Double Face Masks Detection Using Region-Based Convolutional Neural Network

Sa'aadah Sajjana Carita, Raden Budiarto Hadiprakoso
Cryptographic engineering, Polteknik Siber dan Sandi Negara, Jl. Haji Usa, Bogor, 16120, Indonesia

ARTICLE INFO

Article history:

Received April 19, 2023
Revised July 25, 2023
Accepted October 03, 2023

Keywords:

R-CNN;
Deep learning;
Transfer learning;
Face mask detection

ABSTRACT

Because of the fast spread of coronavirus, the globe is facing a significant health disaster of COVID-19. The World Health Organization (WHO) released many suggestions to combat the spread of coronavirus. Wearing a face mask in public places and congested locations is one of the most effective preventive practices against COVID-19. However, according to recent research wearing double face masker even provide better protection than just one mask. Based on this finding, various public places require double masks to proceed more. It is pretty tricky to monitor individuals in crowded public places personally. Therefore, a deep learning model is suggested in this paper to automate recognizing persons who are not wearing double face masks. A faster region-based convolutional neural network model is developed using the picture augmentation approach and deep transfer learning to increase overall performance. We apply deep transfer learning by fine-tuning the low level pre-trained Visual Geometry Group (VGG) Face2 model. This study used the publicly accessible VGGFace2 dataset and the self-processed dataset. The findings in this study show that deep transfer learning and image augmentation can increase detection accuracy by up to 11%. Consequently, the created model achieves 93.48% accuracy and 93.19% F1 score on the validation dataset, demonstrating its excellent performance. The test results show the proposed model for further research by adding the predicted dataset and class.

Corresponding Author:

Sa'aadah Sajjana Carita, Polteknik Siber dan Sandi Negara, Jl. Haji Usa, Bogor, 16120, Indonesia
Email: sajjana.carita@bssn.go.id

1. INTRODUCTION

The COVID-19 pandemic is currently causing major health issues worldwide. Originating from China, at the end of 2019, Covid-19 is spreading very fast [1]. At the beginning of 2021, the COVID-19 tsunami hit almost worldwide. Covid-19 positives increased to around 3 million per day [2]. Covid-19 has become a sustained pandemic due to the coronavirus's mutation into multiple variants [3]. Between early and mid-2021, the delta variant of the coronavirus was the most contagious, causing corona pandemic in many countries, including India, Brazil, and Indonesia [4], [5]. In late 2021, the Omicron variant became the primary cause of the Corona pandemic almost worldwide [6]. According to research [7], the omicron variant of the coronavirus is the fastest-transmitting variant. Conversely, vaccine efficacy against this variant is significantly reduced. Some people who received the third booster vaccine dose are still infected by the omicron coronavirus [8].

The World Health Organization (WHO) has issued several guidelines for self-protection against coronavirus [9]. Wearing a mask in crowded places, especially when physical distance is difficult to maintain, is one COVID-19 prevention measure. Almost all countries have regulations requiring people to wear masks in public places, while others recommend it. WHO also advised fully vaccinated people to continue wearing masks and following health protocols [10]. The reason is that even those vaccinated are still susceptible to Covid-19, especially the new variant.

According to a recent Centers for Disease Control and Prevention (CDC) study, a double mask is recommended for better protection [11]. The second mask will effectively double the material layer that the virus's respiratory droplets must pass before reaching the face and mouth. In their study results, if someone only uses one surgical mask, it will only filter about 60%-80% of air particles. In comparison, a single cloth mask can provide protection around 50-70%. When a person uses a medical mask coated with a cloth mask, both will provide up to 96% protection [12].

Based on these findings, the Indonesian government recently advocated the usage of double masks in public settings through the ministry of health [13]. To limit the potential of Covid-19 transmission, many public facilities, including airports, hospitals, and local train stations, mandate travelers to wear double masks [14]. Nevertheless, controlling massive groups of people is becoming more complex. The monitoring process includes disclosing anyone who is not wearing a face mask. Therefore, face mask detection has become a crucial computer vision task [15].

This issue has attracted the attention of several researchers for efforts to build a detection model of whether a person is wearing a face mask or not. Several studies have built various models of face mask detection through machine learning [16]–[18] and deep learning [19]–[25] approaches. Some even achieve excellent results with up to 95% accuracy [15], [18]. However, according to the CDC's recommendations, efforts to detect whether someone is wearing a double face mask have not been addressed yet. Double face mask detection has become a significant problem in image processing and computer vision that requires extra work compared to existing single face mask detection.

