

LeNet Convolutional Neural Network for Face Mask Usage Classification Using a Low-Cost Device

Radimas Putra Muhammad Davi Labib¹, Aryuanto Soetedjo², Sirojul Hadi³,
Parama Diptya Widayaka⁴

^{1,2}Institut Teknologi Nasional Malang

³Universitas Bumigora

⁴Universitas Negeri Surabaya

radimas@lecturer.itn.ac.id¹, aryuanto@lecturer.itn.ac.id², sirojuhadi@universitasbumigora.ac.id³,
paramawidayaka@unesa.ac.id⁴

Article Info:

Received: May 06, 2023, Revised: May 20, 2023, Accepted: June 01, 2023

Abstract-

Background: One of the efforts to prevent the spread of the COVID-19 virus is to wear a face mask in public places. However, there are still many people who use masks in the wrong way, and some do not even wear masks when in public places. From these problems, we need an image-based classification system that can be employed to identify the use of face masks. The system built must use a low-cost device to be purchased by sundry groups. In previous studies, some classification systems for face mask use were designed using various methods, but there were limitations. The Convolutional Neural Network (CNN) method provides high accuracy. However, it has a heavy computational level and cannot be used in real-time on low-cost devices. In contrast, the haar-cascade method provides a fast processing time but is less accurate than the CNN method.

Objective: In this article, research was conducted on the development of image processing algorithms for the classification process of face mask use using low-cost devices.

Methods: The method used was CNN with LeNet architecture which has a light computational level. In the machine learning process, a dataset of 400 images was used, which was split into 240 images for training needs and 160 images for validation needs.

Result: This study produced a classification with an accuracy rate of 98.75%. The prediction process that is carried out using a low-cost device requires an average time of 0.235 seconds.

Conclusion: This research showed that the system can be run in real time.

Keywords: Convolutional Neural Network, COVID-19, Classification, Face Mask, LeNet

Corresponding Author:

Radimas Putra Muhammad Davi Labib,
Teknik Elektro, Institut Teknologi Nasional Malang, Malang, Indonesia,
Email: radimas@lecturer.itn.ac.id

1. INTRODUCTION

The pandemic of COVID-19 that started at the end of 2019 not only increased the cause of death rates but also caused an economic crisis on a global scale. The first case was found in December 2019 in the city of Wuhan, Hubei Province, China [1]. This pandemic is caused by a new type of coronavirus, SARS-CoV-2, which can spread massively quickly [2]. Efforts to prevent the coronavirus spread can be made through vaccination and implementing health protocols [3]. Vaccination is done to form antibodies that can suppress the growth of the virus in the body. Health protocols are implemented in public places by keeping a distance, washing hands frequently, and using face masks [4, 5].

The use of face masks in public places can prevent the spread of the coronavirus through the air [6]. Many people still wear face masks incorrectly or do not even use face masks when in public places. This causes

How to Cite: R.P.M.D. Labib, A. Soetedjo, S. Hadi, and P. D. Widayaka, "LeNet Convolutional Neural Network for Face Mask Usage Classification using a Low-Cost Device" *Jurnal Bumigora Information Technology (BITe)*, vol. 5, no. 1, pp. 9~16 2023.

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

the recommended health protocols to be ineffective and the transmission rate of the coronavirus to be high. From these problems, a system is needed that can classify the level of correctness in wearing face masks so that they can give warnings if the use of face masks is inappropriate.

In previous studies, some classification systems for face mask use were designed using various methods, but there were limitations [7]. The CNN method provides a high level of accuracy, but it has a heavy computational level, so it cannot be used in real-time on low-cost devices. The haar-cascade method works by detecting two facial features: the nose and lips. If one or both features are detected, it is concluded that the face mask is not being used properly. The system has an accuracy rate of 88.89% and can run on low-cost devices. The main drawback of the system is the limitation in distinguishing whether the nose and lips are being covered by a mask or covered by another object, such as a hand. It provides fast processing time but has a lower accuracy than the CNN method. Further research has been carried out by Ahmadi and his team [8] in 2021 regarding the detection of mask use using the haar cascade classifier method. This study's results are the system's accuracy using the haar cascade classifier method of 90.5%. This study was limited to a maximum facial condition of 30 degrees, turning sideways and looking up. Subsequent research has been conducted by Dicky Giancini and his team [9] regarding the identification of mask use using the CNN YOLOv3-Tiny algorithm. The results of this study are the accuracy of mask detection by the system by 90%. The built system can be used in real-time.

