Green Intelligent Systems and Applications

Research Article
Volume 2(2), 2022, 71-83
https://doi.org/10.53623/gisa.v2i2.109

Attendance System with Face Recognition, Body Temperature, and Use of Mask using Multi-Task Cascaded Convolutional Neural Network (MTCNN) Method

Noor Cholis Basjaruddin[1*], Edi Rakhman[1], Yana Sudarsa[1], Moch Bilal Zaenal Asyikin[2], Septia Permana[3]

[1] Department of Electrical Engineering, Politeknik Negeri Bandung, Indonesia.
[2] Department of Refrigeration, Air Conditioning, and Energy Engineering, National Chin-Yi University of Technology, Taiwan.
[3] Depatment Digital and Next Business, PT. Telkom Indonesia Tbk, Indonesia.

*Correspondence: noorcholis@polban.ac.id

SUBMITTED: 19 August 2022; REVISED: 29 September 2022; ACCEPTED: 1 October 2022

ABSTRACT: The application of health protocols in educational, office, or industrial environments can be made by changing old habits that can spread COVID-19. One of them is the habit of recording attendance, which still requires direct physical contact. In this research, an attendance system based on facial recognition, body temperature checks, and mask use using the multi-task cascaded convolutional neural network (MTCNN) has been developed. This research aims to integrate a facial recognition system, a mask detection system, and body temperature reading into an attendance recording system without the need for direct physical contact. The attendance system offered in this study can minimize the spread of COVID-19. So, it has enormous potential for use in educational, office, and industrial environments. The focus of this research is to create an attendance system by integrating the application of face recognition, body temperature, and the use of masks using a pre-trained model. Based on the research results, an attendance system was successfully developed where the results of face recognition, mask detection, and body temperature were displayed on the machine screen and attendance platform. Facial recognition testing on the original LFW dataset has an accuracy of 66.45%. The accuracy of the dataset reaches 92-100%. In addition, the intelligent attendance platform has been successfully developed with user management, machine service, and attendance service features. The results of the attendance record are successfully displayed on the platform or through the download feature.

KEYWORDS: COVID-19; attendance system; facial recognition; mask detection; body temperature; MTCNN

1. Introduction

The COVID-19 outbreak has become a phenomenon that has a significant impact on all countries in the world. As of February 2021, more than 100 million people have been confirmed to be infected with the virus. The number of deaths worldwide is over two million [1]. Indonesia is one of the countries with many cases reaching more than one million people, with a death toll of more than 32 thousand people [2]. The COVID-19 virus can spread from the mouth or nose of an infected person in the form of small particles (aerosols) when they cough, sneeze, or talk. A person can be infected with this virus if the aerosol enters his mouth or nose [3].

The Indonesian government has taken various measures to prevent the spread of the COVID-19 virus, both at the central and regional levels, such as closing access to arrivals and departures from and to foreign countries [4], implementing large-scale social restrictions, and enforcing regulations on community-based activities [5]. The impact of government policies has limited community activities such as carrying out online learning and working at home. In addition, the government is implementing new life adaptations by promoting the 3M slogans (wearing masks properly, maintaining distance and avoiding crowds, washing hands with soap) and 3T (Testing, Tracing, and Treatment) to suppress the spread of COVID-19 [6]. One of the old habits that must change in the COVID-19 situation is the attendance system that uses objects or body parts in direct contact with specific tools.

Attendance systems are widely applied both in educational, office and industrial environments. [7] proposed contactless attendance with features such as temperature detection and real-time capability. However, the proposed system was not equipped with database management. In [8], an attendance management system using a smart IC card was presented. The latter emphasized the use of unique identification based on the physical tokens which can be transferred from one person to another. Thus, this problem needs to be solved by using real-time photographic identification, as proposed in this paper. So we need a new plan during the COVID-19 pandemic that can record attendance without direct physical contact. The main problem to be solved in this research is to find an algorithm to integrate a face recognition-based attendance system with the detection of mask use and body temperature check. The facial recognition algorithm is challenged, which means it was challenging to perform the face recognition process when the target to be recognized uses a mask. This shows the urgency and challenges of the situation brought on by the COVID-19 pandemic, that recognition must be accomplished with limited features of the human face, such as the nose and lips. The issues discussed in [9] stated the urgent need for face recognition during the pandemic, especially in developing countries, while the cognitive aspects were discussed in [10], stating the risk of inaccurate identification should the identification method not be properly examined.

