

Real-time mask-wearing detection in video streams using deep convolutional neural networks for face recognition

Suhirman¹, Shoffan Saifullah^{2,3}, Ahmad Tri Hidayat¹, M. Apriandi Kusuma¹, Rafał Dreżewski^{2,4}

¹Department of Informatics, Faculty of Science and Technology, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia

²Faculty of Computer Science, AGH University of Krakow, Krakow, Poland

³Department of Informatics, Universitas Pembangunan Nasional Veteran Yogyakarta, Yogyakarta, Indonesia

⁴Knowledge Engineering and Data Science Center, Department of Electrical Engineering and Informatics, Universitas Negeri Malang, Malang, Indonesia

Article Info

Article history:

Received Jun 9, 2023

Revised Jul 20, 2023

Accepted Sep 12, 2023

Keywords:

Coronavirus disease 2019

Convolutional neural networks

Face recognition

Real-time system

Mask detection

ABSTRACT

This research aims to develop a real-time mask-wearing detection system using deep convolutional neural networks (CNNs). This is crucial in the coronavirus disease 2019 (COVID-19) pandemic to alert individuals who are not wearing masks early on, thereby reducing the spread of the virus. Since COVID-19 primarily spreads through respiratory droplets and mask-wearing is recommended, our proposed study utilizes computer vision techniques, specifically image processing, to detect masked and unmasked faces. We employ a customized CNN architecture consisting of five convolutional layers, followed by max-pooling layers and fully connected (FC) layers. The final output layer utilizes softmax activation for classification. The model is updated with optimized layer configurations and parameter values. We are developing an application that uses a digital camera as an input device. The application utilizes a dataset comprising 11,792 image samples, which are used for training and testing purposes with the 80:20 ratio. Real-time testing is conducted using human subjects captured by the camera. The experimental results demonstrate that the CNN method achieves a classification accuracy of 99% on the training data and 98.83% during real-time video testing. These findings suggest that the real-time mask detection system using CNN performs effectively.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Suhirman

Department of Informatics, Faculty of Science and Technology, Universitas Teknologi Yogyakarta

Siliwangi (Ringroad Utara) St., Jombor, Sleman, D.I. Yogyakarta 55285, Indonesia

Email: suhirman@uty.ac.id

1. INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic caused by the novel coronavirus has become a global health crisis, necessitating various preventive measures to curb its spread [1]. Among these measures, health authorities have widely recommended the use of masks to reduce transmission through respiratory droplets effectively [2]. Ensuring compliance with mask-wearing guidelines has become crucial to managing the pandemic [3]. Manual monitoring and enforcing mask-wearing can be challenging and resource-intensive, especially in public spaces and high-traffic areas [4]. To address these challenges, researchers have turned to computer vision techniques and deep learning algorithms, such as deep convolutional neural networks (CNNs), to develop automated systems for real-time mask-wearing detection [5]–[8] in video streams.

This study introduces an effective and efficient system for real-time mask-wearing detection using

deep convolutional neural networks (CNNs) for facial recognition. By leveraging computer vision techniques and image processing, the system accurately identifies individuals wearing masks or not in video streams [9]. Deep CNNs are employed to extract meaningful facial features [10], ensuring precise classification [11]. The system utilizes a custom-built application with a digital camera as an input source for capturing real-time video streams. Extensive training and testing on a comprehensive dataset are conducted, evaluating the system's performance with an 80:20 train-test split [12]. The trained CNN model processes real-time video frames [14], providing timely and accurate mask detection. Instantaneous alerts or notifications are generated for individuals without masks, facilitating proactive intervention to enforce mask-wearing protocols and reduce the risk of virus transmission. This research contributes to computer vision and deep learning by applying deep CNNs for real-time mask-wearing detection [13], offering a valuable tool for enhancing public health and safety during the COVID-19 pandemic.

This study presents a system for real-time mask-wearing detection using CNNs. The paper is organized into five sections. In the introduction, we highlight the importance of mask-wearing and the challenges of manual monitoring. The related works section reviews previous studies, while the method section explains the approach used. Results and discussion discuss training, testing, and real-time scenarios. The conclusion summarizes the findings and future improvements. Overall, our research contributes to computer vision and offers an effective solution for real-time mask detection.

2. RELATED WORKS

Face recognition accuracy has greatly improved in controlled environments but faces some challenges in unconstrained settings due to variations in poses, lighting, expressions, and occlusions [15], [16]. Previous approaches addressed specific issues separately, resulting in limited accuracy. However, introducing deep learning, particularly deep convolutional neural networks (DCNNs), has revolutionized the field by extracting abstract features and effectively handling diverse facial variations, surpassing traditional methods [17].

