Signal and Image Processing Letters
Vol. 5., No. 1, 2023, pp. 8-18
ISSN 2714-6677
https://simple.ascee.org/

Mask detection system at the entry of a room

Erik Herdiyanto [a,1], Abdul Fadlil [a,2,*]

[a] Department of Electrical Engineering, Universitas Ahmad Dahlan, Yogyakarta, Indonesia
[1] erik.herdiyanto2@gmail.com; [2] fadlil@ee.uad.ac.id
* corresponding author

ARTICLE INFO

Keywords
Mask Detection
CNN
TensorFlow
Raspberry pi

ABSTRACT

This study focuses on automatic mask detection tools that can open doors in a room to minimize violations of health protocols, one of which is the use of masks during the pandemic. The method used in this study is the CNN classification method. Where the CNN calcification method has several stages in it, including pre-processing, training, and testing. In the pre-processing, all image data used will be labeled using Labeling.axe. The training process at CNN uses TensorFlow framework version 1.15. In the testing process, the test and data testing will be carried out in real-time by entering new images and models that are made and then a classification process is carried out on objects caught by the camera, classified images are marked with boxes and names of data classes. This data class is divided into two, namely data on wearing masks and without masks. The results of the test were carried out by entering 200 facial image data. The system can correctly detect as much as 190 times from 200 data tested with an Accuracy rate of 95%. Based on the test results, it shows that the resulting model is good and suitable for the classification process of recognizing mask detection images. However, to produce a better model requires data with more variety and a larger amount of data.

1. Introduction

One of the problems the world is currently facing is the spread of the corona virus which has claimed many lives in every country, even in Indonesia, from March 2020 to November 2020 there have been 3,526,607 cases (covid19.kemkes.go.id. Info corona virus. Accessed on November 10, 2020). The government as a policy maker in this case does not remain silent in dealing with the COVID 19 problem, where there have been many regulations and policies issued by the Indonesian government such as the implementation of Large-Scale Social Restrictions (PSBB) which are enforced in almost every region, as well as regulations regarding the use of disinfectant liquid to be sprayed at the entrance, wearing a mask when outside the house and washing hands before entering the house [1].

The government also enacted new regulations so that everyone can leave the house and do work without having to worry about being exposed to the COVID 19 virus. One of the new regulations made by the governor of DKI Jakarta is that everyone who is going out of the house must wear a mask and also maintain physical distancing. This condition also applies to office workers where every employee is required to wear a mask and practice physical distancing to avoid the spread of the COVID 19 virus. Therefore, the use of masks is very important in carrying out daily activities, especially when leaving the house [2]. Most people still don't understand how to use face masks properly and correctly, there are still many people who use masks without covering their noses. Face masks are important during the pandemic, with many types of masks on the market. Here are some

examples of types of masks, including cloth masks, medical masks, surgical masks, electric masks, KN95 masks, N95 masks, Reusable Facepiece Respirators and Mask Filtering Facepiece Respirator (FFR) [3].

Image processing techniques have been widely used for the recognition of objects that have similar characteristics in the process of identifying images. Many methods and algorithms can be used to identify an image such as using Machine learning (ML), Artificial Neural Network (ANN), and Deep Learning (DL). But in general, artificial intelligence is the mother of all these techniques. Artificial intelligence or Artificial Intelligence (AI) is a technique used to imitate the intelligence possessed by living and inanimate objects to solve a problem. One of the most frequently used branches of AI in identifying images is Machine Learning [4]. Machine Learning (ML) is an algorithm that is embedded into a machine so that it can carry out learning independently, without any assistance from humans. Machine Learning has two basic techniques, namely supervised learning and unsupervised learning. Supervised learning is a learning process carried out under supervision. Whereas unsupervised learning is a learning process that focuses on data exploration, this algorithm does not require a model. The first Supervised Learning method is Support Vector Machine (SVM), which is a method used to classify objects by finding the best hyperplane by maximizing the value of the distance between classes. Hyperplane is a function used to separate classes. The distance between classes is the separation distance between two object classes. Apart from that, another Supervised Learning method is Deep Learning (DL), which is a method that uses an Artificial Neural Network (ANN). Artificial Neural Network uses the principles of the human brain to receive information and forward it into the next process. Learning Vector Quantization (LVQ) is one of the classification methods of Artificial Neural Networks (ANN). Apart from Machine Learning (ML), Deep Learning is also currently the most frequently used method, especially with Convolutional Neural Networks (CNN), namely by classifying several objects with Multilayer Perceptron (MLP) [5].