2. Materials and Methods

2.1. Multi-Task Cascaded Convolutional Neural Network (MTCNN)

In performing face recognition, one of the steps that must be passed is face detection. In this research, the method used to perform face detection is the Multi-Task Cascaded Convolutional Neural Network (MTCNN). MTCNN is widely used because it can achieve good results on multiple product datasets. In addition, the MTCNN method has the ability to recognize facial features such as eyes and mouth, which is then called landmark detection.

MTCNN uses a cascading structure. Cascade has the meaning of several networks working together where the output of the first network was the input to the next network and so on. The overall framework of MTCNN can be seen in Figure 1. Based on Figure 1, the given image was resized into several different scales, which are then called image pyramids. Then the image follow the cascade structure in three stages as follows.

Stage 1-At this stage, a Network Proposal (P-Net) is used, which produced a candidate face area with a bounding box. Then bounding box regression is used to filter the candidates. Then non-maximum suppression (NMS) was used to combine several overlapping candidates.
Stage 2-The candidate's face area generated from stage 1 became input in the second stage using Refine Network (R-Net). At this stage, bounding box regression and NMS was used to get rid of several candidates with large error values.
Stage 3 - At stage 3 using the Output Network (O-Net) which has a process similar to stage 2. However, this stage results in five facial landmark positions.

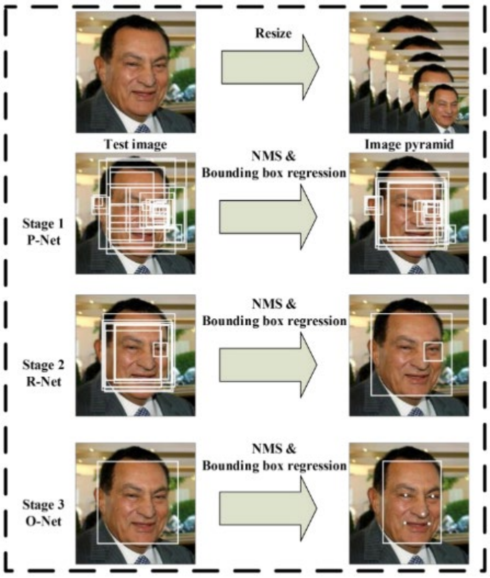


Figure 1. MTCNN structure [11].

2.2. System Design

The block diagram in this research can be seen in Figure 2. Before creating an attendance system platform, several steps need to be taken, such as setting up a facial recognition system, mask detection, and body temperature detection.

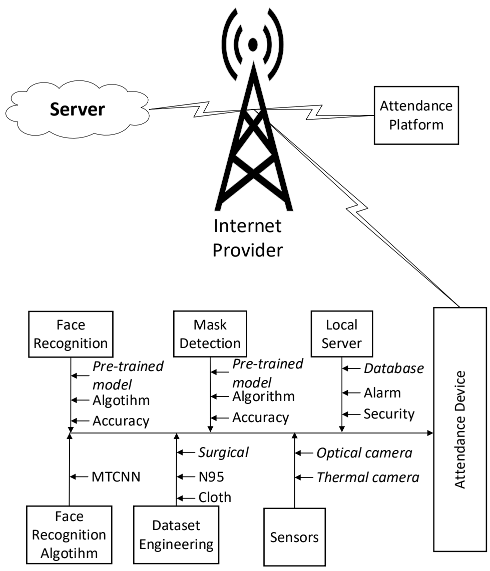


Figure 2. Block diagram.

2.2.1. Face recognition design

To produce an excellent facial recognition system, a strategy in designing facial recognition algorithms is very important. The facial recognition design using MTCNN can be seen in Figure 3.

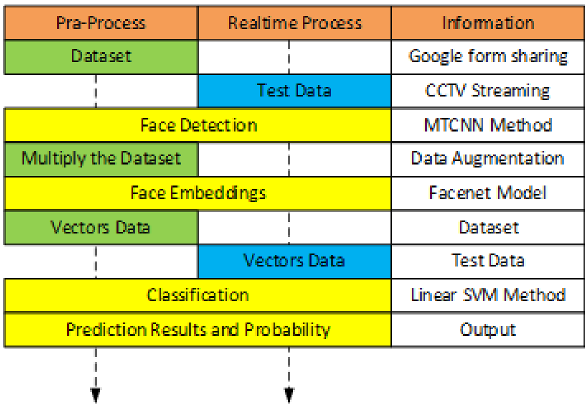


Figure 3. Facial recognition design.