Several studies have been conducted on the detection of mask usage using different methods. Christa *et al.* [18] developed a system that used the MobileNetV2 architecture for driver mask detection. The training results showed accurate classification of 13 out of 13 photos of drivers without masks, but the system failed to correctly classify 1 out of 13 photos of drivers wearing masks. The study achieved high accuracy, but it was limited by the small size of the dataset. YOLOv3-tiny algorithm in a convolutional neural network (CNN) [19] is used for mask identification and the testing results indicated a mean average precision (mAP) of approximately 98.73% on custom dataset, outperforming YOLOv3-tiny by approximately 62%. One advantage of this study was using an augmented dataset, but it suffered from relatively low accuracy and required a long training time due to many iterations.

Furthermore, the development of CNNs using the Viola-Jones method [20], pre-trained CNNs [21], and Haar Cascade [22] has made significant contributions to facial mask detection, achieving accuracy above 90%. However, some of these methods exhibit drawbacks such as lengthy processing times, excessively large numbers of epochs, implementation on static images rather than real-time video, and limitations in terms of the distance factor, which is restricted to only 40 cm, along with the requirement for high light intensity. The Viola-Jones method, combined with CNNs [20], has proven effective in detecting facial masks with an accuracy above 90%. However, it has been observed that this approach often entails a time-consuming process. Additionally, some studies have employed many epochs during the training phase, leading to prolonged training times. Although these methods have achieved high accuracy rates, they have primarily focused on analyzing static images rather than real-time video inputs, limiting their practical applicability.

Moreover, the Haar Cascade method [22], when integrated with CNNs [23], has also demonstrated successful mask detection outcomes with an accuracy above 90%. However, its implementation is subject to certain restrictions. Specifically, the detection performance is optimized when the distance between the camera and the subject is limited to 40 cm. Additionally, these methods require high light intensity to ensure accurate mask detection.

Goyal *et al.* [24] introduce a face mask detection model for static images and real-time videos, achieving 98% accuracy on the Kaggle dataset. It demonstrates improved computational efficiency and precision compared to other models (DenseNet-121, MobileNet-V2, VGG-19, and Inception-V3). The model has practical applications in public and commercial settings, ensuring effective face mask compliance in schools, hospitals, banks, and airports during the COVID-19 pandemic. However, YOLO-v4 improves feature extraction [25],

reduces computation, and incorporates semantic information for face mask recognition. It achieves a high accuracy of 98.3% and a fast frame rate of 54.57 FPS, surpassing existing algorithms.

These studies have contributed to the field of mask detection, highlighting various methods and their respective strengths and limitations. While some achieved high accuracy, there is still room for improvement regarding dataset size, training time, and the use of video or real-time inputs for more practical applications. Further research can focus on addressing these limitations to enhance the effectiveness and applicability of mask detection systems.

3. METHOD

The primary objective of this study is to develop an efficient and accurate system for detecting whether an individual is wearing a face mask. The model under development leverages a sophisticated learning process based on carefully curated datasets obtained from Jangra [30]. These datasets serve as valuable resources for training and testing the model's performance in real-time using video streams to identify individuals wearing face masks. The detection process employs the mighty deep neural network (DNN) concept, explicitly utilizing the CNN method known for its exceptional capabilities in visual analysis. To ensure optimal performance, the initial phase of the study involves preprocessing the data to obtain high-quality inputs for training and testing. This preprocessing step is critical in obtaining an optimal model to detect face masks in real-time scenarios.

Furthermore, a validation process is conducted to validate the accuracy and reliability of the application of the CNN method to face mask detection. The study aims to obtain precise and trustworthy results through meticulous testing and validation. The detailed process steps are outlined in Figure 1, visually representing the sequential stages related to data preprocessing, model development, real-time detection, and validation.