In this study a mask detection system was designed at the entrance of a room by utilizing a microcontroller in the form of a Raspberry pi as the main controller, a webcam as input which would detect the use of a mask on someone's face in front of the entrance and give orders to the microcontroller to control the relay and solenoid at the entrance. In this study the data used were data from people who wore masks and did not wear masks. The data that has been trained will then be run by Raspberry Pi, to then detect whether the person at the door is wearing a mask or not. If the person is wearing a mask, the door will open so that person can enter the room.

2. Research Method

Image processing is a processing operation for images that aims to improve image quality or extract the required information from processed images, especially by using a computer so that image quality becomes better. In other words, image processing is an image processing and analysis process that involves a lot of visual perception [6]. Image processing is carried out to improve image quality so that it is easy to be interpreted by humans/computers. The input of this image processing is the image itself and the output can also be an image or the characteristics of the image, but with better quality than the input image [7]. Image is a representation, resemblance or imitation of an object or things [8]. The visible image is the light reflected from an object [9]. In a mechanism for using a digital system, the image itself is broken down into two parts, namely the names sampling and quantitation. The sampling mechanism itself is a process of taking fractional or discrete spatial coordinate values, for example: x,y, periodically during the sampling phase. Digitization itself is the process of changing an object, be it an image, text, or sound, from an object that can be examined into electronic data and that data can be stored in a 2-dimensional matrix form.[10]

Deep learning is a branch of machine learning that utilizes artificial neural networks to implement problems with representational learning methods that enable computational models composed of many layers (hidden layers) of nonlinear information processing to perform feature extraction, pattern recognition, and classification of commands input to produce output [11]. The difference with ANN itself is the number of hidden layers in deep learning which are modeled in such a way as to provide more accurate output [12]. Deep learning is capable of performing very accurate pattern recognition such as graphic patterns, handwriting and other patterns so it is very well used for developing applications such as image recognition and classification [13]. One of the Deep Learning methods used for image recognition is the Convolutional Neural Network. This method has a high network

depth and is often applied to image recognition data so that it can produce a high level of accuracy and good results [11]. The Convolutional Neural Network method is very good at finding features in the image to the next layer to form nonlinear equations that can increase the complexity of a model [14].

Convolutional Neural Network (CNN) is a classification method for the type of neural network that is usually used in image data processing which belongs to the deep learning group which is designed to process two-dimensional data with high network depth and is widely applied to image data using a convolution layer to convolve an input using filter[13]. A convolution or commonly called a convolution is a matrix that has the function of filtering images. Each convolution has a different number of filters and kernel sizes [15]. Convolutional Neural Network is one of the methods used for image classification, which was first used by Hubel and Wiesel who were inspired by the visual cortex of the cat's sense of sight. The visual cortex is the part of the brain that functions to process visual information which contains a complex arrangement of cells. This model can reduce a number of independent parameters and can handle input image deformations such as scale, rotation, and translation [16]. CNN is used to analyze visual images, detect and recognize objects in images, which are high-dimensional vectors that involve many parameters, namely several layers that are used to filter each process. The process is called two main stages, namely feature learning and classification. The feature learning stage consists of 3 stages, namely the convolution layer, ReLU (activation function) and the pooling layer, while the classification stage consists of flatten and fully-connected layers [17].

The programming language used in this study is python. This programming language can be used to create image processing and recognition features. This language is packaged with several Libraries which are required for some functions. One of the most powerful and efficient libraries is the Scikit-Learn Machine Learning Library. Apart from image recognition, the library can be used for smarter functions such as facial recognition and motion detection [18]. This language was designed by a man named Guido van Rossum and first appeared in 1991. Development of Python is still being carried out by the Python Software Foundation. almost all operating systems are supported by the Python language, even for the Linux operating system, almost all distributions include Python in it. With a code that is simple and easy to implement, a programmer can prioritize the development of the application he is making, instead of being busy looking for syntax errors [19][20].

2.1. System Design

This research is generally divided into two, the first is the system design stage and the second is the system testing stage. In the first stage there are two stages, namely hardware design and software design. To obtain ideal results and in accordance with what is expected, therefore the design of this system will refer to the theoretical basis that has been obtained from various reliable sources. In the first step is the stage in the design such as block diagrams, and assembly / preparation of tools. In the second stage of design is the design and manufacture of software including flowcharts, programmers, and tool testing. At this stage the system testing is carried out after the tool has been successfully made and tested first. This aims to ensure that the system is made as expected. The steps of system testing start from several stages as follows: hardware testing includes testing of tool components. Software testing includes testing the program and compatibility between hardware and programs that are made. The following is a system hardware design to detect the use of masks on the face can be seen in Fig. 1.