Based on previous research, the novelty of our research is that the detection of mask use is carried out using the LeNet-based CNN method. Using this method can increase the system's accuracy when detecting the use of masks. The system built also works in real-time, and the system is low-cost. The purpose of mask detection using the Convolutional Neural Network (CNN) method with a LeNet network structure is to provide a low computational level but have high enough accuracy so that it can still be embedded in low-cost devices.

2. METHODS

2.1. Research Stages

The stages involved in this research include literature study, dataset collection, algorithm design, implementation, and evaluation. The stages of the research can be seen in Figure 1. A literature study is needed to design programming algorithms to suit existing theories. The collection of datasets is needed as learning data for the system, as well as material for evaluation and validation of the programming algorithms that have been designed. The programming algorithm was designed using the Python programming language.

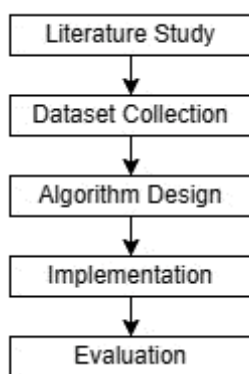


Figure 1. Research Stages

2.2. LeNet

The LeNet network structure is a CNN model originally used as an image-based handwriting classification system [10]. In its development, the LeNet network structure can also be used as a classification system for various objects such as diseases, odor types, and also cataract maturity levels [11–14]. This network structure

consists of a convolution layer, a pooling layer, and a fully-connected layer [15–18]. The LeNet network structure used in this study is shown in Figure 2.

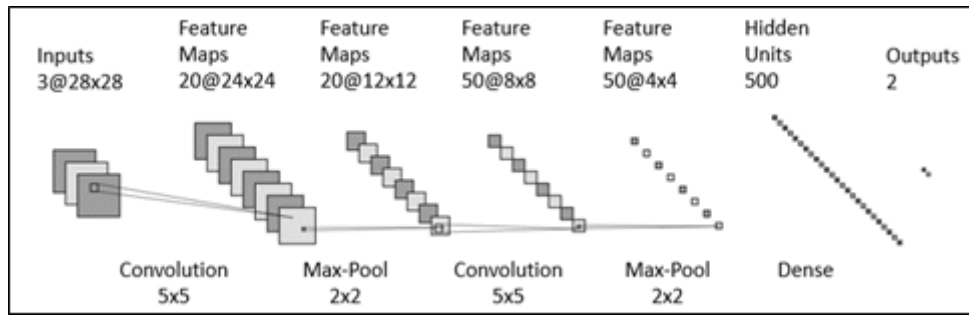


Figure 2. LeNet Network Structure

The LeNet network structure is built with seven layers of artificial neural networks. The first layer is the input which contains an image matrix with three channels representing the basic colors Red, Green Blue (RGB). The second layer is a feature map measuring 24×24 pixels with a depth of 20 channels which is the convolution result of the first layer. The third layer is a feature map measuring 12×12 pixels with a depth of 20 channels, resulting from max pooling from the second layer. The fourth layer is a feature map measuring 8×8 pixels with a depth of 50 which is the convolution result of the third layer. The fifth layer is 4×4 pixels feature map with a depth of 50 resulting from max pooling from the fourth layer. The feature extraction process occurs from the second to the fifth layer. Then a dense process is carried out on the feature map in the fifth layer to produce 500 hidden units in the sixth layer. Moreover, the last process is fully connected to get the prediction results which will be displayed on the seventh layer.

2.3. Dataset

The dataset is a collection of data used in the process of machine learning in artificial neural network systems [19]. This study used a dataset of 400 images with details of 200 positive data images and 200 negative data images. Positive data consists of images of various faces with the correct use of face masks. Meanwhile, negative data consists of images of various faces using the wrong face mask and images of various faces without using a face mask. The dataset is obtained by capturing the process manually using the internal webcam camera from the laptop. The amount of this dataset is divided into 60% for machine learning purposes and 40% for validation purposes. Some of the datasets can be seen in Figure 3.



Figure 3. Dataset: a) Negative Data, b) Positive Data

2.4. Machine Learning

Artificial neural networks are designed to mimic the work of the human brain through a combination of input data, weights, and biases [20]. These elements work together to identify, classify, and accurately describe objects. A deep neural network consists of several layers of interconnected nodes. Artificial neural network models require a learning process based on the dataset that has been prepared to get the desired predictive results. The stages of machine learning can be seen in Figure 4.

