Detection of Mask Usage Using Image Processing and Convolutional Neural Network (CNN) Methods

Ervan Asri[1], Yance Sonatha[2]*✉, Indri Rahmayuni[3], Meri Azmi[4]
[1,2,3,4]Information Technology Department, Politeknik Negeri Padang, Indonesia
Corresponding Author: yance@pnp.ac.id

Article Information	ABSTRACT
	Since the first instance of COVID-19 was detected in Indonesia in March 2020, the number of patients who have been positively infected has been growing by the day. On April 13, 2021, 1,571,824 patients were positively contaminated with COVID-19. Various efforts have been made to suppress the spread of this dangerous virus, one of which is by implementing health protocols that must be obeyed when someone is in a crowd. The recommendation to use masks is a form of health protocol. The research was conducted to develop a system that can detect the proper use of health masks. The system was created utilizing a deep learning and convolutional neural network algorithm (CNN). CNN will categorize data items in photograph images and assess whether or not they are appropriately wearing masks. The audio of system will be supplied via speakers situated next to the camera that will scan the user's face.

1. INTRODUCTION

The World Health Organization (WHO) classified COVID-19 as a worldwide pandemic in mid-March 2020. The virus, which originated in the Wuhan area of China, spread fast to 123 nations in just three months[1]. The first positive case report in Indonesia was announced in March 2020. Since these findings, the addition of positive COVID-19 cases has been increasing every day[2]. It was recorded that on January 26, 2022, 4,301,193 cases of patients were confirmed positive for COVID-19. Of these cases, 96% or around 4,127,662 patients have been declared cured, and 3.4% or approximately 144,254 people have died[3].

To avoid the spread of the COVID-19 case, the government has implemented a large-scale social restriction program in various districts, limiting citizens' movement. However, the virus's transmission is still on the rise[4]. In addition, the government is increasingly encouraging the government to comply with health protocols to raise public awareness in protecting themselves to prevent transmission and break the chain of the spread of the virus [5][6].

Various health protocols have often been promoted through various media, such as reducing visits to crowded, narrow, and closed places, maintaining social distance, always using face masks, and diligently washing hands properly[7]. Instructions on properly washing hands and using the correct mask are everywhere. But there are still many people who do not heed it. For example, in using a face mask, many still see it but do not meet the specified requirements, such as covering the nose and chin[8].

This research departs from the concern that many people have not adhered to the proper use of masks as recommended. Advances in information technology are snowballing, including the development and processing of digital information. A system with digital image classification processing is needed to provide the best solution using technology to overcome these problems. Image classification aims to classify images into specific categories[9]. Image classification is currently one of the problems that have long sought a solution in computer vision[10], how to replicate the human ability to comprehend digital picture information such that computers, like humans, can detect items in the image. The feature engineering technique is often quite constrained, as it can only be applied to specific datasets and has no generalizability[11].

Because of its enormous skills in modeling complicated data such as pictures and sounds, Deep Learning has recently become one of the hottest subjects in the area of Machine Learning. The CNN is the Deep Learning approach that is presently producing the most significant results in image recognition (CNN)[8][9], also frequently utilized with picture data. Essentially, CNN is regarded as one of the finest models for solving object detection and identification issues [13]. On some datasets, the CNN study showed digital image recognition with accuracy equivalent to humans in 2012. CNN typically leverages the convolution method by applying a specified size convolution kernel (filter) on an image. By multiplying that area of the image with the filter, the computer acquires new representative information [15].

2. RESEARCH METHOD

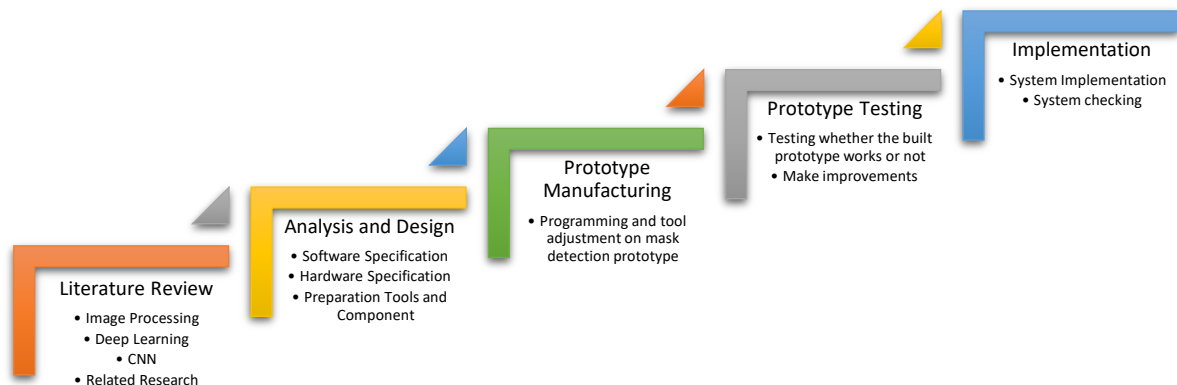Figure 1 below shows the steps taken in completing this research.