From the data obtained by the camera/webcam which is then sent to the Raspberry Pi to be able to carry out the programmed commands. If the camera detects the use of a mask on someone's face, the relay will activate the solenoid door lock so that the door can be opened, but if the camera detects that someone is not wearing a mask, the speaker will turn on as a warning sign to use a mask. The following is the design of the prototype and schematic of this research series can be seen in Fig. 2.

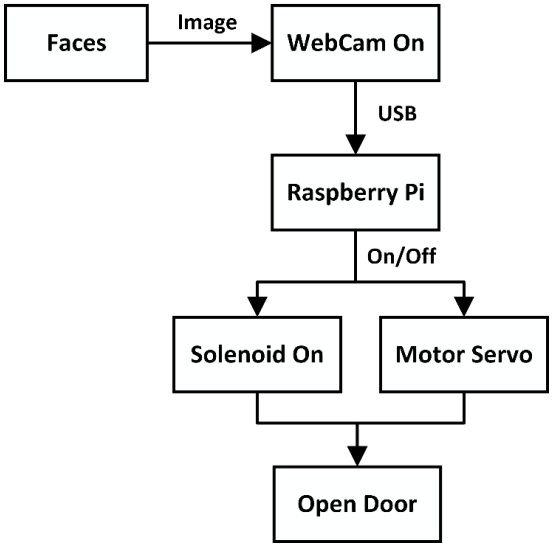

Fig. 1. Block diagram of the mask detection system

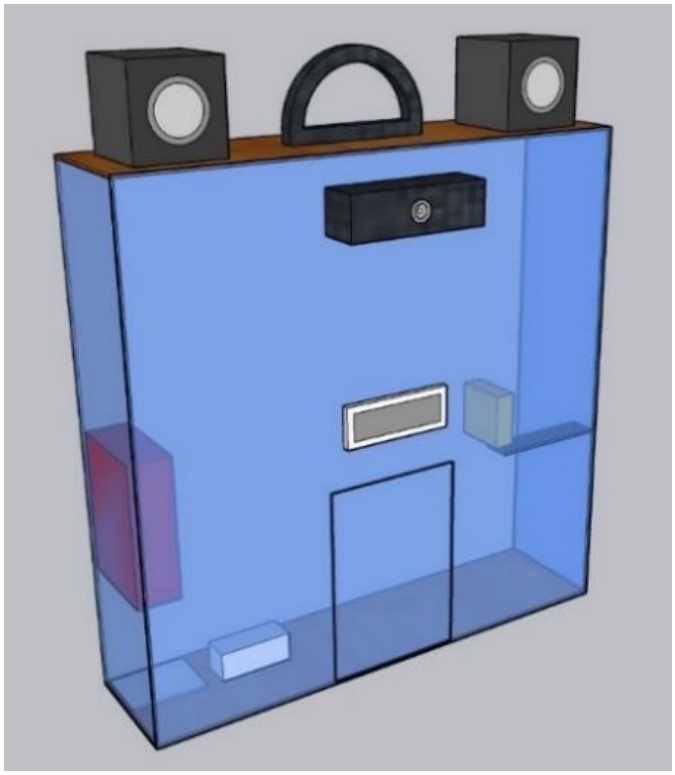

Fig. 2. Mask detector cover design

The system development process using the CNN method is shown in the Fig. 3.

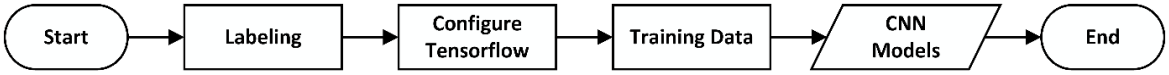

Fig. 3. Block diagram of the training model making process

The Fig. 3 shows that the CNN system development process is divided into several stages that are quite complex. The preparation process requires good accuracy so that configuration errors do not occur before training.

2.2. Flowchart

The design of the software for detecting the use of masks at the entrance to a room is a design for the Raspberry pi program, which can be seen in Fig. 4.