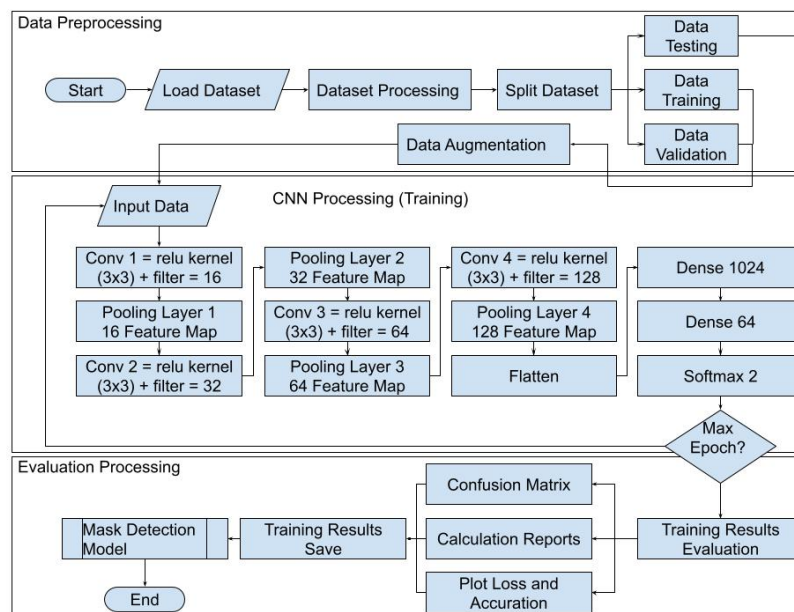


Figure 1. The processing workflow of datasets for CNN modeling and the evaluation process

Based on the CNN modeling in Figure 1, the preprocessed dataset is then processed by a CNN model using the U-net architecture. This allows the developed model to detect face masks by applying a face recognition-based model evaluation method. The U-net architecture is a popular choice for image segmentation tasks, where the objective is to classify each pixel in an image into different categories. In the context of face mask detection, the U-net model can be used to segment the face region and classify whether a mask is present.

A model evaluation method based on face recognition is employed to evaluate the developed model's performance. This evaluation method assesses the model's ability to accurately recognize and classify faces

with and without masks. Various evaluation metrics such as accuracy, precision, recall, and F1-score can be utilized to measure the model's performance. The developed CNN model can effectively detect face masks by combining the U-net architecture for image segmentation and the face recognition-based evaluation method. This approach leverages the strengths of both techniques to achieve accurate and reliable results in face mask detection.

3.1. Face recognition

Face recognition is a crucial step in using a face mask detection system. The face detection process is the initial and essential stage, since mask identification focuses on the facial area. Masks are worn on the face, making it crucial to detect faces accurately [10]. Several challenges are involved in the face detection process, one of which is the presence of factors that affect the detection of facial objects, such as the face position. Various aspects of facial position need to be considered, including the face orientation, sideways and tilted positions, and factors like facial hair, mustaches, and accessories like eyeglasses [26].

Moreover, facial expressions such as laughter, smiles, speech, sadness, and so on further complicate face detection. Additionally, indoor and outdoor environmental conditions can hinder detection, as crowded places or public areas may obstruct the view of facial objects. Lighting conditions also play a significant role, including the lighting within a room, the direction and intensity of light, and the characteristics of the camera sensor and lens used in the face detection process. Considering these factors, face recognition in the context of mask detection requires robust algorithms and techniques capable of handling variations in facial positions, expressions, environmental conditions, and lighting [27]. By addressing these challenges, a practical face recognition system can accurately detect faces and contribute to successfully detecting face masks.

3.2. Deep convolutional neural networks

CNNs are a specialized type of neural network architecture designed to process and analyze visual data efficiently. They have revolutionized the field of computer vision and are widely used in tasks such as image classification, object detection, and face recognition [28]. The key characteristics of deep CNNs include:

- Convolutional layers: CNNs employ convolutional layers that apply a set of learnable filters (kernels) to the input image. These filters capture local patterns and features, allowing the network to extract relevant visual information automatically.
- Pooling layers: pooling layers are used to downsample the spatial dimensions of the feature maps generated by the convolutional layers. Common pooling operations include max pooling and average pooling, which help reduce the computational complexity and extract the most salient features.
- Non-linear activation functions: activation functions, such as rectified linear unit (ReLU), are applied to introduce non-linearity into the network. Non-linear activation functions enable the model to learn complex relationships and improve its ability to capture intricate patterns in the data.
- Fully connected layers: towards the end of the network, fully connected layers are employed to process the high-level features extracted by the convolutional layers. These layers combine the learned features and make predictions based on them, mapping the features to specific classes or outputs.
- Backpropagation and training: CNNs are trained using backpropagation, where the gradients of the loss function with respect to the network parameters are computed and used to update the weights through optimization algorithms like stochastic gradient descent (SGD) or Adam. This iterative process allows the network to learn the optimal set of weights that minimizes the prediction error.

The advantages of deep CNNs include their ability to automatically learn hierarchical features, robustness to variations in input data, and the ability to generalize well to unseen examples. They have significantly improved the performance of computer vision tasks, achieving state-of-the-art results in various benchmarks and competitions.