This paper proposes a deep learning-based double face mask detection model. The proposed model can detect people wearing single or double face masks, or none at all, using real-time surveillance cameras. We used image augmentation and deep transfer learning to improve overall model accuracy. This study's innovation uses the faster Region-based CNN (Faster R-CNN) architecture to implement semantic segmentation. Unlike traditional CNN structures, Faster R-CNN can handle more complex object detection and image segmentation tasks. Faster R-CNN is designed to perform faster than conventional R-CNNListed below are our study's significant contributions. (1) The model developed is highly reliable because it uses deep transfer learning with image augmentation to analyze the signs of wearing a double face mask. (2) Surveillance cameras can use our method for real-time detection. (3) Our face mask detection system works well with the research results on double masks to prevent Covid-19.

The rest of this work is organized as follows. Section 2 reviews prior related research. Section 3 elaborates on the dataset and research method. Section 4 describes the results of the experiments, and Section 5 concludes and suggests future work.

2. METHODS

This study proposes a method to detect double face masks using a deep learning model based on the region-based CNN architecture [26]. The schematic of how our model works is quite simple, and our proposed methodology includes several general steps consisting of (1) data collection, (2) data pre-processing, (3) model engineering, (4) model training and evaluation. The plot of the research carried out is illustrated in Fig. 1.

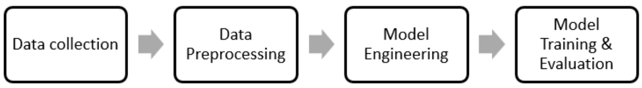

Fig. 1. Research stages

2.1. Data Collection

The datasets were collected in three categories: face datasets with masks, face datasets without masks, and double face masks. We used publicly available datasets from previous studies [27]. This dataset contains 5,850 images of 3,000 mask-wearing faces and 2,850 non-mask-wearing faces. The images in this dataset vary in resolution but are all JPG files.

On the other hand, the collection of facial datasets using double masks is done semi-manually. Ten participants were asked to wear double masks. Five men and five women aged 19 to 44 were asked to collect data voluntarily. Video of respondents wearing double masks was recorded for 5 minutes via webcam. The video will automatically generate images taken every 1 second to produce 3,000 datasets of face images with face masks. The data collection process was carried out during November 2021. An overview of the number of datasets used in each class is described in Table 1.

Table 1. The Dataset

Dataset	Amount
Double masks	3,000
Single masks	3,000
Without masks	2,850
Total	8,850

2.2. Data Preprocessing

Following the collection of datasets, we pre-process the data. The data is pre-processed in a variety of ways during this stage:

1. Eliminate image noise.
2. Adjust the image size.
3. Image augmentation
4. Split the dataset.

First, we resize the image to ensure that all inputs are uniform in size. The image has been reduced to 100×100 pixels. We reduce the image size to a manageable amount while retaining the information contained within. This process is required to ensure that the model runs quickly when later trained on the architectural model. The following reduces image noise by employing a gaussian blur filter. This filter will increase the model's accuracy, allowing it to forecast the image more accurately.

Image augmentation is a handy strategy for increasing the size of a model's training data without having to hunt for new data. In general, picture augmentation reproduces an image with different changes to increase the data. We enhance images by zooming +/- 20%, rotating +/-10%, altering brightness +/-20%, and rotating the image -/+10%. This way, the trained model is more representative of real-world settings and can adapt to changing conditions. Finally, we partition the dataset into subsets. 80% of the data is utilized for training the model, and 20% is used to test the model. In this study, we used 5-fold cross-validation. The 5-fold cross-validation separates the dataset into five equal-sized folds that do not overlap during training and testing.

2.3. Data Preprocessing

The architectural model used in this research is the faster R-CNN (Region-Based Convolutional Neural Network), which develops the R-CNN architecture. The R-CNN process itself consists of 3 stages [20]:

1. Look for regions or parts of the image that may object to the proposal region method. The regional proposal algorithm used is selective search.
2. Each region is then used as input for CNN as feature extractors from each region.
3. Each feature generated then becomes an input for SVM, which will produce a class from that region and a linear regressor that will produce a bounding box.