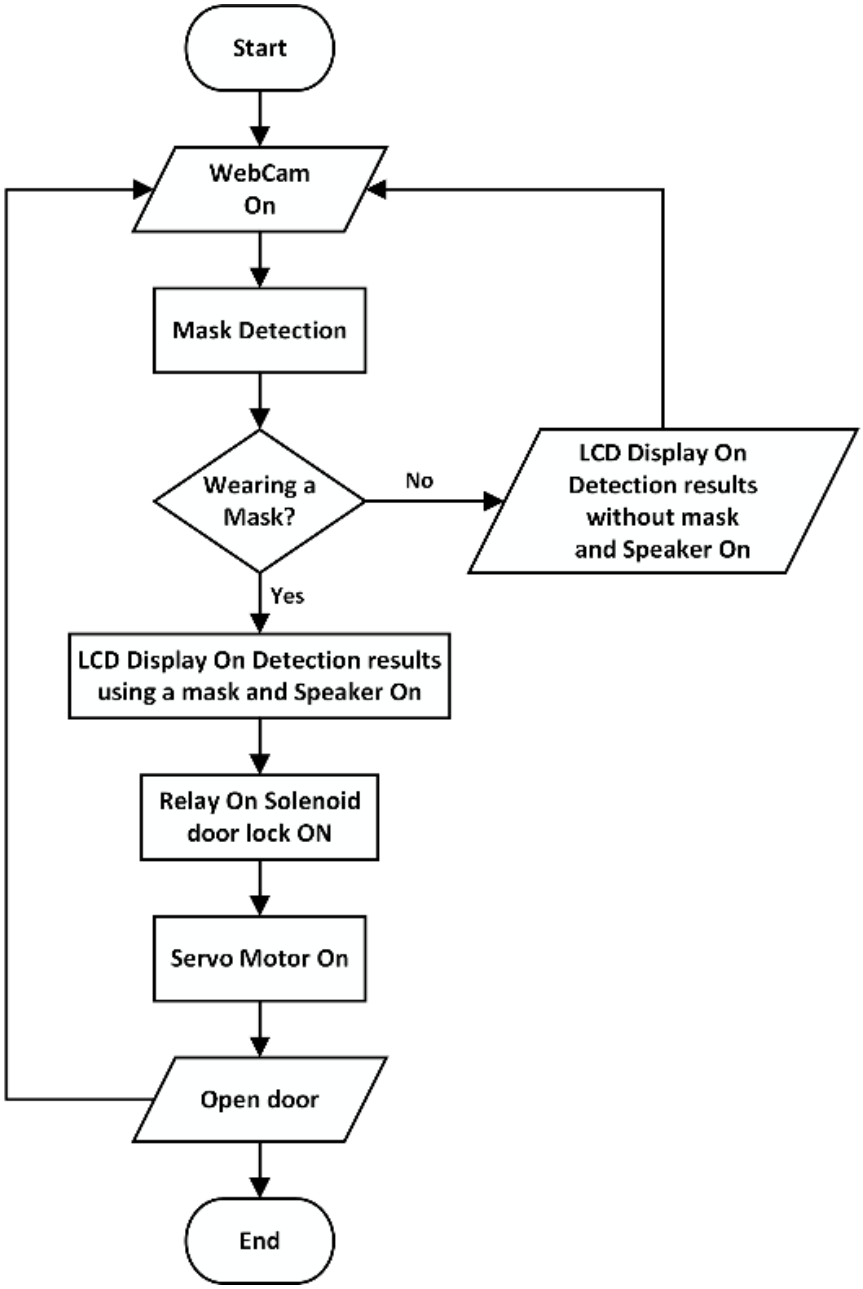

Fig. 4. Flowchart of the mask detection system

In making programs that will be entered into the Raspberry pi using TensorFlow Lite and MobileNetV2. In this software, programming settings will be carried out to include ports to be used as input, output and other features, this can be done directly and written in the programming pages of TensorFlow Lite and MobileNetV2. The following is a flowchart of Raspberry pi programming, which can be seen in Fig. 5.