The first stage is the initialization of the current epoch, max epoch, and model of the CNN which will be used. The second stage is preparing the dataset by loading it from memory. The third stage is splitting the dataset into sixty percent for training purposes and forty percent for validation purposes. The next stage is training the model, which consists of two processes, including forward propagation and backpropagation. The training is carried out as many as the max epoch specified in the first stage. The results of training the CNN model will be stored in memory. Machine learning algorithms are built using the Python programming language combined with the open-source Tensorflow library.

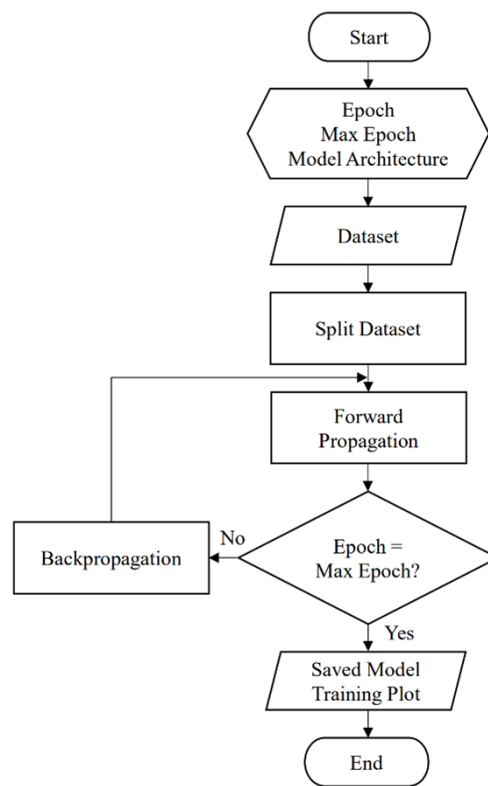


Figure 4. Stages of The Machine Learning Process

3. RESULTS AND DISCUSSION

3.1. CNN Model Training Process

The model training process was carried out with 30 epochs and a learning rate of 0.0002. The dataset is divided into 240 images for training needs and 160 images for validation needs. This process produces an artificial neural network model with a training accuracy of 98.05% and a validation accuracy of 98.75%. The results of the model training can be seen in Figure 5. Meanwhile, the model training process for each epoch is shown in Figure 6.

```

Command Prompt
6/6 [=====] - 0s 25ms/step - loss: 0.1276 - accuracy: 0.9415 - val_loss: 0.1236
- val_accuracy: 0.9625
Epoch 25/30
6/6 [=====] - 0s 25ms/step - loss: 0.1168 - accuracy: 0.9610 - val_loss: 0.1093
- val_accuracy: 0.9625
Epoch 26/30
6/6 [=====] - 0s 25ms/step - loss: 0.1273 - accuracy: 0.9415 - val_loss: 0.1042
- val_accuracy: 0.9625
Epoch 27/30
6/6 [=====] - 0s 25ms/step - loss: 0.1176 - accuracy: 0.9561 - val_loss: 0.0993
- val_accuracy: 0.9750
Epoch 28/30
6/6 [=====] - 0s 28ms/step - loss: 0.0906 - accuracy: 0.9659 - val_loss: 0.0950
- val_accuracy: 0.9750
Epoch 29/30
6/6 [=====] - 0s 25ms/step - loss: 0.0981 - accuracy: 0.9659 - val_loss: 0.0819
- val_accuracy: 0.9750
Epoch 30/30
6/6 [=====] - 0s 25ms/step - loss: 0.0908 - accuracy: 0.9805 - val_loss: 0.0848
- val_accuracy: 0.9875
[INFO] serializing network . . .
    
```

Figure 5. Training Results

These results prove that a classification system created using the LeNet network structure can be used to identify the use of face masks that are good and correct. This system also has a higher accuracy level than previous studies, which had an accuracy rate of 88.89%.