Faster R-CNN uses the Region Proposal Network (RPN) to use a neural network that replaces the role of selective search to propose regions (which parts of an image need further processing). RPN produces several bounding boxes. Each box has two probability scores whether there are objects at that location or not. The resulting regions will be input for an architecture similar to Fast R-CNN.

Faster R-CNN also abandons the use of SVM as classifiers, replacing it with a region of interest (ROI) pooling and fully connected layers. This approach with 1 CNN, ROI pooling layer, and feed-forward network accelerates R-CNN performance, increases R-CNN's capability to be end-to-end differentiable, and simplifies the training process because there is no need to train for SVM and only trains 1 CNN architecture instead of multiple CNNs. Faster R-CNN architecture can be seen in Fig. 2.

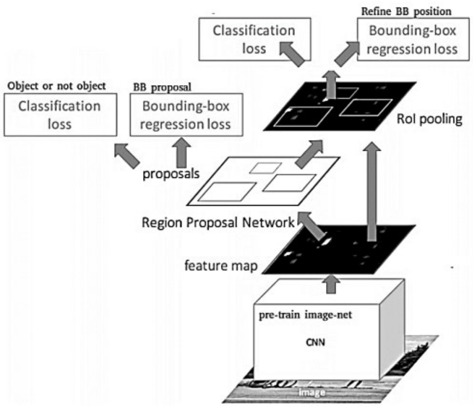

Fig. 2. Faster R-CNN Architecture

Instead of processing many cropped images, R-CNN reduces the number of things to process to some regions. Faster R-CNN uses only 1 CNN. The resulting feature map is then matched with the Region of Interest (ROI) obtained from the RPN results, then its class is classified, and its bounding box is detected. In other words, Faster R-CNN performs feature extraction before submitting regions. The use of RPN to replace selective search reduces computational requirements significantly and makes the entire model trainable end-to-end because no particular method is used as with the previous method proposal region. In addition, Faster R-CNN also produces faster and more accurate performance when compared to R-CNN [28].

Our proposed R-CNN model uses the pre-trained Visual Geometry Group (VGG) Face2 model [29]. The VGG Face2 dataset contains 3.31 million photos of 9131 people, with an average of 362.6 photographs for each subject. Images are downloaded from Google Image Search and have considerable differences in age, gender, ethnicity, and profession, including actors, athletes, politicians. The pre-trained model will detect a person's facial image. After passing the faster R-CNN model, it will produce a bounding box output on the face image. Next, a CNN architecture-based model will be created to determine whether someone is wearing a double mask or not. The CNN model that we propose consists of 2 convolution layers with 64 filters, two max pool layers with a size of 4×4 px, and 2 MLP (Multi-Layer perceptron) layers, as shown in Fig. 3.

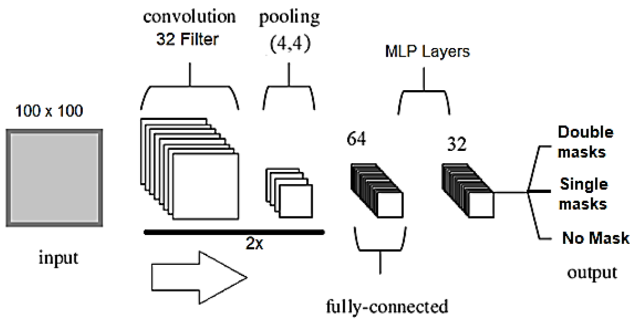

Fig. 3. The Proposed CNN Model

In Fig. 3, the model receives an input image with a size of 100×00 px, which is the output of the pre-processing stage of the data. This input is passed to the CNN layer, which has 64 filters—then forwarded to the max pool layer with a pool size of 4×4. We duplicate this CNN and max pool layer to repeat the process twice. After that, the data will be processed at the MLP layer. There are 2 MLP layers in the first layer having 64 neurons, while the second layer has 32 neurons. All these neurons are fully connected. Finally, the model is activated with the softmax activation function. The softmax activation function will produce output with labels 1, 2, or 3, representing class labels wearing double masks, single masks, and not wearing masks.