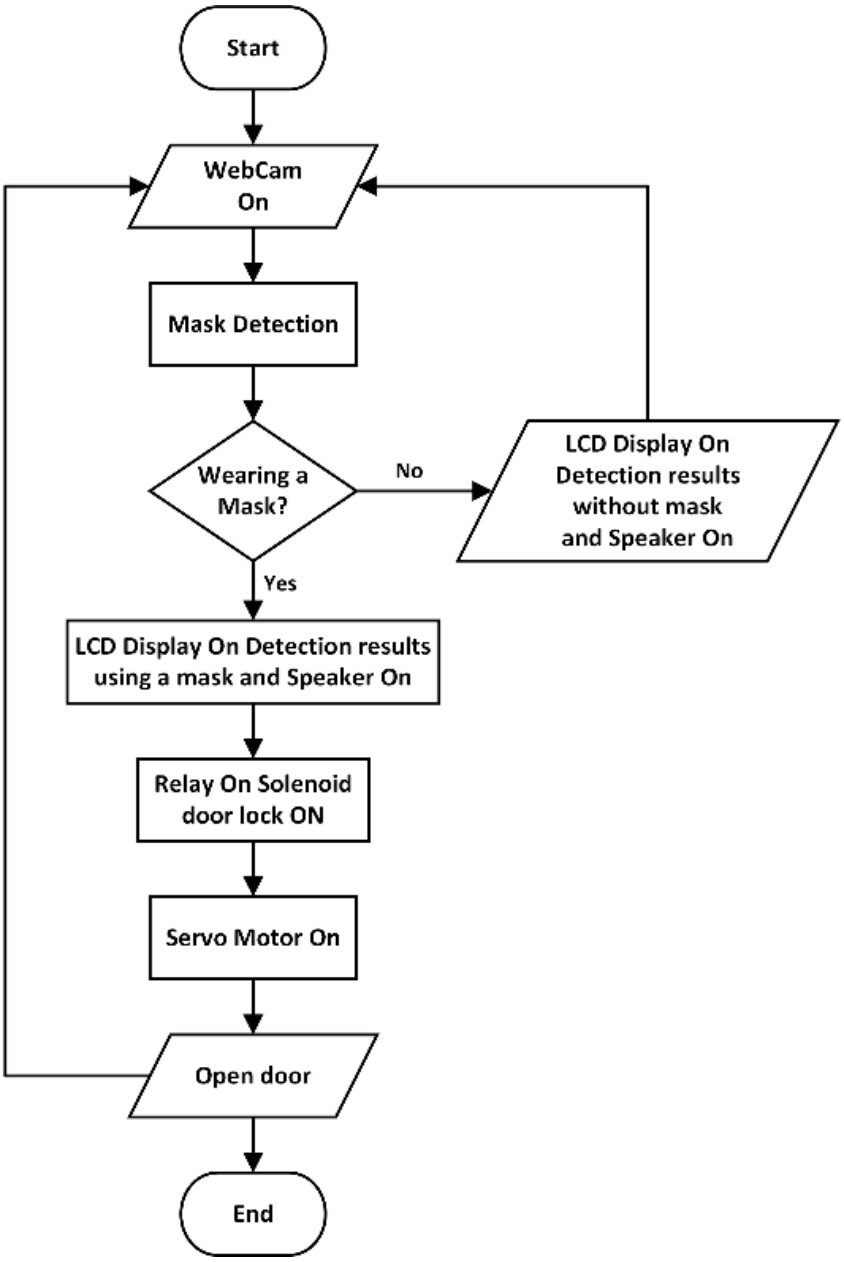

Fig. 5. Flowchart of the mask detection system

In the flowchart Fig. 5, it can be explained that the system will start to activate and run the program to capture video streams on the Webcam, then the results from the Webcam stream are sent to the Raspberry pi which has been programmed in advance so that it can detect whether a person is wearing a mask or not, if the person is not wearing a mask, the words "Detection Results Without a Mask" will appear on the LCD display screen and when the person is wearing a mask, the LCD Display will display "Results of Detection Wearing a Mask" which will then activate the relay so that the solenoid door lock is active so that the door can be opened. using a servo motor to open the door as the output of the system in this study.

This detector circuit scheme is arranged according to the flowchart that was made and is a presentation of the flowchart that is embodied in hardware form for testing based on the program that has been made. Fig. 6 is a schematic of this research series