Based on Figure 3, there are two stages for facial recognition: the preprocessing and the second is real-time processing. According to that stage, all preprocessing stages must be carried out in real-time to carry out facial recognition. The pre-processing stages that must be done consist of 7 stages as follows:

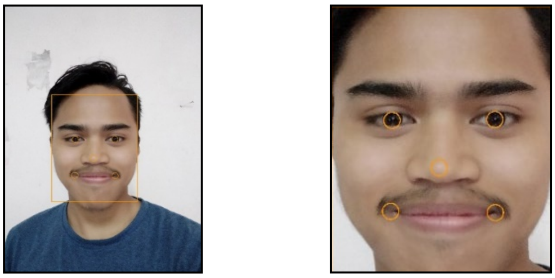
1) Dataset acquisition
To complete the preprocessing stage, the dataset collection is done by sharing a Google form. Volunteers willing to fill out the form must complete their identity data and upload at least one photo. All photos collected was used as a dataset. In the simulation, only those who filled in the dataset could carry out the facial recognition process.

2) Face recognition
In carrying out facial recognition, one of the steps that must be passed is face detection. One of the deep learning methods used to perform facial recognition is the Multi-Task Cascaded Convolutional Neural Network (MTCNN) [11]. The realization is made in a face image where part of the body is still there. In the testing process, the library used is provided by Iván de Paz Centeno in his MTCNN project [11]. The collected photos are selected to be cut in the face area using the MTCNN method, as shown in Figure 4. This stage aims to obtain maximum facial data in the face embedding process in the next step.

3) Multiply the dataset by two.
Data augmentation is a technique of multiplying data from one image into many by rotating, cropping, resizing, and other means [12]. Techniques and strategies for performing data augmentation are essential and may even be more important than having a network structure to be used [13]. Google has said that the alignment method can improve the accuracy of face recognition with faceNet models from 98.87% to 99.63% [14, 15].

Because the minimum number of photos requested is one, multiplying the dataset with data augmentation is carried out. The techniques used are:

a. Horizontal flip
b. 3, 6, and 9 degrees left rotation
c. 3, 6, and 9 degrees right rotation
d. Translation
e. 50% brightness reduction



(**a**) (**b**)

Figure 4. (a) bounding box and face landmark by using MTCNN; **(b)** The result of cutting the face area.

4) Face embeddings
This research used the pre-trained faceNet model [15]. FaceNet is a model developed by Schroff, et al. which can be used for face recognition, verification, and clustering. FaceNet is a deep convolutional neural network with 22 layers that produced 128-D embedding output. The loss function used in faceNet is known as 'triplet loss'. The model has been trained on the MS-Celeb-1M Dataset [16]. The dataset used in training is ~ 10 million faces with 100,000 classes. Meanwhile, the CNN architecture used in the faceNet model is the Inception-ResNet-V1 architecture [17].

5) Vectors data
The face embeddings stage produced data from each photo incorporated in the vector form dataset. The vector data was used for the next step, namely classification.

6) Classification
The classification technique compares the vector data that was achieved from the test data with the vector data attained from the dataset. The classification method is Linear SVM. By using this technique, the system achieved facial predictions and probabilities.

7) Prediction result and probability
The final output of facial recognition is face prediction from the test data against the dataset. The prediction results was used for recording the attendance system.

2.2.2. Mask detection design

Mask detection is carried out from the extract of facial images received by the system through the camera. Like the face recognition algorithm, the mask detection algorithm compare the results of the extract of the face with a deep learning model trained on the dataset with the mask object. Figure 5 shows the flow of mask detection in this study.

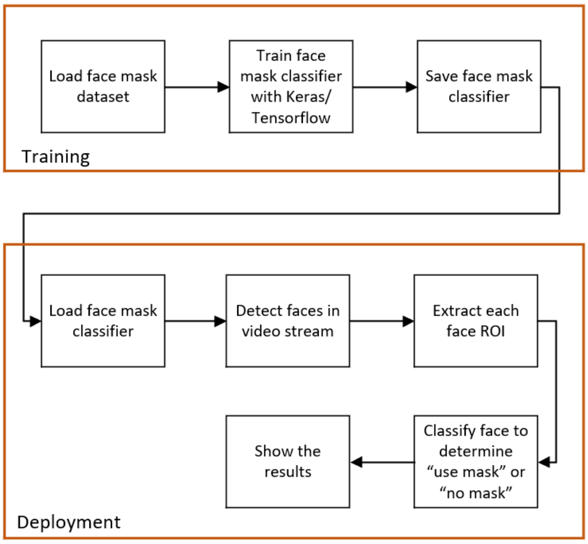


Figure 5. Mask detection flow.