The softmax function accepts input i vector z of K real numbers and normalizes it to a probability distribution of K probabilities proportional to the input numbers' exponentials. Before applying softmax, some vector components may be negative or greater than one and may not sum to one; however, after applying softmax, each component will be in the interval [0, 1] display style [0,1] [0,1], and the components will sum to one, allowing them to be interpreted as probabilities. Additionally, the greater the input components, the greater the probability. Equation (1) represents the conventional softmax function.

$$\theta(z)i = \sum_{z}^{i} k1, k2, \ldots \; for \; i = 1, \ldots k \; and \; z(z_1, z_2, \ldots.) \in R\, i^j \tag{1}$$

2.4. Model Training and Evaluation

After the model is created, it will be trained with the pre-processed dataset. The model training process uses an initial parameter of 200 epochs, Adam optimizer, and a learning rate of 0.01. These parameters will later be re-evaluated to find the model parameters with the best results. The model was trained using the TensorFlow 2.7 library on an Nvidia Quadro T1000 graphics card. We use accuracy, precision, recall, and F1 score as metrics for evaluating the model. The assessments include evaluating the results of deep transfer learning and image augmentation. The equations used to calculate the accuracy, precision, recall, and F1 scores are shown in (2) to (5),

$$\text{Accuracy} = \text{TP} + \text{TN}/\text{TP} + \text{FP} + \text{FN} + \text{TN} \tag{2}$$

$$\text{Precision} = \text{TP}/\text{TP} + \text{FP} \tag{3}$$

$$\text{Recall} = \text{TP}/\text{TP} + \text{FN} \tag{4}$$

Vol. 9, No. 4, December 2023, pp. 904-911

$$\text{F1 score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision}) \tag{5}$$

with TP (True- Positive), TN (True-Negative), FP (False-Positive), FN (False-Negative).

3. RESULTS AND DISCUSSION

During the data collection stage, we processed 8,850 image datasets consisting of 3,000 face images with double masks, 3,000 with single masks, and the rest without masks. This entire dataset will be used to train and test the model. Seven thousand eighty data were used to train the model, and the remaining 1,770 were used to validate the model.

Two methodologies are used in the testing strategy. Our first approach validates the CNN region-based model built on the dataset directly. The dataset will be separated into two parts: 80 percent will be used to train the model, and the remainder will be used to test the model. The second testing strategy used the same dataset but included 5-fold cross-validation, picture augmentation, and deep transfer learning techniques stated in the methodological research sub-chapter. This approach is necessary to determine how performance improvement occurs when image augmentation and deep transfer learning are used.

In the model training phase, we quantify the model's resulting loss. True-Positive (TP), False-Positive (FP), True-Negative (TN), and False-Negative (FN) are the other performance metrics. These metrics are represented by the confusion matrix, a grid-like structure. This work creates two confusion matrices to evaluate the model's achievement during training and testing. Confusion matrices for double masks with non-double masks as presented in Fig. 4.

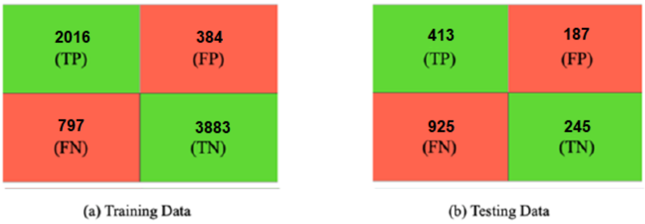

Fig. 4. The Confusion Matrices

The results in Fig. 4 show that the model correctly predicted 2,019 face images with image masks from a total of 2,400 images during the training process. The remaining 384 or 16% of the total images were incorrectly classified as faces with single masks. Of the 4,680 images other than double masks, 797 or about 17% were incorrectly classified as wearing double masks. All misclassified labels show a face with a single mask incorrectly labeled as a double mask or vice versa. These results show that the model is very good at recognizing people who are not wearing masks. However, the model still has an error of around 17-16% when detecting a face mask between single or double masks. We performed the image augmentation technique to improve the overall model performance described in the previous section. Fig. 5 shows an example of an augmented dataset result.

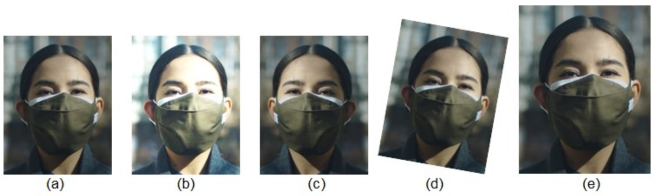

Fig. 5. Image Augmentation

Fig. 5(a), the most left image is the initial image before the image augmentation process. Then it is augmented by adding +20% brightness to produce something like Fig. 5(b). Fig. 5(c) is a horizontal flip of the original image. Fig. 5(d) shows the results of image augmentation in a 10% rotation clockwise manner, while Fig. 5(e) is the result of zooming 120% of the original image. After the pre-processing stage, the image dataset is used to train and test the model. This time we implemented 5-fold cross-validation with image augmentation and deep transfer learning. The model validation results are shown in Table 2, and the Receiver Operating Characteristic (ROC) curve is in Fig. 6.