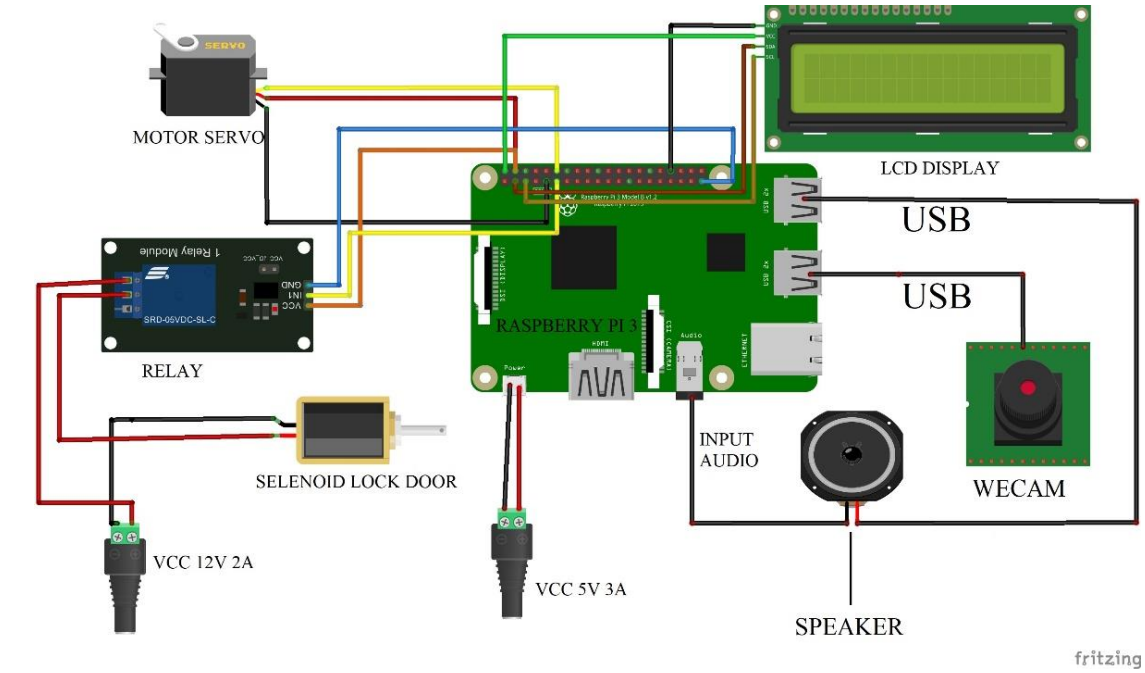

Fig. 6. Schematic of the mask detector network

3. Result and Discussion

This study uses the CNN classification method in which there are several stages, namely pre-processing, training, and testing. In general, the making of the system is as shown in the Fig. 7.

Fig. 7. System Building Block Diagram

The initial process for building a CNN system is preprocessing. In this step, the image data used is first marked with LabelImg.exe. The purpose of labeling is to provide the system with information about the object being used. Labeling is performed on all image data. After the labeling process is complete, the next step is to prepare the model and configuration for the learning process. The training process at CNN uses the TensorFlow framework version 1.15. The results of the learning process will later become a data model that can be used for classification. The results of the classification will be calculated accuracy using Equation (1).

$$Accuracy(\%) = \frac{TP}{Total\ Test\ Data} \times 100\% \quad (1)$$

Where:

TP is the number of classification test data that is correctly predicted.

3.1. Software Testing

The first software test will be carried out on the image database from this study which is divided into two classes, namely the first class in the form of photos of people wearing masks and the second class in the form of photos of people without masks with a total of 2000 photos as image data. Image data is collected based on each class in a folder which is divided into two folders, namely the training and testing folders. The distribution of image data can be seen in the Table 1.

Table 1. Face image database

No	Classification	Amount of data		Number of Image Data
		Training	Testing	
1.	Wear a mask	900	100	1000
2.	No masks	900	100	1000
	Total	1800	200	2000

The training image data in the table above will be subdivided into 2, namely training image data and training image data which are divided into 800 training and 100 validations. This is because there is a validation process during training to measure the accuracy of the model to be produced. The next process is the image labeling pre-processing process using the Labeling.exe software. Image data is displayed into the software, and image objects will be selected using the drag image tool located on the left. Then, the selected image data is named according to its respective class. Do this repeatedly until all of the image data used has been labeled. The marking process is shown in Fig. 8.

Fig. 8. Image labeling process using Lebeling.axe

The results of the labeling process in Fig. 8 will be in .xml format with the data information needed for the training process. After all the annotations are complete, the next step is to convert the .xml data into .csv data. Perform this procedure to combine all the .xml data generated by the tagging process into one file. The results of this conversion process are converted into two files, namely Excel files in .csv format for training data and validation data. The next process is the data training process using the TensorFlow framework version 1.15 and running the train.py program. Training for this method runs automatically and generates a loss value for each completed step. Training is carried out up to 200000 steps or until it has a small loss value. The loss value generated during this training process is 5.3. This loss value is still quite good even though it is more than 1. The smaller the resulting loss value, the better the resulting model. The training process diagram is shown in Fig. 9.