3.3. Real-time and the proposed system

Real-time refers to the operational state of a system where actions and events are processed and perceived immediately as they occur, without noticeable delays. In computer systems, real-time processing entails the prompt input, processing, and output of data within predefined time constraints [29]. For instance, when typing a document on a computer, the characters entered via the keyboard are swiftly displayed on the screen within the established tolerance, imperceptible to the human eye. The system is considered non-real-time if the latency surpasses the predetermined threshold, resulting in noticeable delays. Real-time systems find

widespread applicability across various domains and are commonly embedded in specialized devices such as cameras, MP3 players, aircraft, GPS devices, and automobiles.

This research proposes a real-time mask detection system, leveraging the possible real-world implementation scheme depicted by the following steps:

- Deployment of surveillance cameras at designated points,
- Continual monitoring of individuals' activities outside their homes through the camera feeds,
- Automated detection of individuals wearing or not wearing masks, and
- Provision of visual and audible alarms on the monitor if an individual is detected without a mask.

By automating the detection process and enabling real-time alerts, the proposed system aims to enhance the effectiveness and efficiency of mask-wearing enforcement, ultimately reducing the risk of virus transmission and safeguarding the well-being of residents within the complex.

4. RESULT AND DISCUSSION

In this section, we will provide examples of datasets used, discuss the results obtained, and analyze the face mask detection using the CNN method with the U-net architecture. For this study, we utilize a secondary dataset obtained from kaggle.com. This dataset is selected for its suitability in modeling and developing a real-time face detection system. The dataset comprises images annotated with information on whether a person is wearing a face mask.

To evaluate the performance of our approach, we conduct experiments and simulations using Python. The experimental results will be presented and analyzed using various evaluation metrics, including the confusion matrix. The confusion matrix allows us to assess the accuracy and performance of our face mask detection system by comparing the predicted labels with the ground truth labels. Furthermore, we also evaluate the system in a real-time setting by implementing our proposed scenario. This involves deploying the trained CNN model with the U-net architecture in a real-time video stream to detect face masks. The performance and effectiveness of the system will be discussed, taking into account factors such as speed, accuracy, and robustness. Utilizing the selected dataset, experimental simulations, and real-time testing, we aim to demonstrate the effectiveness of the CNN method with the U-net architecture for face mask detection. The results and discussions presented in this section will provide insights into our proposed approach's performance and potential applications.

4.1. Dataset training

In this research, the dataset "Face Mask 12K Images Dataset," available at www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset was used [30]. The dataset comprises facial images depicting individuals both wearing and not wearing masks. These images were collected from various countries worldwide, ensuring a diverse representation of individuals and contexts. The dataset consists of 11,792 images, with 5,883 images depicting individuals wearing masks and 5,909 images depicting individuals without masks in Figure 2. Including many images in each category enables the model to learn and generalize patterns effectively. This dataset is a valuable resource for training and evaluating the face mask detection model, providing a wide range of scenarios and variations to enhance its robustness and accuracy.

4.2. Training results and confusion matrix

The training phase of the research utilized a carefully curated dataset and employed rigorous training procedures, including 20 epochs and a batch size of 32. The choice of batch size is crucial as it affects both the accuracy and memory usage of the training process. The training results are plotted in a graph as shown in Figure 3. The Adam optimization algorithm with a learning rate of 0.0001 was employed to fine-tune the model. The results obtained from the classification report during training demonstrated excellent performance across all classes, with consistently high precision, recall, and F1-score values, as shown in Table 1. This indicates that the model could learn and generalize from the training data.

The confusion matrix obtained from the training process provided insights into the distribution of data samples in each class. The balanced number of samples for each class, with 509 images without masks and 483 images with masks, ensured a fair evaluation of the model's performance. The overall accuracy achieved during training was an impressive 99%, indicating the effectiveness of the deep convolutional neural network architecture in accurately detecting face masks.

