance in the whole stage (training, validation, and testing), the confusion matrices were calculated in Figure 5 [20].

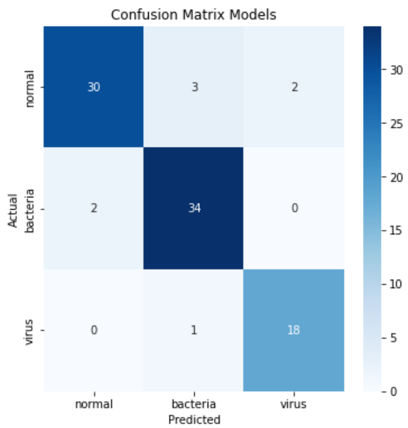


FIGURE 6. **Classification Test**

The depthwise separable convolution effectively retains the performance accuracy 72.96%, loss 12.43%, recall 74.67%, precision 77.67%, and F1-score 75.33%. The CXR augmentation is also successfully increase the performance accuracy 74.55%, loss 11.37%, recall 77.67%, precision 79.56%, and F1-score 78.33%.

IV. CONCLUSION

In general, a Convolutional Neural Network (CNN) classifier is used to classify an image tensor based on the mapped labels. Because of the classifier's simplicity, CNN is frequently used to classify images, particularly in the biomedical field. As a result, CNN is commonly used to classify computer tomography (CT) and chest X-ray (CXR) images against mapped labels. Several transfer learning models, such as ResNet, Inception, Xception, and others, were used to classify CXR images for preliminary detection of COVID-19 infection. However, a transfer learning model has a maximum and minimum input resolution. As a result, the computational cost is usually enormous and cannot be optimized. As a result, configuring the feature extraction layers in a custom CNN model can help to reduce computational costs. This study proposed Depthwise Separable Convolution Network as an efficient feature extraction reduction method for COVID-19 CXR. Furthermore, numerous strategies were used to reduce computational cost while maintaining accuracy, such as customizing the Batch Normalization (BN) layer and replacing the convolution layer with a separable convolution layer. The proposed model successfully reduced the feature extraction represented by the decreases in trainable parameters from 28.640 trainable parameters to 4.640 trainable parameters. The CNN was successfully trains COVID-19 CXR with accuracy 72.96%, loss 12.43%, recall 74.67%, precision 77.67%, and F1-score 75.33%. Meanwhile, the depthwise separable convolution effectively retains the performance accuracy 71.07%, loss 11.55%, recall 70.00%, precision 77.67%, and F1-score 70.00%. The CXR augmentation is also successfully increase the performance accuracy 74.55%, loss 11.37%, recall 77.67%, precision 79.56%, and F1-score 78.33%.

ACKNOWLEDGMENT

The author expresses gratitude to the Electrical Engineering Department and Research Center of Universitas Mercu Buana, Jakarta, Indonesia. Under the support, the authors have been completed this research.