2.2.3. Body temperature detection design

Body temperature is detected using a thermal imaging infrared temperature sensor. Body temperature checks are carried out when the system detects faces captured through the camera lens at the face recognition stage. The temperature reading results was compared with the threshold value that has been set as the reference value. Users are allowed to enter if the results of checking body temperature do not exceed the threshold value. Figure 6 shows the body temperature detection algorithm used in this study.

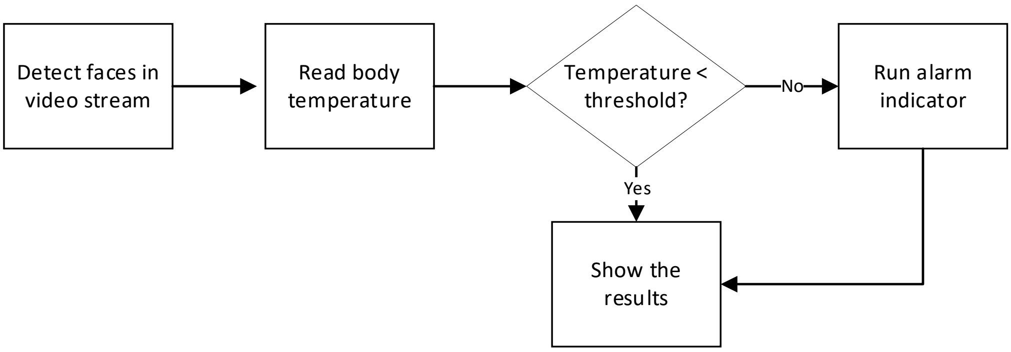


Figure 6. Body temperature detection algorithm.

2.2.4. Hardware (Machine)

The selection of hardware that was used is critical because it must consider specifications that meet three main features: face recognition, mask detection, and body temperature detection. Figure 7 shows the hardware used in this study. The device integrates sensors such as optical cameras, thermal imaging, infrared temperature sensors, and other features such as displays and speakers. In addition, the device is also equipped with an LCD touch screen and an Android-based processor. The hardware specifications used can be seen in Table 1.

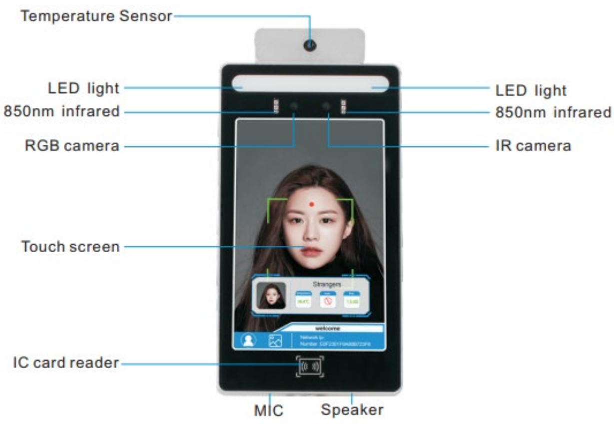


Figure 7. Hardware (Machine).

Table 1. Machine Specification

Parameters		Specifications
Thermal	Sensor	Thermal imaging infrared temperature sensor
	Temperature measurement	0° - 98°C
	Detection distance	0.5m – 3m
Visible Parameters	Lens	Focal length 1.8mm, Field of View: 118°
	Dynamic range	≥ 100dB
	S/N Ratio	≥ 43dB (AGC OFF)
	Exposure mode	Program mode, shutter mode (1/5 – 1/20.000s)
	White balance	Auto, indoor, outdoor, sodium lamp mode, manual
	Digital noise reduction	Support DNR, 3DNR
	Day and night mode	Fixed color

2.2.5. Attendance system platform design

All face recognition, mask detection, and body temperature detection were recorded on the attendance platform. The data was processed and analyzed by the platform to become user attendance data registered on the platform. The machine service handles the integration between the system hardware (machine) and the platform. In addition, the API gateway takes the integration between platform and client features, in this case, using a browser. The software architecture of the platform is shown in Figure 8.