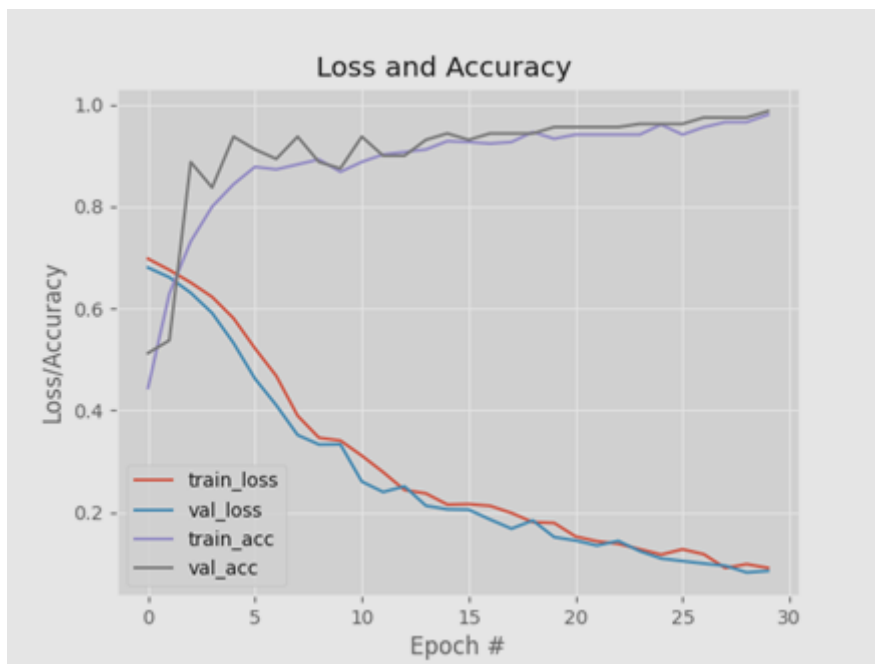


Figure 6. Training Process




In this process, there is a change in the value of the accuracy of training and validation in each epoch. In Figure 6, the accuracy value is improving, and the loss value is decreasing. Therefore, it can be concluded that the artificial neural network model that has been built has good classification quality because there is no under-fitting or over-fitting during the machine learning process.

3.2. Classification Results

System testing is carried out by applying image processing algorithms combined with the LeNet network model that has been built. The input is a static image of the face with various face mask usage conditions.

This experiment was carried out using a low-cost device, namely the raspberry pi 3 model B. The test results can be seen in Table 1.

Table 1. Face Mask Usage Classification Results

No	Input	Condition	Classification	Processing time (s)
1		No face mask	Unsafe	0.235664129
2		Nose and lips covered with a hand	Unsafe	0.235263824
3		Nose and lips covered with hands	Unsafe	0.235945702
4		The Face mask was used properly	Safe	0.234496593
5		The Face mask was used properly	Safe	0.237072229
6		The Face mask was used improperly	Unsafe	0.233533382
7		The Face mask was used improperly	Unsafe	0.235379457
8		The Face mask was used improperly	Unsafe	0.23592782
9		The Face mask was used improperly	Unsafe	0.232234716
10		The Face mask was used improperly	Unsafe	0.232787609
11		The Face mask was used properly	Safe	0.236531019
12		The Face mask was used properly	Safe	0.234724045

The results that have been presented indicate that the classification system built can provide optimal performance. The system can recognize whether the face mask has been used properly or not, and the system can also distinguish whether the nose and lips are being covered by a face mask or only being covered by hands. Based on the results of the tests that have been carried out, it can be concluded that the use of the CNN algorithm with the LeNet architecture can increase accuracy in detecting mask use in real-time compared to previous studies [7–9].

4. CONCLUSION

The face mask use classification algorithm based on the built CNN can work optimally with a validation accuracy rate of 98.75%. In addition, the system is also able to distinguish objects that are covering the nose and lips. Thus, it can be concluded that the development of the algorithm that has been carried out can correct the deficiencies of previous research. The prediction process that is carried out using a low-cost device requires an average time of 0.235 seconds; this shows that the system can be run in real time. However, the speed is felt to be lacking, so it is necessary to optimize the CNN parameters to reduce computation which might increase the prediction speed. Suggestions for further research, namely that a system for mask detection can be carried

out using the deep learning method.

ACKNOWLEDGEMENT

We want to thank Lembaga Penelitian dan Pengabdian Masyarakat (LPPM) of Institut Teknologi Nasional Malang for funding our research.