REFERENCES

[1] J. Zhao, X. He, X. Yang, Y. Zhang, S. Zhang, and P. Xie, "COVID-CT-Dataset: A CT image dataset about COVID-19," *arXiv*, pp. 1–14, 2020.
[2] S. Hassantabar, M. Ahmadi, and A. Sharifi, "Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches," *Chaos, Solitons and Fractals*, vol. 140, p. 110170, 2020, doi: 10.1016/j.chaos.2020.110170.
[3] L. Wang, Z. Q. Lin, and A. Wong, "COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images," *Sci. Rep.*, vol. 10, no. 1, pp. 1–12, 2020, doi: 10.1038/s41598-020-76550-z.
[4] N. K. Chowdhury, M. M. Rahman, and M. A. Kabir, "PDCOVIDNeT: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images," *arXiv*, 2020, doi: 10.1007/s13755-020-00119-3.
[5] Y. Zhong, "Using Deep Convolutional Neural Networks to Diagnose COVID-19 from Chest X-Ray Images," *arXiv*, 2020.
[6] E. E. D. Hemdan, M. A. Shouman, and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," *arXiv*, 2020.
[7] M. A. Zulkifley, S. R. Abdani, and N. H. Zulkifley, "COVID-19 screening using a lightweight convolutional neural network with generative adversarial network data augmentation," *Symmetry (Basel).*, vol. 12, no. 9, pp. 1–17, 2020, doi: 10.3390/SYM12091530.
[8] M. Rahimzadeh and A. Attar, "a New Modified Deep Convolutional Neural Network for Detecting Covid-19 From X-Ray Images," *arXiv*, vol. 19, p. 100360, 2020, doi: 10.1016/j.imu.2020.100360.
[9] J. C. Y. Yujin Oh, Sangjoon Park, "Deep Learning Covid-19 Features on CXR using Limited Training Data Sets." IEEE Transaction on Medical Imaging, Vol 39, No 8, pp. 2688–2700, 2020.
[10] S. Widodo *et al.*, "Detection of Covid-19 on X-Ray Images Using a Deep Learning Convolution Neural Network," pp. 255–259, 2021.
[11] Z. Iklima, T. M. Kadarina, and E. Ihsanto, "Realistic Image Synthesis of Covid-19 Chest X-Ray Imaging using Depthwise Boundary Equilibrium Generative Adversarial Networks," vol. 9, no. 3, pp. 1–11, 2021, doi: 10.11591/eei.v9i3.xxxx.
[12] "Chest X-ray (Covid-19 & Pneumonia) | Kaggle." https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia?resource=download (accessed May 12, 2022).
[13] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 39, no. 7, pp. 1379–1389, 2020, doi: 10.1007/s10096-020-03901-z.
[14] B. Liu, D. Zou, L. Feng, S. Feng, P. Fu, and J. Li, "An FPGA-based CNN accelerator integrating depthwise separable convolution," *Electron.*, vol. 8, no. 3, 2019, doi: 10.3390/electronics8030281.
[15] A. H. Panahi, A. Rafiei, and A. Rezaee, "FCOD: Fast COVID-19 Detector based on deep learning techniques," *Informatics Med. Unlocked*, vol. 22, p. 100506, 2021, doi: 10.1016/j.imu.2020.100506.
[16] L. Dang, P. Pang, and J. Lee, "Depth-wise separable convolution neural network with residual connection for hyperspectral image

classification," *Remote Sens.*, vol. 12, no. 20, pp. 1–20, 2020, doi: 10.3390/rs12203408.
[17] D. Hossain, M. H. Imtiaz, T. Ghosh, V. Bhaskar, and E. Sazonov, "Real-Time Food Intake Monitoring Using Wearable Egocnetric Camera," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2020-July, no. March 2021, pp. 4191–4195, 2020, doi: 10.1109/EMBC44109.2020.9175497.
[18] D. S. Kermany *et al.*, "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning," *Cell*, vol. 172, no. 5, pp. 1122-1131.e9, Feb. 2018, doi: 10.1016/J.CELL.2018.02.010.
[19] P. Saha and S. Neogy, "Concat_CNN: A Model to Detect COVID-19 from Chest X-ray Images with Deep Learning," vol. 3, p. 305, 2022, doi: 10.1007/s42979-022-01182-1.
[20] S. Sukegawa *et al.*, "Deep neural networks for dental implant system classification," *Biomolecules*, vol. 10, no. 7, pp. 1–13, 2020, doi: 10.3390/biom10070984.

ZENDI IKLIMA received the M.Sc. degree in Software Engineering from Beijing Institute of Technology in 2018. His research interests include robotics, deep learning, and cloud computing. He has an experience as the Chief Technology Officer of Diaspora Connect Indonesia. Indonesian Diaspora Connect is one of the controversial start-up companies in 2018. The idea of connecting Indonesia is to help the Indonesian government find out about citizens who are studying or living abroad which integrates with Artificial Intelligence. Based on the idea, he achieved 1st place in ALE Hackathon Competition in 2018 which held by Alcatel-Lucent (Jakarta). He can be contacted at email: zendi.iklima@mercubuana.ac.id.

TRIE MAYA KADARINA received her bachelor's degree in Electrical Engineering from Institut Teknologi Nasional in 2001. She received her master's degree in Biomedical Engineering from the Department of Electrical Engineering, Institut Teknologi Bandung, in 2005. Currently, she is a lecturer in the Department of Electrical Engineering at Universitas Mercu Buana. Her research interests include biomedical instrumentation, electronics and control systems, machine learning, and Internet of Things. She can be contacted at email: trie.maya@mercubuana.ac.id.

RINTO PRIAMBODO received his bachelor degree in electrical engineering from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 2004 and his master degree in information technology from Universitas Indonesia in 2011. He is currently pursuing the doctoral degree in the faculty of computer science in Universitas Indonesia. His research interest includes information system, business intelligence and Internet of Things in healthcare. He can be contacted at email: rinto.priambodo@mercubuana.ac.