Figure 1. Research Methodology

a. The literature review is carried out by reading and understanding both books, websites, and scientific works related to this research.

b. In this analysis and design stage, it is explained about the software and hardware needed and the preparation of tools for prototype development

c. The prototype manufacturing stage will produce a system prototype that will be tested

d. The testing phase of the mask detection prototype aims to determine whether the tool built is functioning or not. If there are problems, repairs are carried out as early as possible

e. If the prototype has been tested, the next step is to implement the actual conditions and check whether the system output is the expected target.

3. RESULTS AND DISCUSSION
3.1. System Design
3.2.1. Application Block Diagram Design

Numerous components in the system architecture assist the attainment of a device where the camera as input is then processed using a laptop and provides an output shown on the laptop monitor. Figure 2 depicts the primary block diagram design:
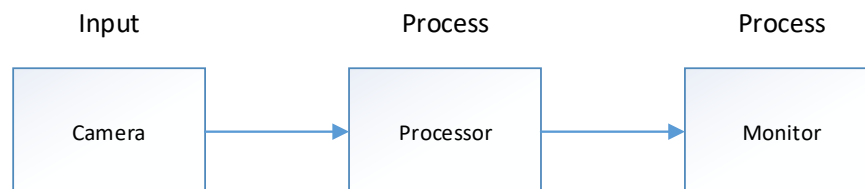

Figure 2. Application Block Diagram

3.2.2. Flowchart Model Convolutional Neural Network

After gathering the data, the CNN model must be trained. In general, CNN comprises two stages: feature learning and classification. The CNN model's image input utilizes a 64x64x3 picture. The number three represents a picture with three channels: Red, Green, and Blue (RGB). The input image will then be processed first through the convolution and pooling processes at the Feature Learning stages. The complete flowchart can be seen in Figure 3 below:
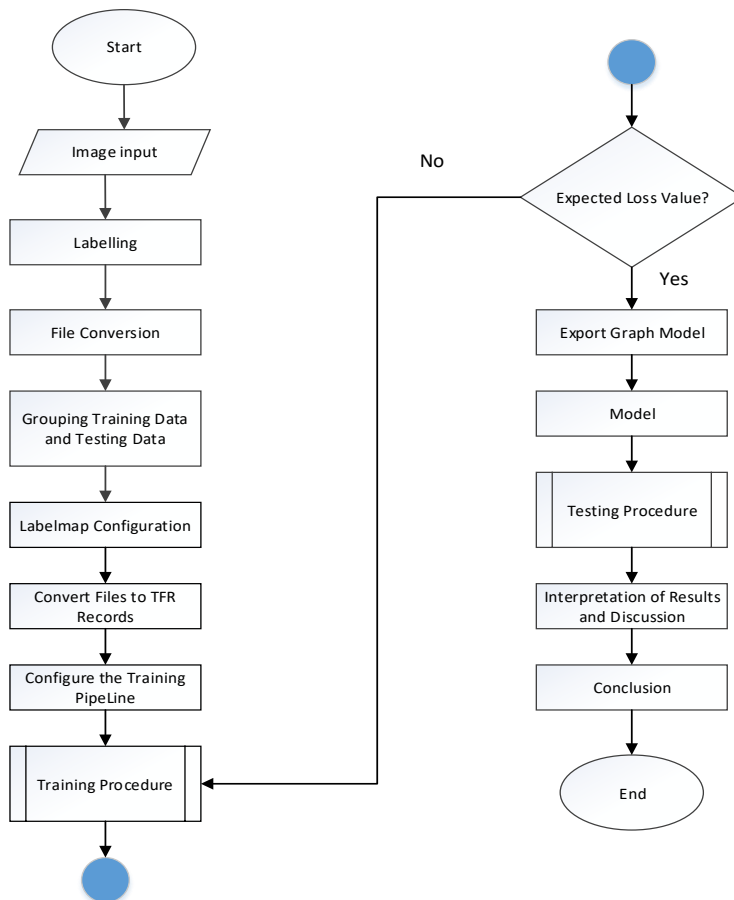