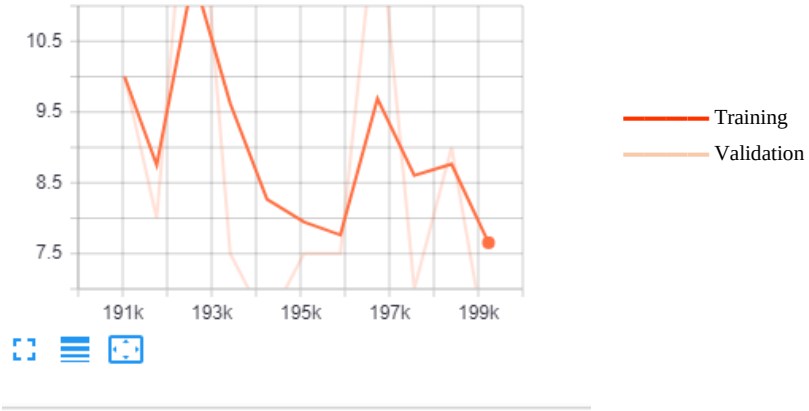

Fig. 9. Graph of loss value during training

From the graph in Fig. 9 the loss value is still not stable, but it can be said to be quite good. This graph is a feature of TensorFlow that displays real-time loss training data. The training process is automatically saved at the checkpoint so that it can be stopped or restarted at any time. After the

training process is complete, the training results are converted into a model. The transition is made to the inference graph process. The result of this transformation is a file called freeze_inference_graph.pb, which is the model of the CNN that was created. This file can be used to classify image use masks. The results of the model that has been made are very large, where all the files in one folder are more than 100MB in size. Data testing is done using 10 new images in each class. The program code used to run the test is named Object_detection_image.py. The program tests the system by entering new images and models that have been created. then the classification process is carried out, as shown in Fig 10 and Fig. 11.

Fig. 10. Results of image classification of wearing masks

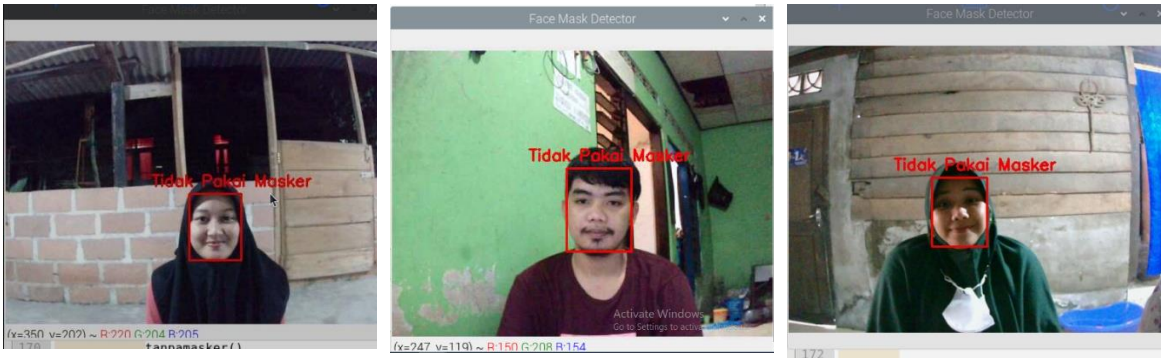

Fig. 11. Image classification results not wearing masks

Based on Fig. 10 and Fig. 11, classified images are marked with boxes and class names. The database used in this study is 2000 data. The data is divided into 80% for training data and 20% for the model training process. While the data for system testing is 200 image data, the data is divided into 100 image data using masks and 100 image data without using face masks on 5 men and 5 women. The results of these data can be seen in Table 2.

Table 2. Test Data Results

	No	Image Name	PREDICTION		Total
			Wear a Mask	Wear a Mask	
ACTUAL	1.	Wear a Mask	90	10	90
	2.	No Mask	0	100	110
		Total	90	110	200

Based on the test results in Table 2, the system can correctly predict 190 times from the 200 data tested. The test results that have been carried out using the CNN method produce an accuracy of 95%. Based on these results, it appears that the resulting model is good and suitable for the classification process of image recognition using masks. However, to produce a better model, data with more variations and a larger amount of data are needed.

4. Conclusion

Based on the results of testing the data that has been obtained, it can be concluded that the system has successfully detected masks on a person's face automatically using the Convolutional Neural Network method which is done offline by entering 200 facial image data for 5 men and 5 women, the system has successfully detected 190 times out of 200 image data tested with an accuracy of 95%.

References

[1] J. Wang *et al.*, "Prevention and control of COVID-19 in nursing homes, orphanages, and prisons," *Environmental Pollution*, vol. 266, p. 115161, 2020.

