Identification of Proper Use of Masks Based on Mouth and Nose Detection

Denny Hardiyanto[1], Ihtiari Prastyaningrum[2], Umi Kholifah[3], Dyah Anggun Sartika[4]

***Abstract*—During the pandemic, the proper use of face masks is very critical to reduce and prevent the transmission of the COVID-19 between people. Unfortunately, many people are still careless about this proper use of masks like using masks to cover only their mouth or chin. There are also people who do not wear masks when traveling or interacting. Such conducts indicate the lack of concern for the mask use. This study aims to obtain an identification algorithm for using the face mask appropriately through digital photos/images. The basic algorithm used was the face, nose, and mouth detection algorithm developed by Viola and Jones. These algorithms were then combined so that they formed a strong algorithm for detecting the proper use of the face mask. The data tested were classified into five categories, namely images of proper use of masks, images of masks with visible noses, images of masks with visible mouths, images of faces with masks worn on the chin, and mixed images with various accessories. Results of the study employed sixty testing images with various variations of attributes, the result obtained an accuracy value of 90%, a sensitivity value of 100%, and a specificity value of 62.5%. The low specificity value was caused by many detection errors in the false positive (FP) attribute, meaning that the system can detect objects other than the mouth and nose. This research is expected to be developed and synergized with other applications so that it can raise public awareness about the proper use of masks.**

***Keywords*—Mask Identification, COVID-19, Viola-Jones, Face Detection.**

I. INTRODUCTION

The COVID-19 is a type of disease caused by one of the coronavirus types, namely SARS-CoV-2 which is also known as Corona. This virus can cause respiratory system disorders, ranging from mild to severe symptoms, namely lung infections like pneumonia. In severe infections, this disease may even cause death. According to data compiled by the COVID-19 Handling Task Force of the Republic of Indonesia, the number of confirmed positive cases as of March 25, 2021, was 1,476,452 people, with a death toll of 39,983 people. Meanwhile, the case fatality rate caused by COVID-19 was around 2.7% [1]. The COVID-19 case was first detected in Wuhan, China. This virus transmits quickly from person-to-person and is currently spreading in most countries, including Indonesia. On March 2, 2020, the Indonesian government announced the first two cases of COVID-19 positive patients in Indonesia [2]. Indonesia has been battling the COVID-19 outbreak for more than two years now. To combat the COVID-19, the government has issued a policy to enforce the 3M health protocol, which is now 5M, including wearing masks, maintaining distance, washing hands, staying away from crowds, and reducing mobility [3]. This health protocol must be adhered to by all levels of society to reduce and prevent the transmission of the virus.

Fig. 1 Masks are not used properly (source: Republika).

During the pandemic, the appropriate use of face masks is very critical to reduce and prevent the transmission of the COVID-19 between people. Coronavirus has the potential to spread through droplets [4]. Unfortunately, a prevalent phenomenon in the community is that there are still many people who are careless about the proper use of masks, even though they are crucial for maintaining health and preventing the COVID-19 transmission during the ongoing pandemic. Examples of the improper use of face masks include masks that are used only to cover the mouths, chins, or nostrils, as illustrated in Fig. 1. Even worse, there are still many people who do not bother to wear masks in the crowd [5].

This phenomenon serves as the background underlying researchers to build an identification system for the proper use of face masks to raise public awareness regarding the necessity of using masks during the pandemic. To achieve this goal, the researchers carried out the process of developing face mask detection algorithms based on references from earlier studies and developed them to improve the accuracy of the proposed detection process so that it can detect whether the masks are worn appropriately. After that, researchers tested the system that had been developed before being socialized and introduced

[1,2,3] *Department of Electrical Engineering Education, Faculty of Teacher Training and Education, Universitas PGRI Madiun, Auri Street No. 14-16, Kanigoro, Kec. Kartoharjo, Madiun, INDONESIA (email: [1]denny.hardiyanto@unipma.ac.id, [2]ihtiari.prastya@unipma.ac.id, [3]umikholifah@unipma.ac.id)*

[4] *Department of Computer Control Engineering, Politeknik Negeri Madiun, Serayu Street No. 84, Pandean, Kec. Taman, Madiun, Jawa Timur, INDONESIA (email: dyahanggun@pnm.ac.id)*

[Received: 25 October 2021, Revised: 27 August 2022]

to the public through scientific forums so that the algorithm that had been developed could be used immediately in the community.

II. Development of the Mask Detection Algorithm

This research is based on several important studies, namely the research on face detection which was first introduced by Viola and Jones in 2001 [6]. Then, it was developed over several years later [7]. Furthermore, this research can be further developed as it has many benefits and applications, for example face detection for security systems, safety systems, recognition systems, presence systems [8], and other applications.