Figure 3. Flowchart of Object Detection with CNN

3.2. System Implementation
3.2.1. Data Set

A data set is a collection of objects representing data and their relationships in storage media. There are various ways to describe datasets, such as attributes used to define types of things, both qualitative and quantitative. This research uses a public dataset, which

means using a dataset from a public repository that has been approved by previous researchers and can be accessed globally[15].

The data utilized to evaluate the mask detection system is two-dimensional picture data. For this experiment, 1484 images were used, which were separated into training data and testing data. Training data is used to train machine learning; in this example, the CNN method employed is consistent with the predicted predictions. At the same time, data testing is used to test the accuracy and performance of machine learning.

The dataset used is a public type created by a researcher named Prajna Bhandary. The dataset can be accessed at https://github.com/prajnasb/observations. The dataset consists of 1484 images and is classified into 2 classes, namely the wear class and the class without a mask. Table 1 shows the amount of data for each class for training data, and Table 2 shows the amount of data for each class for data testing

Table 1. Details of Total Training Data

Class	Amount
With mask	703
Without mask	557
Total	1260

Table 2. Details of Total Testing Data

Class	Amount
With mask	125
Without mask	99
Total	224

Figure 4 below is an example of a dataset used for the data training process. The data set is divided into 2, namely images with masks and without masks. The dataset has been labeled data using the LabelImg module in Python.
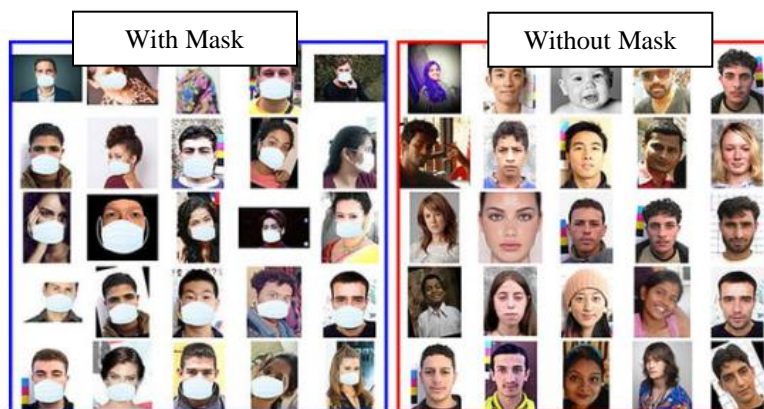
Figure 4. Sample Dataset for Training Data

3.3. System Testing
3.3.1. Testing Procedure

Tests are carried out on training datasets, testing images in real-time to determine the accuracy produced and the results through analysis. Figure 5 is an overview of the system testing scheme.
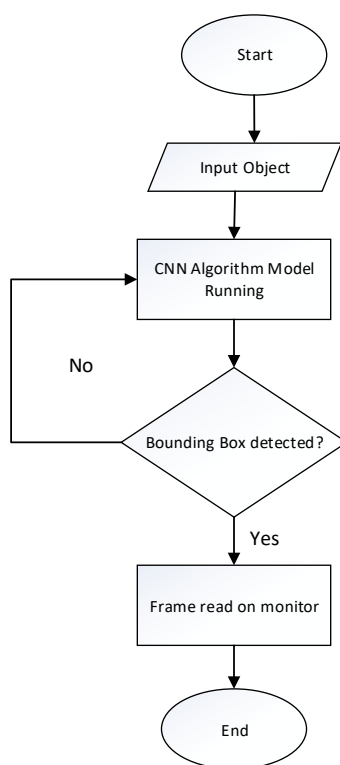

Figure 5. Test Scheme

3.3.2. Mask Detection System

At this training stage, it was using pre-trained CNN. The stage that will be carried out is training for the face dataset wearing a mask and not in each class with pre-trained CNN. In the data, 1484 images have been divided to test the accuracy of the dataset. The first process is dividing the facial dataset into train and test data. The following process is training using a network that has been trained or previously trained (pretrained) to determine the accuracy of the dataset using transfer learning.