Table 2. Cross-Validation Result

	Fold-1	Fold-2	Fold-3	Fold-4	Fold-5	Avg.
Accuracy	0.9368	0.9321	0.9397	0.9260	0.9397	0.9348
Precision	0.9409	0.9407	0.9240	0.9231	0.9345	0.9324
Recall	0.9365	0.9365	0.9292	0.9324	0.9220	0.9313
F1 Score	0.9387	0.9386	0.9266	0.9277	0.9282	0.9319

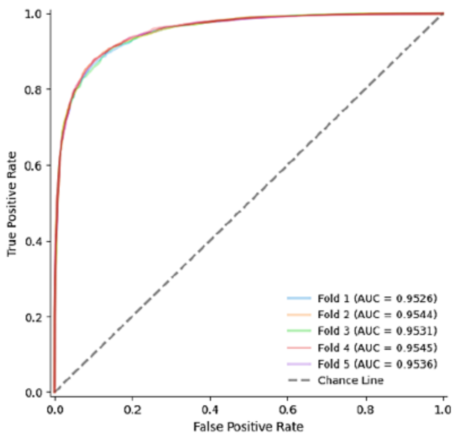

Fig. 6. ROC Curve

According to Table 2, the average value of accuracy, precision, recall, and F1 scores is between 92 and 94%. Between each fold, the results are relatively steady. Additionally, the accuracy value and the F1 score are relatively consistent, remaining within a narrow range. The results demonstrate that the proposed model is compatible with the test dataset. The ROC curve is used to assess the performance of classification issues. The Area Under Curve (AUC) measures separability, whereas the ROC is a probability curve. This value indicates a model’s ability to separate classes. Fig. 6 depicts a pretty broad AUC of roughly 95%. The value shows that the generated model is suitable for the input categorization.

When image augmentation and deep transfer learning were implemented, a significant increase in model performance of approximately 9-11 percent was observed. This result is based on the general performance of deep learning models when trained on massive datasets. This result can be accomplished using image augmentation, which adds diversity to the image collection to imitate real-world settings.

Transfer learning for computer vision is based on the premise that models trained on vast datasets of available images can be utilized as foundational models for recognizing properties of objects in the actual world. We can use these learned features without retraining the model from scratch. Deep transfer learning is the process of applying a model that has been pre-trained on another dataset to our own. This fact makes it possible to improve accuracy results significantly.

4. CONCLUSION

Numerous research studies have established that wearing double face masks in public reduces the virus’s transmission rate. As a result, governments and the private sector have made double masks essential in public and populated places. It is challenging to keep an eye on crowds in these locations. This paper presents a deep learning model for detecting individuals who are not covering double face masks. The deep learning model is constructed utilizing a faster R-CNN. In this work, picture augmentation techniques proven boost the model’s performance by diversifying the training data. In addition, deep transfer learning improves the overall model accuracy by 11-12%. The suggested deep learning model achieved 94,99% accuracy and 94.83% F1 score on the validation dataset.

The limitation of this study is that the model built can only distinguish the use of double, single, and non-mask masks. There are reports that double masks are less effective on the KF94 and KN95/N95 types of masks [30]. Additionally, the datasets for the double mask sample are still scarce. The limitations of this dataset are likely to influence real-world settings with widely varying inputs. Future work could be improved further by entering a large amount of data and expanding it to classify these types of face masks. The Multi-view Convolutional Neural Networks architecture [31] can be investigated further for optimal detection accuracy.

REFERENCES

[1] M. Ciotti, M. Ciccozzi, A. Terrinoni, W.-C. Jiang, C.-B. Wang, and S. Bernardini, “The COVID-19 pandemic,” *Crit. Rev. Clin. Lab. Sci.*, vol. 57, no. 6, pp. 365–388, 2020, https://doi.org/10.1080/10408363.2020.1783198.