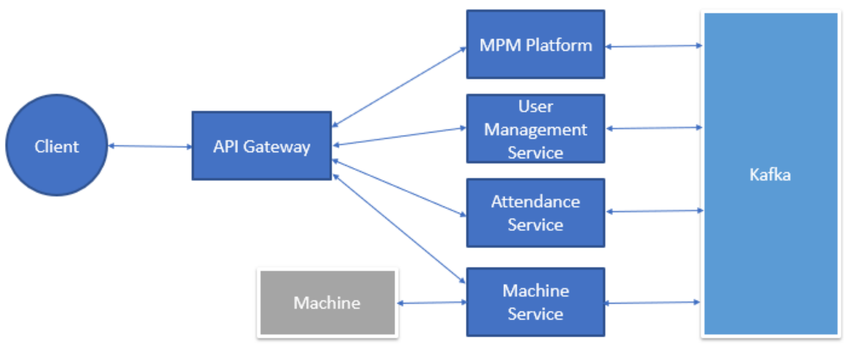


Figure 8. Platform algorithm.

3. Results and Discussion

3.1. Facial Recognition with MTCNN method and mask detection

In testing the faceNet model used, a prediction test was carried out on photos by trying a dataset sourced from the internet, namely the original LFW [18], which has a dataset of more than 13 thousand photos. Besides, in testing, predictions were also carried out on the datasets collected before the augmentation process and after going through the augmentation process. Table 2 shows the test results and comparisons with other studies with the same concept. Face recognition has been successfully implemented in Xiao and Weiwei's study. The system developed with the CNN model reached an accuracy rate of 98.1% [19]. In addition, previous study succeeded in implementing the application of the faceNet model on the surveillance system with an accuracy of 97% [20].

Table 2. Comparison with other studies.

Dataset	Number of Photos	Number of classes	Accuracy (%)
Original LFW	13233	5749	66.45
Xiao and Weiwei study	126	13	98.10
Edwin Jose et al. study	5000	10	97.00
Own dataset	270	64	92.22
Own dataset + Augmentation	2700	64	100.00

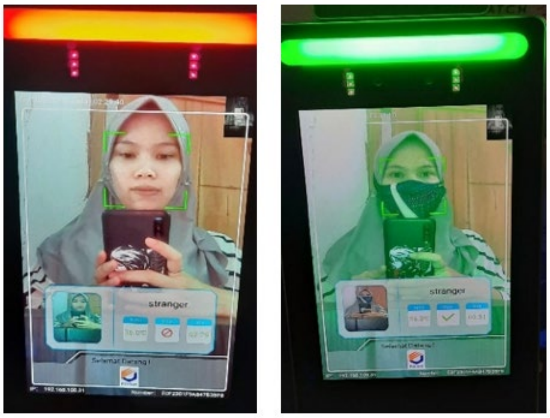
According to Table 2, facial recognition testing on the original LFW dataset has an accuracy of 66.45%. This value is much smaller than the accuracy of the own dataset, which reaches 92-100%. However, this happens because the original dataset, LFW, has far more classes than the own dataset. Face recognition testing is carried out on the system hardware (machine) as shown in Figure 9.

Figure 9. The machine.

The test results show that registered users can be recognized when the machine displays the user name, as shown in Figure 10 (a). While the user is not registered, the stranger's name was displayed by the machine in Figure 10 (b). Testing on the use of masks is carried out on users with masks and without masks. The test results show that the system detects the user without a mask when the LCD interface, which shows the "Mask" description, does not have a check symbol, as shown in Figure 11 (a). In addition, the system has also succeeded in detecting masks for users who are wearing masks. On the LCD interface, which shows the "Mask" description, there is a check symbol as shown in Figure 11 (b).

(**a**) (**b**)

Figure 10. Face recognition test result for (**a**) Registered user; (**b**) Unregistered user

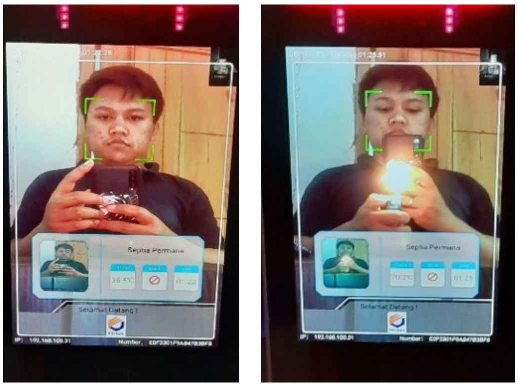


(**a**) (**b**)

Figure 11. Mask detection test results for (**a**) Not wearing a mask; (**b**) Wearing a mask.

3.3. Body temperature check

Body temperature detection tests are carried out by reading normal and abnormal body temperatures (using a gas lighter). The measurements at normal body temperature can be seen in Figure 12 (a), where the LCD shows temperature information of 36.5 °C with green text. While measuring abnormal body temperature by lighting a gas fire, the system measures a temperature of 70.2 °C with red text, as shown in Figure 12 (b).