[2] C. Eisenmann, C. Nobis, V. Kolarova, B. Lenz and C. Winkler, "Transport mode use during the COVID-19 lockdown period in Germany: The car became more important, public transport lost ground," *Transport policy*, vol. 103, pp. 60-67, 2021.

[3] C. J. Kähler and R. Hain, "Fundamental protective mechanisms of face masks against droplet infections," *Journal of aerosol science*, vol. 148, p. 105617, 2020.

[4] G. Z. Karimova and V. P. Goby, "The adaptation of anthropomorphism and archetypes for marketing artificial intelligence," *Journal of Consumer Marketing*, vol. 38, no. 2, pp. 229-238, 2021.

[5] D. Goularas and S. Kamis, "Evaluation of Deep Learning Techniques in Sentiment Analysis from Twitter Data," *2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML)*, Istanbul, Turkey, 2019, pp. 12-17, doi: 10.1109/Deep-ML.2019.00011.

[6] J. Yang, C. Wang, B. Jiang, H. Song and Q. Meng, "Visual Perception Enabled Industry Intelligence: State of the Art, Challenges and Prospects," in *IEEE Transactions on Industrial Informatics*, vol. 17, no. 3, pp. 2204-2219, March 2021, doi: 10.1109/TII.2020.2998818.

[7] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara and D. Rueckert, "Self-supervised learning for medical image analysis using image context restoration," *Medical image analysis*, vol. 58, p. 101539, 2019.

[8] T. Hu, C. Liang, G. Min, K. Li and C. Xiao, "Generating video animation from single still image in social media based on intelligent computing," *Journal of Visual Communication and Image Representation*, vol. 71, p. 102812, 2020.

[9] W. Wang, X. Wu, X. Yuan and Z. Gao, "An Experiment-Based Review of Low-Light Image Enhancement Methods," in *IEEE Access*, vol. 8, pp. 87884-87917, 2020, doi: 10.1109/ACCESS.2020.2992749.

[10] O. T. Tawose, B. Li, L. Yang, F. Yan and D. Zhao, "Topological Modeling and Parallelization of Multidimensional Data on Microelectrode Arrays," *2022 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, Lyon, France, 2022, pp. 795-805, doi: 10.1109/IPDPS53621.2022.00082.

[11] P. Wang, E. Fan and P. Wang, "Comparative analysis of image classification algorithms based on traditional machine learning and deep learning," *Pattern Recognition Letters*, vol. 141, 61-67, 2021.

[12] K. Amasyali anad N. El-Gohary, "Machine learning for occupant-behavior-sensitive cooling energy consumption prediction in office buildings," *Renewable and Sustainable Energy Reviews*, vol. 142, p. 110714, 2021.

[13] M. A. Iqbal Hussain, B. Khan, Z. Wang and S. Ding, "Woven fabric pattern recognition and classification based on deep convolutional neural networks," *Electronics*, vol. 9, no. 6, p. 1048, 2020.

[14] Y. Liu, H. Pu and D. W. Sun, "Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices," *Trends in Food Science & Technology*, vol. 113, pp. 193-204, 2021.

[15] A. Agrawal and N. Mittal, "Using CNN for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy," *The Visual Computer*, vol. 36, no. 2, pp. 405-412, 2020.

[16] X. Bai, X. Zhou, F. Zhang, L. Wang, R. Xue and F. Zhou, "Robust Pol-ISAR Target Recognition Based on ST-MC-DCNN," in *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 12, pp. 9912-9927, Dec. 2019, doi: 10.1109/TGRS.2019.2930112.

[17] R. S. Latha *et al*., "Automatic Detection of Tea Leaf Diseases using Deep Convolution Neural Network," *2021 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2021, pp. 1-6, doi: 10.1109/ICCCI50826.2021.9402225.

[18] P. Elechi, "Facial Recognition Based Smart Door Lock System," *FUPRE Journal of Scientific and Industrial Research (FJSIR)*, vol. 6, no. 2, pp. 95-105, 2022.

[19] B. S. Chowdary, V. N. Subhadra and S. Kavitha, "Age and Gender Detection to Detect the Manipulated Images using CNN," *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, 2023, pp. 1187-1192, doi: 10.1109/ICSSIT55814.2023.10060905.

[20] I. Neforawati and D. Arnaldy, "Message Digest 5 (MD-5) Decryption Application using Python-Based Dictionary Attack Technique," *2021 4th International Conference of Computer and Informatics Engineering (IC2IE)*, Depok, Indonesia, 2021, pp. 424-428, doi: 10.1109/IC2IE53219.2021.9649020.