[2] Worldometer, "COVID Live - Coronavirus Statistics." https://www.worldometers.info/coronavirus/ (accessed Jul. 21, 2023).
[3] T. Harder *et al.*, "Effectiveness of COVID-19 vaccines against SARS-CoV-2 infection with the Delta (B.1.617.2) variant: second interim results of a living systematic review and meta-analysis," *Eurosurveillance*, vol. 26, no. 41, pp. 1–10, 2021, https://doi.org/10.2807/1560-7917.ES.2021.26.41.2100920.
[4] S. Shiehzadegan, N. Alaghemand, M. Fox, and V. Venketaraman, "Analysis of the Delta Variant B.1.617.2 COVID-19," *Clin. Pract.*, vol. 11, no. 4, pp. 778–784, 2021, https://doi.org/10.3390/clinpract11040093.
[5] T. K. Burki, "Lifting of COVID-19 restrictions in the UK and the Delta variant," *Lancet Respir. Med.*, vol. 9, no. 8, p. e85, 2021, https://doi.org/10.1016/S2213-2600(21)00328-3.
[6] S. S. A. Karim and Q. A. Karim, "Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic," *Lancet*, vol. 398, no. 10317, pp. 2126–2128, 2021, https://doi.org/10.1016/S0140-6736(21)02758-6.
[7] E. Mahase, "Covid-19: Do vaccines work against omicron—and other questions answered," *BMJ*, vol. 375, p. n3062, 2021, https://doi.org/10.1136/bmj.n3062.
[8] I. Torjesen, "Covid-19: Omicron may be more transmissible than other variants and partly resistant to existing vaccines, scientists fear," *BMJ*, vol. 375, p. n2943, 2021, https://doi.org/10.1136/bmj.n2943.
[9] M. Abboah-Offei, Y. Salifu, B. Adewale, J. Bayuo, R. Ofosu-Poku, and E. B. A. Opare-Lokko, "A rapid review of the use of face mask in preventing the spread of COVID-19," *Int. J. Nurs. Stud. Adv.*, vol. 3, p. 100013, 2020, https://doi.org/10.1016/j.ijnsa.2020.100013.
[10] M. Liao *et al.*, "A technical review of face mask wearing in preventing respiratory COVID-19 transmission," *Curr. Opin. Colloid Interface Sci.*, vol. 52, p. 101417, 2021, https://doi.org/10.1016/j.cocis.2021.101417.
[11] J. T. Brooks *et al.*, "Maximizing Fit for Cloth and Medical Procedure Masks to Improve Performance and Reduce SARS-CoV-2 Transmission and Exposure," *MMWR. Morb. Mortal. Wkly. Rep.*, vol. 70, no. 7, pp. 254–257, 2021, https://doi.org/10.15585/mmwr.mm7007e1.
[12] M. A. Honein *et al.*, "Summary of Guidance for Public Health Strategies to Address High Levels of Community Transmission of SARS-CoV-2 and Related Deaths," *MMWR. Morb. Mortal. Wkly. Rep.*, vol. 69, no. 49, pp. 1860–1867, 2020, https://doi.org/10.15585/mmwr.mm6949e2.
[13] R. Rachmaniar and S. Susanti, "Diskursus Video Masker Ganda untuk Pencegahan COVID-19 di Channel Youtube CNN Indonesia," *PARAHITA J. Pengabdi. Kpd. Masy.*, vol. 2, no. 1, pp. 8, 2021, https://doi.org/10.25008/parahita.v2i1.56.
[14] C. C. Lai, M. Y. Yen, P. I. Lee and P. R. Hsueh, "How to keep COVID-19 at bay: a Taiwanese perspective," *Journal of Epidemiology and Global Health,* vol. 11, no. 1, p. 1, 2021, https://doi.org/10.2991%2Fjegh.k.201028.001.
[15] A. Oumina, N. El Makhfi, and M. Hamdi, "Control The COVID-19 Pandemic: Face Mask Detection Using Transfer Learning," in *IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pp. 1–5, 2020, https://doi.org/10.1109/ICECOCS50124.2020.9314511.
[16] Y. Chen *et al.*, "Face Mask Assistant: Detection of Face Mask Service Stage Based on Mobile Phone," *IEEE Sens. J.*, vol. 21, no. 9, pp. 11084–11093, 2021, https://doi.org/10.1109/JSEN.2021.3061178.
[17] R. Vaishya, M. Javaid, I. H. Khan, and A. Haleem, "Artificial Intelligence (AI) applications for COVID-19 pandemic," *Diabetes Metab. Syndr. Clin. Res. Rev.*, vol. 14, no. 4, pp. 337–339, 2020, https://doi.org/10.1016/j.dsx.2020.04.012.
[18] S. A. Sanjaya and S. Adi Rakhmawan, "Face Mask Detection Using MobileNetV2 in The Era of COVID-19 Pandemic," in *International Conference on Data Analytics for Business and Industry: Way Towards a Sustainable Economy (ICDABI)*, pp. 1–5, 2020, https://doi.org/10.1109/ICDABI51230.2020.9325631.
[19] X. Fan and M. Jiang, "RetinaFaceMask: A Single Stage Face Mask Detector for Assisting Control of the COVID-19 Pandemic," in *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 832–837, 2021, https://doi.org/10.1109/SMC52423.2021.9659271.
[20] A. Chavda, J. Dsouza, S. Badgujar, and A. Damani, "Multi-Stage CNN Architecture for Face Mask Detection," in *6th International Conference for Convergence in Technology (I2CT)*, pp. 1–8, 2021, https://doi.org/10.1109/I2CT51068.2021.9418207.
[21] J. Zhang, F. Han, Y. Chun, and W. Chen, "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19," *IEEE Access*, vol. 9, pp. 42975–42984, 2021, https://doi.org/10.1109/ACCESS.2021.3066538.
[22] G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, "Face Mask Detection Using Transfer Learning of InceptionV3," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12581, pp. 81–90. 2020, https://doi.org/10.1007/978-3-030-66665-1_6.
[23] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, and J. Hemanth, "SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2," *Sustain. Cities Soc.*, vol. 66, p. 102692, 2021, https://doi.org/10.1016/j.scs.2020.102692.
[24] M. Loey, G. Manogaran, M. H. N. Taha, and N. E. M. Khalifa, "Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection," *Sustain. Cities Soc.*, vol. 65, p. 102600, 2021, https://doi.org/10.1016/j.scs.2020.102600.
[25] J. Yu and W. Zhang, "Face Mask Wearing Detection Algorithm Based on Improved YOLO-v4," *Sensors*, vol. 21, no. 9, p. 3263, 2021, https://doi.org/10.3390/s21093263.
[26] T. E. Liang, U. U. Sheikh, and M. Norzali Haji Mohd, "Malaysian car plate localization using region-based convolutional neural network," *Bull. Electr. Eng. Informatics*, vol. 9, no. 1, pp. 411–419, 2020,