(**a**) (**b**)

Figure 12. Body temperature checks result; (**a**) Normal temperature (**b**) High temperature

3.4. Attendance system platform

System management is a significant part of the development of an attendance system. The attendance system platform has the primary function of a management system. Features such as adding users, analyzing attendance levels, and displaying the number of attendees in real-time can be displayed on this platform. In this research, platform development and integration with machines have been carried out. All machine reading results are sent to the server and stored in the database. While at the client level, the platform can be accessed via a browser. The features that have been added to the attendance platform are listed in Table 3.

Table 3. Attendance System Platform Features

Machine Service	Attendance Services	User Management
Attendance listener	Attendance list	CRUD User
Machine configuration	Real-time integration record	CRUD class
User management	Mapping user by id, class, date	
Photo management	Attendance report	

The solution proposed by the author has a unique data structure, which differentiates it from the above-listed platforms. Therefore, in all experiments, the data used was self-generated by including several participants in the testing. Testing on the attendance platform was carried out by running the entire attendance system. The stages in using the attendance system were as follows:

1) User registration

User registration is done on the attendance platform by selecting the user menu, as shown in Figure 13. Next, complete the required account data, including email, id, name, password, and class.

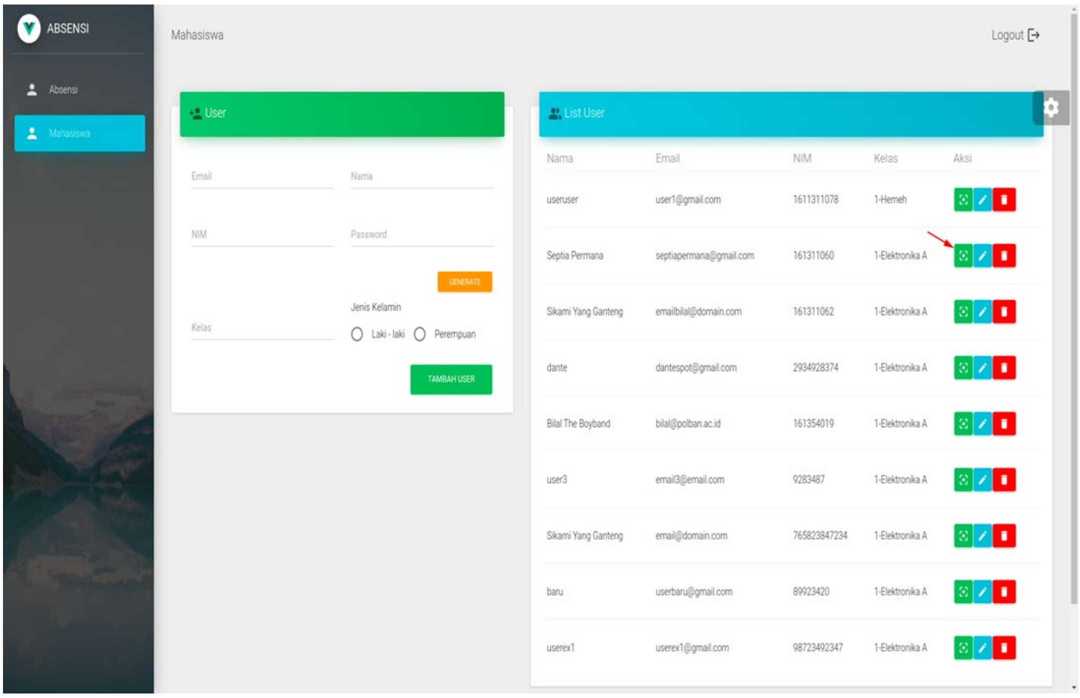


Figure 13. Registration form.

2) User's data input

After creating a new user account during the registration stage, the next step is to enter the new user data. The data that must be completed includes photos and machine IP as shown in Figure 14.