REFERENCES

- [1] Y. C. Wu, C. S. Chen, and Y. J. Chan, “The outbreak of COVID-19: An overview,” *Journal of the Chinese Medical Association*, vol. 83, no. 3, pp. 217–220, 2020.
- [2] S. Annas, M. Isbar Pratama, M. Rifandi, W. Sanusi, and S. Side, “Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia,” *Chaos, Solitons and Fractals*, vol. 139, pp. 1–7, 2020.
- [3] S. Manigandan, M. T. Wu, V. K. Ponnusamy, V. B. Raghavendra, A. Pugazhendhi, and K. Brindhadevi, “A systematic review on recent trends in transmission, diagnosis, prevention and imaging features of COVID-19,” *Process Biochemistry*, vol. 98, no. August, pp. 233–240, 2020.
- [4] D. K. Chu *et al.*, “Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis,” *The Lancet*, vol. 395, no. 10242, pp. 1973–1987, 2020.
- [5] W. Lyu and G. L. Wehby, “Community Use Of Face Masks And COVID-19: Evidence From A Natural Experiment Of State Mandates In The US,” *Health affairs (Project Hope)*, vol. 39, no. 8, pp. 1419–1425, 2020.
- [6] S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, “Rational use of face masks in the COVID-19 pandemic,” *The Lancet Respiratory Medicine*, vol. 8, no. 5, pp. 434–436, 2020.
- [7] R. P. M. D. Labib, S. Hadi, and P. D. Widayaka, “Low Cost System for Face Mask Detection Based Haar Cascade Classifier Method,” *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 21, no. 1, pp. 21–30, 2021.
- [8] A. Ahmadi, A. C. Saputra, and A. Lestari, “Rancang Bangun Aplikasi New Normal Covid-19 Deteksi Penggunaan Masker Menggunakan Haar Cascade Classifier,” *Jurnal Teknologi Informasi: Jurnal Keilmuan dan Aplikasi Bidang Teknik Informatika*, vol. 15, no. 2, pp. 199–209, 2021.
- [9] D. Giancini, E. Y. Puspaningrum, and Y. V. Via, “Identifikasi Penggunaan Masker Menggunakan Algoritma CNN YOLOv3-Tiny,” *Prosiding Seminar Nasional Informatika Bela Negara*, vol. 1, pp. 153–159, 2020.
- [10] A. E. Sawy, H. El-Bakry, and M. Loey, “CNN for handwritten arabic digits recognition based on LeNet-5,” *Advances in Intelligent Systems and Computing*, vol. 533, pp. 565–575, 2019.
- [11] G. Wei, G. Li, J. Zhao, and A. He, “Development of a LeNet-5 gas identification CNN structure for electronic noses,” *Sensors (Switzerland)*, vol. 19, no. 1, 2019.
- [12] T. Wang, C. Lu, G. Shen, and F. Hong, “Sleep apnea detection from a single-lead ECG signal with automatic feature-extraction through a modified LeNet-5 convolutional neural network,” *PeerJ*, vol. 2019, no. 9, pp. 1–17, 2019.
- [13] D. Rongshi and T. Yongming, “Accelerator Implementation of Lenet-5 Convolution Neural Network Based on FPGA with HLS,” *3rd International Conference on Circuits, System and Simulation, ICCSS 2019*, pp. 64–67, 2019.

- [14] R. Putra, M. Davi, S. Hadi, P. D. Widayaka, and I. Suryani, “Convolutional Neural Network for Cataract Maturity Classification Based on LeNet,” *Jurnal Bumigora Information Technology (BITe)*, vol. 5, no. 2, pp. 97–106, 2022.
- [15] D. Bau, J. Y. Zhu, H. Strobelt, A. Lapedriza, B. Zhou, and A. Torralba, “Understanding the role of individual units in a deep neural network,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 48, pp. 30 071–30 078, 2020.
- [16] S. Albawi, T. A. M. Mohammed, and S. Alzawi, “Layers of a Convolutional Neural Network,” *Ieee*, p. 16, 2019.
- [17] H. Gholamalinezhad and H. Khosravi, “Pooling Methods in Deep Neural Networks, a Review,” *IEEE*, pp. 1–16, 2020.
- [18] S. Kanai, Y. Yamanaka, Y. Fujiwara, and S. Adachi, “SigsofTmax: Reanalysis of the softmax bottleneck,” *Advances in Neural Information Processing Systems*, vol. 2018-Decem, no. NeurIPS, pp. 286–296, 2018.
- [19] G. Liu and J. Guo, “Bidirectional LSTM with attention mechanism and convolutional layer for text classification,” *Neurocomputing*, vol. 337, pp. 325–338, 2019.
- [20] A. F. Agarap, “Deep Learning using Rectified Linear Units (ReLU),” *IEEE*, no. 1, pp. 1–7, 2018.