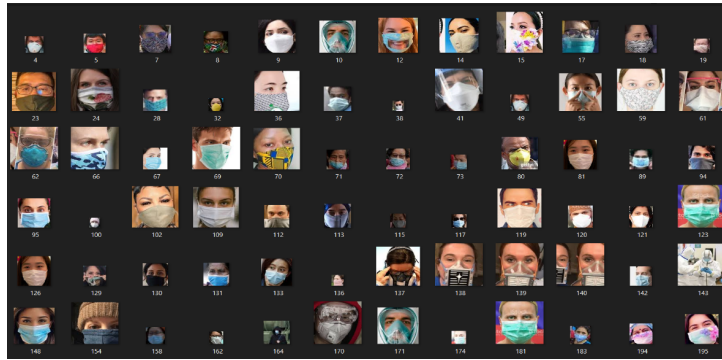


Figure 2. Sample images of faces with masks

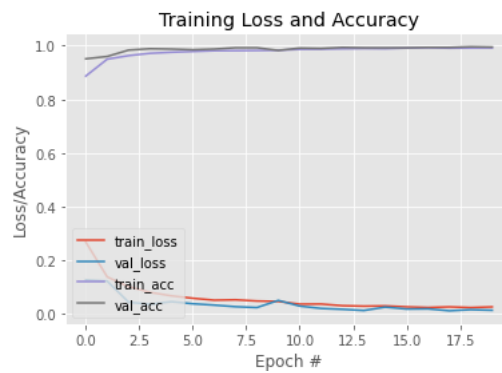


Figure 3. Loss and accuracy results for the training phase

Table 1. The values of metrics for the training phase

Type	Evaluation results (%)				Support
	Accuracy	Precision	Recall	F1 Score	
With mask	99	99	99	99	483
Without mask	99.2	99	99	99	509

4.3. Real-time model testing with single or multiple objects (people)

The system model was evaluated through real-time testing using a directly connected camera that provided live video input. This allowed for the detection process to operate on dynamic images, enabling instantaneous determination of whether an individual was wearing a face mask or not. During the testing phase, various factors were considered as they could influence the accuracy of the proposed model, such as the distance between the object (person) and the camera and the number of objects within a single frame [5].

In real-time testing, the system's performance was analyzed under different scenarios. The system demonstrated its ability to detect face masks within a specific range. However, it exhibited limitations in detecting masks at longer distances. This finding suggests that the proximity between the camera and the detected objects influences the model's performance. Additionally, the impact of lighting conditions on mask detection was observed, with testing conducted in well-lit conditions during daylight. These factors should be considered when deploying the system in real-world settings.

A comprehensive testing scenario was designed to assess the model's performance, consisting of 10 different variations of object-to-camera distances as outlined in Table 2. The dataset used for testing was obtained from Kaggle.com and was utilized for system modeling to develop real-time face detection capabilities. Experimental results were obtained using Python simulations and the confusion matrix technique. Additionally, the system was tested in real-time using a proposed scenario.

Table 2 presents the testing scenario and results for real-time face mask detection for people with and without masks. The factors evaluated include the distance between the object and the camera and the number of

objects within a frame. The system's accuracy in detecting face masks was measured regarding the percentage of correctly classified objects. The average accuracy across all scenarios was calculated to assess the overall model's performance.

Table 2. The results of real-time face mask detection system for different scenarios

Factor	Accuracy (%)				Average accuracy
	Single persons	2-5 people	More than 5 people	Average	
1 meter	100	100	100	100	98.83
1.5 meter	100	100	100	100	
2 meter	100	100	100	100	
2.5 meter	100	100	100	100	
3 meter	100	100	100	100	
3.5 meter	100	100	100	100	
4 meter	100	100	100	100	
4.5 meter	100	100	91.67	97.2	
5 meter	100	100	91.67	97.2	
5.5 meter	95	95	91.67	93.89	

Based on the results presented in Table 2, the testing scenario demonstrated an accuracy of 98.61% in detecting individuals not wearing face masks. The objects were tested at various distances from the camera, ranging from 1 meter to 5.5 meters. However, it was observed that the camera's capturing capability was compromised beyond a distance of 5.5 meters, which limited the system's accuracy. Hence, 5.5 meters was considered the maximum distance for this study. These findings are consistent with previous studies that have reported decreased accuracy as the distance between the object and the camera increases [6].

The accurate results were obtained by conducting experiments on individuals wearing and not wearing face masks. It was found that the accuracy of detecting individuals wearing face masks was generally higher compared to those not wearing masks. Notably, for distances ranging from 4.5 meters to 5.5 meters, the accuracy of detecting individuals not wearing face masks reached optimal levels (100%), while the detection of individuals wearing masks exhibited some errors in identifying people with masks or not. This observation aligns with previous research that has reported challenges in accurately detecting face masks due to variations in mask designs and materials [7].

Furthermore, the detection system maintained optimal performance (100% accuracy) for objects of fewer than 6 individuals at a distance of 5 meters. However, when multiple objects were present at distances of 4.5 meters and above, the real-time accuracy results indicated slightly lower performance (though still above 90%). This can be attributed to the increased complexity of the scene, potential occlusions, and the decreased resolution of objects due to the increased distance from the camera. It is worth mentioning that factors such as the camera's specifications and image acquisition settings had minimal impact on the detection performance, as indicated by the consistent accuracy results across the tested scenarios. These findings corroborate previous studies that have emphasized the robustness of deep convolutional neural networks in handling variations in imaging conditions and camera settings [8].