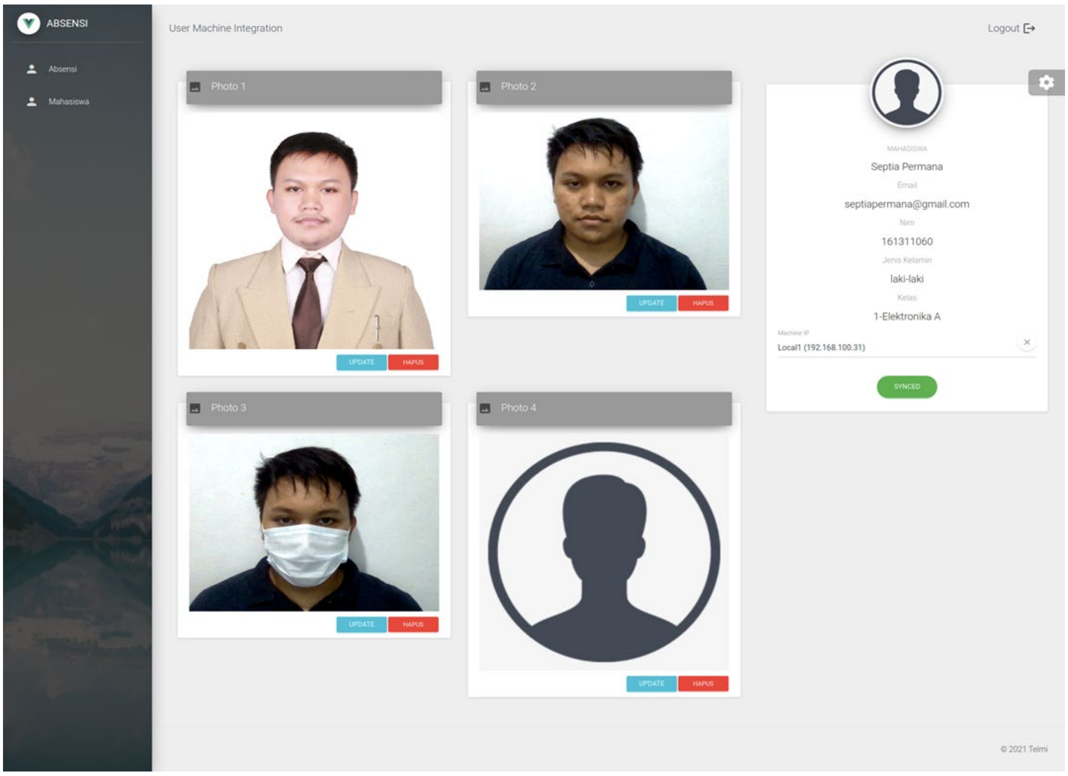
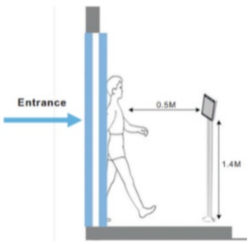


Figure 14. Data input form.

3) Machine Scanning

The machine scanning stage is the stage where the user stands right in front of the machine. The machine installation can be seen in Figure 15a. As can be seen from the figure, the device is designed to work 0.5 meters from the user and is placed 1.4 meters from the ground. The results of the system reading include face recognition, mask detection, and body temperature detection. All tasks are displayed on the LCD machine. In addition, all data is also sent to the server to be monitored on the platform and can also be downloaded, as shown in Figure 15b.



(a)

B	C	D	E	F	G
Nama	NIM	Kelas	Masker	Temperature	Waktu
Septia Permana	161311060	1-Elektronika A	TRUE	36	2021-10-04 18:2
Septia Permana	161311060	1-Elektronika A	TRUE	36	2021-10-04 18:3
Septia Permana	161311060	1-Elektronika A	FALSE	36.3	2021-10-04 18:4
Septia Permana	161311060	1-Elektronika A	FALSE	36.2	2021-10-04 18:4
Septia Permana	161311060	1-Elektronika A	FALSE	36.4	2021-10-04 18:5

(b)

Figure 15. Machine installation (a) and attendance recap (b)

4. Conclusions

Based on the research results, an attendance system was successfully developed where the results of face recognition, mask detection, and body temperature were displayed on the machine screen and attendance platform. Facial recognition testing on the original LFW dataset has an accuracy of 66.45%. The accuracy of the dataset reaches 92-100%. In addition, the attendance platform has been successfully developed with user management, machine service, and attendance service features. The results of the attendance record are successfully displayed on the platform or through the download feature. However, testing is performed at the laboratory level. Further testing is needed for real-world applications, including increasing the number of classes in the test.

Acknowledgments

The author would like to thank the Center for Research and Community Service at the Politeknik Negeri Bandung for the financial support through the Penelitian Terapan Unggulan Jurusan (PTUJ) contract number B/80.1/PL1.R7/PG.00.03/ 2021.