The introduction of the use of face masks is being tested, with the categories of wearing and wearing a mask. The technique is carried out by taking an image from the camera's video. Figure 6 depicts an example of facial recognition findings.

Figure 6. Mask Detection Results

From this test, it was concluded that the system developed was not only able to recognize the use of masks for one face object but could be more than that.

3.3.3. Determination of Test Parameters

The parameters for testing in this CNN algorithm model are determined to produce the best system. The parameters used are the value of the accuracy per class, accuracy per epoch, and loss per epoch value.

3.3.4. Testing Results

The following are the results of several test parameters tested on 1260 image data consisting of 703 class data for wearing masks and 557 classes without masks. The results are as follows:

1. Accuracy per class
    Accuracy is defined as the ratio of correct predictions (both positive and negative) to total data, This is stated as follows:
    $$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad (1)$$

    Where :
    TP = True Positive
    TN = True Negative
    FP = False Positive
    FN = False Negative

    Based on testing on 224 testing data consisting of 125 mask classes and 99 data without masks, the accuracy value is 1, which means the system recognizes all

images and produces output according to predictions, as shown in Table 3 below.

Table 3. Accuracy Results per Class

Class	Number of samples	Accuracy
With mask	125	1.00
Without mask	99	1.00

2. Accuracy per epoch

The epoch parameter specifies how many times the algorithm will iterate through the full dataset. A one-time epoch occurs when all datasets are put through a single forward and backward method across all neural network nodes at the same time. The amount of training datasets utilized in one batch is called batch size. In the Neural Network training phase, one Epoch is too long since all data is incorporated; hence, it will take a long time. Typically, the dataset is separated into batches to simplify and speed up the training process. The learning rate is one of the training parameters used to determine the weight correction value during the training process. In this test, the epoch value is 50, the batch size is 16, and the learning rate is 0.001. Figure 7 depicts the accuracy value for each epoch:
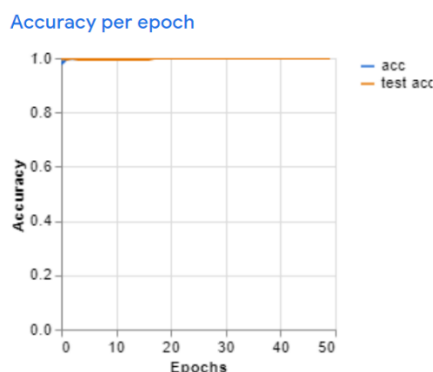
Figure 7. Accurate value per epoch

Figure 7 shows that the training data accuracy for each epoch is very high, close to or equal to 1.00, with a learning rate of 0.001. The testing data shows similar findings.

3. Loss per epoch

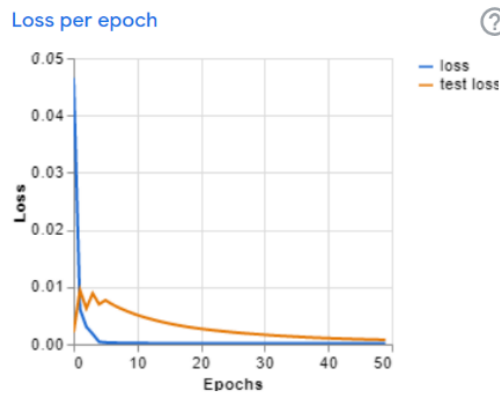With the same epoch value and batch size, the data loss value is sought, and the following Figure 8 is obtained:

Figure 8. Loss per epoch value

Figure 8 indicates that by epoch 10, the loss value for the training data (blue line) is already zero. A loss value test is also performed on the testing data to prevent overfitting. Overfitting happens when a model is developed that is overly focused on a specific training dataset, such that it cannot generate accurate predictions when given other comparable datasets.

From the same picture seen for the testing data (yellow line), the loss value decreases as the epoch value increases and finally close to 0 at the epoch value of 50. This proves that the system works well (good fit) on the testing data.