In conclusion, the experimental results from the real-time testing of the face mask detection system, based on the U-net architecture and deep convolutional neural network, demonstrated high accuracy in detecting individuals wearing face masks. The testing scenarios considered variations in distances and the number of objects, providing insights into the system performance under different conditions. The results confirm the effectiveness of the proposed approach in real-time face mask detection, with implications for applications in public health, security, and safety domains.

5. CONCLUSION

The study's findings indicate that the CNN method is highly effective in the real-time detection of individuals wearing face masks. The deep neural network architecture achieved an impressive accuracy rate of 99% during the training, where extensive training and testing were conducted using appropriate datasets. This signifies the robustness and reliability of the CNN approach in accurately detecting face masks.

The validation of the model through real-time implementation further substantiated its performance, yielding an accuracy rate of 98.83%. This validation process involved testing the model's ability to detect face

masks in live video streams captured directly from a camera connected to a computer. The results obtained from this validation process confirm the model's capability to effectively determine whether an individual is wearing a face mask or not in real-time scenarios. One crucial factor that significantly influenced the accuracy of the proposed model was the distance between the camera and the objects being detected. Through various experiments, it was determined that the optimal performance of the model was achieved when the distance between the camera and the objects was within 5.5 meters. Beyond this distance, the camera's capture capability gradually diminished, which had a noticeable impact on the accuracy of the detection process.

In conclusion, the developed CNN model has demonstrated its ability to detect face masks in real-time settings accurately. To further enhance the model's performance, future work can focus on optimizing object detection in low-light conditions, ensuring that the system would operate optimally under varying lighting conditions. Additionally, the slight video streaming delay caused by limited memory can be addressed by converting the TensorFlow model to TensorFlow.js, which reduces memory requirements and facilitates smoother and faster video streaming. By considering these improvements, the overall accuracy, efficiency, and usability of the CNN-based face mask detection system would be enhanced, making it a valuable tool for real-time applications in various scenarios.

ACKNOWLEDGEMENT

The authors gratefully acknowledge the support and permission of Universitas Teknologi Yogyakarta, Universitas Pembangunan Nasional Veteran Yogyakarta, and AGH University of Krakow to conduct this research. Their collaboration and resources were essential in the successful completion of this study. The authors would also like to express their appreciation for the publication of this article, which showcases the research findings.





REFERENCES

- [1] A. A. Alahmari, A. A. Khan, F. A. Alamri, Y. S. Almuzaini, A. K. Habash, and H. Jokhdar, "Healthcare policies, precautionary measures and outcomes of mass gathering events in the era of COVID-19 pandemic: expedited review," *Journal of Infection and Public Health*, Mar. 2023, doi: 10.1016/j.jiph.2023.03.026.
- [2] H. Karimian, Q. Fan, Q. Li, Y. Chen, and J. Shi, "Spatiotemporal transmission of infectious particles in environment: a case study of COVID-19," *Chemosphere*, vol. 335, Sep. 2023, doi: 10.1016/j.chemosphere.2023.139065.
- [3] M.-C. Couture, L. Walicek, K. L. L'Engle, and A. K. Regan, "High engagement in nonpharmaceutical interventions and their associations with reduced COVID-19 among US college students," *BMC Public Health*, vol. 23, no. 1, May 2023, doi: 10.1186/s12889-023-15916-0.
- [4] C. Cheng, "Real-time mask detection based on SSD-MobileNetV2," in *2022 IEEE 5th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE)*, Nov. 2022, pp. 761–767, doi: 10.1109/AUTEEE56487.2022.9994442.
- [5] M. A. S. Ai *et al.*, "Real-time facemask detection for preventing COVID-19 spread using transfer learning based deep neural network," *Electronics*, vol. 11, no. 14, Jul. 2022, doi: 10.3390/electronics11142250.
- [6] H. Lin, R. Tse, S.-K. Tang, Y. Chen, W. Ke, and G. Pau, "Near-realtime face mask wearing recognition based on deep learning," in *2021 IEEE 18th Annual Consumer Communications and Networking Conference (CCNC)*, Jan. 2021, pp. 1–7, doi: 10.1109/CCNC49032.2021.9369493.
- [7] X. Kong *et al.*, "Real-time mask identification for COVID-19: an edge-computing-based deep learning framework," *IEEE Internet of Things Journal*, vol. 8, no. 21, pp. 15929–15938, Nov. 2021, doi: 10.1109/JIOT.2021.3051844.
- [8] D. S. R. G and J. A. V., "Social distance detection using deep learning and risk management system," *arXiv:2304.10259*, Apr. 2023.
- [9] R. N. S and M. N., "Computer-vision based face mask detection using CNN," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, Jul. 2021, pp. 1780–1786, doi: 10.1109/ICCES51350.2021.9489098.
- [10] G. Kaur *et al.*, "Face mask recognition system using CNN model," *Neuroscience Informatics*, vol. 2, no. 3, Sep. 2022, doi: 10.1016/j.neuri.2021.100035.
- [11] N. Fafous, M.-R. Vemparala, A. Frickenstein, L. Frickenstein, M. Badawy, and W. Stechele, "BinaryCoP: binary neural network-based COVID-19 face-mask wear and positioning predictor on edge devices," in *2021 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW)*, Jun. 2021, pp. 108–115, doi: 10.1109/IPDPSW52791.2021.00024.
- [12] S. Saifullah *et al.*, "Nondestructive chicken egg fertility detection using CNN-transfer learning algorithms," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 9, no. 3, pp. 854–871, 2023, doi:




- 10.26555/jiteki.v9i3.26722.
- [13] S. Asif, Y. Wenhui, Y. Tao, S. Jinhai, and K. Amjad, "Real time face mask detection system using transfer learning with machine learning method in the era of COVID-19 pandemic," in *2021 4th International Conference on Artificial Intelligence and Big Data*, pp. 70–75, 2021, doi: 10.1109/ICAIBD51990.2021.9459008.
- [14] W. Fang *et al.*, "A deep learning-based approach for mitigating falls from height with computer vision: convolutional neural network," *Advanced Engineering Informatics*, vol. 39, pp. 170–177, Jan. 2019, doi: 10.1016/j.aei.2018.12.005.
- [15] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: a review," *Electronics*, vol. 9, no. 8, Jul. 2020, doi: 10.3390/electronics9081188.
- [16] M. O. Oloyede, G. P. Hancke, and H. C. Myburgh, "A review on face recognition systems: recent approaches and challenges," *Multimedia Tools and Applications*, vol. 79, no. 37–38, pp. 27891–27922, Oct. 2020, doi: 10.1007/s11042-020-09261-2.
- [17] Y.-F. Li and H. Ying, "Disrupted visual input unveils the computational details of artificial neural networks for face perception," *Frontiers in Computational Neuroscience*, vol. 16, Nov. 2022, doi: 10.3389/fncom.2022.1054421.
- [18] G. H. Christa, J. J. A. K., and K. M. Sagayam, "CNN-based mask detection system using OpenCV and MobileNetV2," in *2021 3rd International Conference on Signal Processing and Communication (ICSPC)*, May 2021, pp. 115–119, doi: 10.1109/ICSPC51351.2021.9451688.
- [19] S. Kumar, D. Yadav, H. Gupta, M. Kumar, and O. P. Verma, "Towards smart surveillance as an aftereffect of COVID-19 outbreak for recognition of face masked individuals using YOLOv3 algorithm," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 8381–8403, 2023, doi: 10.1007/s11042-021-11560-1.
- [20] M. SivaKumar, N. Saranprasad, N. S. Sridharan, and V. S. Praveen, "Comparative analysis of CNN and Viola-Jones for face mask detection," *Journal of Physics: Conference Series*, vol. 1916, no. 1, May 2021, doi: 10.1088/1742-6596/1916/1/012043.
- [21] A. Rahim, K. Kusriani, and E. T. Luthfi, "Convolutional neural network untuk kalasifikasi penggunaan masker," *Inspiration: Jurnal Teknologi Informasi dan Komunikasi*, vol. 10, no. 2, Dec. 2020, doi: 10.35585/inspir.v10i2.2569.
- [22] A. H. Pratomo, W. Kaswidjanti, A. S. Nugroho, and S. Saifullah, "Parking detection system using background subtraction and HSV color segmentation," *Bulletin of Electrical Engineering and Informatics (BEEI)*, vol. 10, no. 6, pp. 3211–3219, 2021, doi: 10.11591/eei.v10i6.3251.
- [23] N. Anusha, S. Gupta, Y. N. Naidu, M. Ruchitha, and R. Pandey, "Face mask and social distance detection using deep learning models," in *Computational Vision and Bio-Inspired Computing*, 2023, pp. 461–484.
- [24] H. Goyal, K. Sidana, C. Singh, A. Jain, and S. Jindal, "A real time face mask detection system using convolutional neural network," *Multimedia Tools and Applications*, vol. 81, no. 11, pp. 14999–15015, 2022, doi: 10.1007/s11042-022-12166-x.
- [25] J. Yu and W. Zhang, "Face mask wearing detection algorithm based on improved YOLO-v4," *Sensors*, vol. 21, no. 9, May 2021, doi: 10.3390/s21093263.
- [26] R. C. Lanfranco, H. Rabagliati, and D. Carmel, "The importance of awareness in face processing: a critical review of interocular suppression studies," *Behavioural Brain Research*, vol. 437, Feb. 2023, doi: 10.1016/j.bbr.2022.114116.
- [27] M. K. Rusia and D. K. Singh, "A comprehensive survey on techniques to handle face identity threats: challenges and opportunities," *Multimedia Tools and Applications*, vol. 82, no. 2, pp. 1669–1748, 2023, doi: 10.1007/s11042-022-13248-6.
- [28] S. Saifullah and R. Drezewski, "Enhanced medical image segmentation using CNN based on histogram equalization," in *Proceedings of the 2nd International Conference on Applied Artificial Intelligence and Computing*, pp. 121–126, 2023, doi: 10.1109/ICAIC56838.2023.10141065.
- [29] B. U. H. Sheikh and A. Zafar, "RRFMDS: rapid real-time face mask detection system for effective COVID-19 monitoring," *SN Computer Science*, vol. 4, no. 3, Mar. 2023, doi: 10.1007/s42979-023-01738-9.
- [30] A. Jangra, "Face mask detection 12K images dataset," *Kaggle*, 2020. <https://www.kaggle.com/datasets/ashishjangra27/face-mask-12k-images-dataset> (accessed Jan. 11, 2023).