References

[1] WHO Coronavirus Disease (COVID-19) Dashboard. (accessed on 11 February 2021). Available online: https://covid19.who.int/.

[2] World Health Organization. (accessed on 11 February 2021). Available online: https://covid19.who.int/table.

[3] How COVID spreads. Centers for Disease Control and Prevention. (accessed on 11 February 2021). Available online https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/how-covid-spreads.html.

[4] COVID-19 developments in Indonesia, AHK Indonesien, (accessed on 24 December 2021) Available online: https://indonesien.ahk.de/en/infocenter/news/news-details/covid-19-developments-in-indonesia.

[5] No one was listening: How did Indonesia's COVID-19 crisis get this bad?, ABC News, (accessed on 22 July 2022). Available online: https://www.abc.net.au/news/2021-07-23/indonesia-covid-19-coronavirus-epicentre-lockdown-delta-variant/100310858.

[6] Indonesia to ramp up testing, tracing ahead of reopening. (accessed on 22 July 2022). Available online: https://www.thejakartapost.com/news/2021/07/22/government-to-ramp-up-testing-tracing-ahead-of-reopening.html.

[7] Patra, D.; Agrawal, A.; Srivastav, A.; Kathirvelan, J. (2021). Contactless attendance cum temperature detection system with real-time alerts. *2021 Innovations in Power and Advanced Computing Technologies (i-PACT)*. https://doi.org/10.1109/i-PACT52855.2021.9696594.

[8] Geng, S.; Li, G.; Liu, W. (2012). Design and implement of Attendance Management System based on contactless smart IC Card. *2012 Proceedings of International Conference on Computer Science and Electronics Engineering*. https://doi.org/10.1109/ICCSEE.2012.196.

[9] Farman, H.; Khan, T.; Khan, Z.; Habib, S.; Islam, M.; Ammar, A. (2022). Real-time face mask detection to ensure COVID-19 precautionary measures in the developing countries. *Applied Sciences, 12*, 3879. https://doi.org/10.3390/app12083879.

[10] Manley, K.D.; Chan, J.C.; Wells, G.L. (2022). Improving face identification of mask-wearing individuals. *Cognitive Research: Principles and Implications, 7*, 27. https://doi.org/10.1186/s41235-022-00369-7.

[11] Zhang, K.; Zhang, Z.; Li, Z.; Qiao, Y. (2016). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, *23 (10)*, 1499-1503, https://doi.org/10.1109/LSP.2016.2603342.

[12] Khan, S.; Rahmani, H.; Shah, S.A.A.; Bennamoun, M. (2018). A Guide to Convolutional Neural Networks for Computer Vision. *Synthesis Lectures on Computer Vision, 8*, 1-207. http://doi.org/10.2200/S00822ED1V01Y201712COV015.

[13] Shijie, J.; Ping, W.; Peiyi, J.; Siping, H. (2017). Research on data augmentation for image classification based on convolution neural networks. *2017 Chinese Automation Congress (CAC).* https://doi.org/10.1109/CAC.2017.8243510.

[14] Parkhi, O. M.; Vedaldi, A.; Zisserman, A. (2015). Deep Face Recognition. *Proceedings of the British Machine Vision Conference, 41*, 1-12. https://doi.org/10.5244/C.29.41.

[15] Schroff, F.; Kalenichenko, K.; Philbin, J. (2015). FaceNet: A unified embedding for face recognition and clustering. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 815-823. https://doi.org/10.1109/CVPR.2015.7298682.

[16] Harvey, A.; LaPlace, J. (2019). Origins, Ethics, and Privacy Implications of Publicly Available Face Recognition Image Datasets, Available online:megapixels.cc.

[17] Szegedy, C.; Wei, L.; Yangqing, J.; Pierre, S.; Scoot, R.; Dragomir, A.; Dumitri, E.; Vincent, V.; Andrew, R. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1-9. https://doi.org/10.1109/CVPR.2015.7298594.

[18] Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E. (2007). Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments. Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, Erik Learned-Miller and Andras Ferencz and Frédéric Jurie, Oct 2008, Marseille, France.

[19] Xiao, J.; Weiwei, G. (2019). Face recognition algorithm based on Prewitt and convolutional neural network. *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*, 2007–2009. https://doi.org/10.1109/EITCE47263.2019.9095075.

[20] Jose, E.; Greeshma, M.; Haridas, M.T.P.; Supriya, M.H. (2019). Face Recognition based Surveillance System Using FaceNet and MTCNN on Jetson TX2. *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)*. https://doi.org/10.1109/ICACCS.2019.8728466.