4. CONCLUSION

The CNN approach is employed in this study utilizing 1484 image data, divided into 1260 training data and 224 testing data, and divided into two classes, wearing and not wearing a mask. The developed system model is capable of identifying and distinguishing several face objects from an image/video source displayed on the camera screen. Each object can be recognized and grouped based on 2 predefined classes. From the results of training data processing, an accuracy value of 1 means that all predictions are by the actual. In other words, the system can recognize all images and correctly determine the class of the picture. The accuracy value per epoch on the training data is close to or equal to 1.00 at the time of epoch 10 and testing on the testing data. In the training data, at epoch 10 the loss value is already 0. While in the testing data, there is a slight difference because the loss value 0 was only obtained at epoch 50, it is not much different because the loss value is almost close to the value 0, in the sense that the developed model is already in good fit condition and not included in the underfitting or overfitting category.

REFERENCES

[1] World Health Organization, "Looking back at a year that changed the world: WHO's response to COVID-19," no. January, p. 66, 2021.

[2] S. Lin, R. Kantor, and E. Clark, "Coronavirus Disease 2019," *Clinics in Geriatric Medicine*, vol. 37, no. 4, pp. 509–522, 2021, DOI: 10.1016/j.cger.2021.05.001.

[3] Kemenkes RI, "COVID 19 di Indonesia." https://covid19.kemkes.go.id/ (accessed Jan. 27, 2022).

[4] A. D. Samala, F. Ranuharja, I. P. Dewi, R. Watrianthos, and Y. Indarta, "Learn Algorithm and Programming in Higher Education Using E-Learning During the COVID-19 Pandemic: Students' Perspective," *Adv. Soc. Sci. Educ. Humanit. Res.*, vol. 608, no. ICTVET, pp. 97–102, 2021.

[5] P. Wu, H. Li, N. Zeng, and F. Li, "FMD-Yolo: An efficient face mask detection method for COVID-19 prevention and control in public," *Image and Vision Computing*, vol. 117, Jan. 2022, DOI: 10.1016/j.imavis.2021.104341.

[6] S. Sethi, M. Kathuria, and T. Kaushik, "Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread," *Journal of Biomedical Informatics*, vol. 120, Aug. 2021, DOI: 10.1016/j.jbi.2021.103848.

[7] B. Varshini, H. Yogesh, S. D. Pasha, M. Suhail, V. Madhumitha, and A. Sasi, "IoT-Enabled smart doors for monitoring body temperature and face mask detection," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 246–254, Nov. 2021, DOI: 10.1016/j.gltp.2021.08.071.

[8] M. M. Lambacing and Ferdiansyah, "Rancang Bangun New Normal Covid 19 Masker Detektor Dengan Notifikasi Telegram Berbasis Internet of Things," *Jurnal DINAMIK*, vol. 25, no. 2, pp. 77–84, 2020.

[9] N. Sharma, V. Jain, and A. Mishra, "An Analysis of Convolutional Neural Networks for Image Classification," in *Procedia Computer Science*, 2018, vol. 132, pp. 377–384. DOI: 10.1016/j.procs.2018.05.198.

[10] A. Y. Wicaksono, N. Suciati, C. Fatichah, K. Uchimura, and G. Koutaki, "Modified Convolutional Neural Network Architecture for Batik Motif Image Classification," 2017.

[11] Prof. B. S.P., S. Kotambkar, A. Lakade, K. Mande, and T. Deshmukh, "Parking Management System using Image Processing and Distributed Approach," *IJARCCE*, vol. 6, no. 3, pp. 584–586, Mar. 2017, DOI: 10.17148/ijarcce.2017.63136.

[12] L. Deng and D. Yu, "Deep learning: Methods and applications," *Foundations and Trends in Signal Processing*, vol. 7, no. 3–4. Now Publishers Inc, pp. 197–387, 2013. DOI: 10.1561/2000000039.

[13] Y. Lu, "Deep neural networks and fraud detection," Uppsala, Sweden, Oct. 2017.

[14] I. W. Suartika E.P, A. Y. Wijaya, and R. Soelaiman, "Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) pada Caltech 101," *Teknik ITS*, vol. 5, no. 1, pp. A65–A69, 2016.

[15] Ö. İnïk, M. Altıok, E. Ülker, and B. Koçer, "MODE-CNN: A fast converging multi-objective optimization algorithm for CNN-based models," *Applied Soft Computing*, vol. 109, Sep. 2021, DOI: 10.1016/j.asoc.2021.107582.

[16] Darmatasia, "Deteksi Penggunaan Masker Menggunakan Xception Transfer Learning," *INSTEK - Informatika Sains dan Teknologi*, vol. 5, no. 2, pp. 279–288, 2020.