BIOGRAPHIES OF AUTHORS






Suhirman     graduated with a Bachelor of Computer Science from STMIK AKAKOM Yogyakarta, Indonesia in 1998. M.Sc. in Computer Science from Gadjah Mada University, Indonesia in 2004. Holds Ph.D. degree in Computer Science from Universiti Malaysia Pahang, Malaysia in 2016. He has been Associate Professor at the University Technology of Yogyakarta since 2008. He is currently the Director of the Postgraduate Program at the University Technology of Yogyakarta. His research interests are artificial intelligence, information systems, and data mining. He can be contacted via email: suhirman@uty.ac.id.






Shoffan Saifullah    received a bachelor's degree in informatics engineering from Universitas Teknologi Yogyakarta, Indonesia, in 2015 and a master's degree in computer science from Universitas Ahmad Dahlan, Yogyakarta, Indonesia, in 2018. He is a lecturer at Universitas Pembangunan Nasional "Veteran" Yogyakarta, Indonesia. His research interests include image processing, computer vision, and artificial intelligence. He is currently a Ph.D. student at AGH University of Krakow, Poland, with a concentration in the field of artificial intelligence (bio-inspired algorithms), image processing, and medical analysis. He can be contacted at email: shoffans@upnyk.ac.id and saifulla@agh.edu.pl.






Ahmad Tri Hidayat    is a lecturer with the position of Assistant Professor at Yogyakarta University of Technology (UTY), Yogyakarta, Indonesia. He obtained a bachelor's degree in informatics engineering from Yogyakarta University of Technology in 2011. He then earned a master's degree in informatics from the Islamic University of Indonesia in 2017. His research areas include information systems, information systems audit, and enterprise information systems. He can be reached via email at ahmad.tri.h@uty.ac.id.



M. Apriandi Kusuma    received a bachelor's degree in informatics from Universitas Teknologi Yogyakarta, Indonesia, in 2022. He developed a real-time mask detection system using CNN. Now, he is a Web Developer at Desa Digital, a digital village development company in Indonesia with highly skilled and experienced web developer, with expertise in HTML, CSS, JavaScript, PHP, and MySQL. He can be contacted at email: apriandikusuma50@gmail.com.



Rafał Dreżewski    received the M.Sc., Ph.D., and D.Sc. (Habilitation) degrees in computer science from the AGH University of Krakow, Poland in 1998, 2005, and 2019, respectively. Since 2019, he has been an Associate Professor with the Institute of Computer Science, AGH University of Krakow, Poland. He is the author of more than 80 papers. His research interests include bio-inspired artificial intelligence algorithms and agent-based modeling and simulation of complex and emergent phenomena. He can be contacted at email: drezew@agh.edu.pl.