https://doi.org/10.11591/eei.v9i1.1862.

[27] K. Suresh, M. Palangappa, and S. Bhuvan, "Face Mask Detection by using Optimistic Convolutional Neural Network," in *6th International Conference on Inventive Computation Technologies (ICICT)*, pp. 1084–1089, 2021, https://doi.org/10.1109/ICICT50816.2021.9358653.

[28] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep Face Recognition," in *Procedings of the British Machine Vision Conference*, pp. 41.1-41.12, 2015, https://doi.org/10.5244/C.29.41.

[29] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A Dataset for Recognising Faces across Pose and Age," in *13th IEEE International Conference on Automatic Face & Gesture Recognition,* pp. 67–74, 2018, https://doi.org/10.1109/FG.2018.00020.

[30] P. W. Clapp *et al.*, "Evaluation of Cloth Masks and Modified Procedure Masks as Personal Protective Equipment for the Public During the COVID-19 Pandemic," *JAMA Intern. Med.*, vol. 181, no. 4, p. 463, 2021, https://doi.org/10.1001/jamainternmed.2020.8168.

[31] H. Su, S. Maji, E. Kalogerakis, and E. Learned-Miller, "Multi-view Convolutional Neural Networks for 3D Shape Recognition," in *IEEE International Conference on Computer Vision (ICCV)*, pp. 945–953, 2015, https://doi.org/10.1109/ICCV.2015.114.

BIOGRAPHY OF AUTHORS

Sa'aadah Sajjana Carita is a lecturer in Politeknik Siber dan Sandi Negara, specializing in Mathematics and Cryptography. She is also a member of Indonesian Algebraic Society. Her email address is sajjana.carita@bssn.go.id.

Raden Budiarto Hadiprakso is a lecturer in Politeknik Siber dan Sandi Negara, specializing in Cryptography and Software Engineering. She is also a member of IEEE. Her email address is budiarto.hadiprakoso@bssn.go.id.