In the next few years, face detection research will be developed along with the development of supporting applications such as android applications developed by smartphone vendors. Today, smartphone technology has used many faces detection features for security and camera systems [9]. At the same time, the development of face detection algorithms is getting more complex, which is followed by facial recognition algorithms [10], eye detection algorithms [11], nose and mouth detection algorithms [12], even to the point where multiple algorithms are combined to create an algorithm to detect human facial expressions through digital images and videos (moving images) [13].

Previous research on mask detection was the development of face detection through video analysis [14]. The study focused on masked face detection using four different steps of estimating distance from the camera, eye line detection, facial part detection, and eye detection. In this research, principles used in each of these steps and usages of generally available human detection and face detection algorithms were outlined. Further research on the application of YOLO in detecting masks was conducted [15]. This study focused on mask detection using the YOLO model to increase the processing speed of face mask detection in real time. Then, [16] focused on removing the mask object and identifying the face. This research is quite interesting because it can identify and recognize the faces of mask users. After that, another research on the detection framework to determine the use of masks on the face was conducted [17]. This study focused on detecting the conditions of proper use of masks using the region-based convolutional neural network (R-CNN) algorithm. Reference [18] conducted research on face mask recognition using the convolutional neural network (CNN) algorithm. This research is quite interesting because it focuses on facial recognition when using medical masks or masks. Face identification is very useful for revealing a person's identity through CCTV cameras when committing a crime. The methods used included multi-task cascaded convolutional neural networks (MTCNN), Google FaceNet, and support vector machine (SVM). Another research studied the automatic masked face detection using the MobileNet algorithm and global pooling block [19]. Meanwhile in [20], a software-based masked face detection system was developed, and the detection data could be stored in a cloud system. The developed system was also equipped with personal identity features such as gender and age of the mask wearer [20]. Then, in [21], masked and unmasked faces were detected using the principal component analysis (PCA) algorithm. PCA is an algorithm that is frequently used in facial identification and expressions as it has a high accuracy.

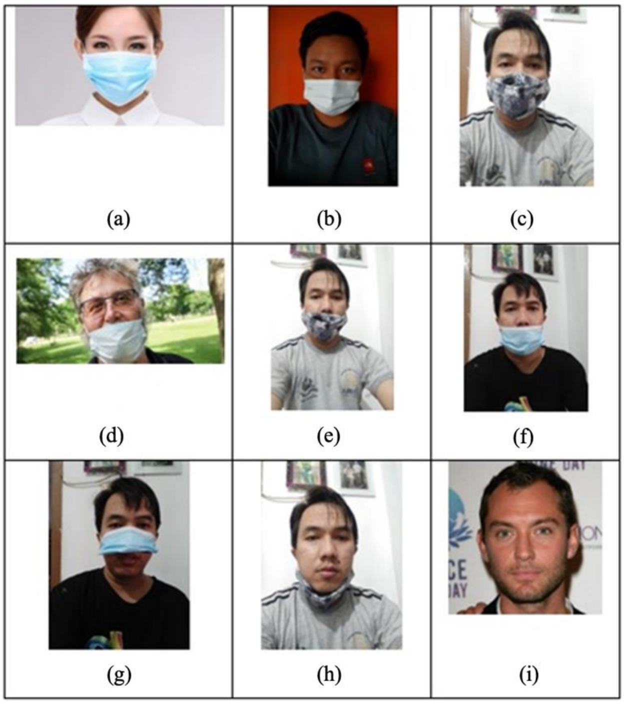

Fig. 2 Research image samples.

Based on research previously mentioned, it can be known that there is a gap that has not been explored by many researchers, namely developing algorithms to detect the proper use of the face mask. For that reason, this research attempts to improve the accuracy of previous research on the proper use of face masks. The results of the research can be useful, especially in dealing with the current pandemic since the proper use of masks is critical to prevent the transmission of the COVID-19.

III. Methods

A. Dataset

The image used in this study is focused on testing images. On the other hand, training images are not used because this research used the Viola-Jones algorithm which has trained more than 10,000 images of body parts such the face, nose, mouth, and eye in a detection system which is then stored in an image processing database and can be used for research and development.

The testing images used in this research were limited only to facial images with the face category facing forward (frontal face) and face testing was not carried out from other points of view. The testing images used in this study consisted of sixty images which were separated into several groups.

- Ten images of faces wearing masks correctly.
- Ten images of faces wearing masks with visible noses.
- Ten images of faces wearing masks with visible mouths
- Ten mixed images containing images of faces wearing masks placed on the neck/chin and images of unmasked people (nose and mouth visible).

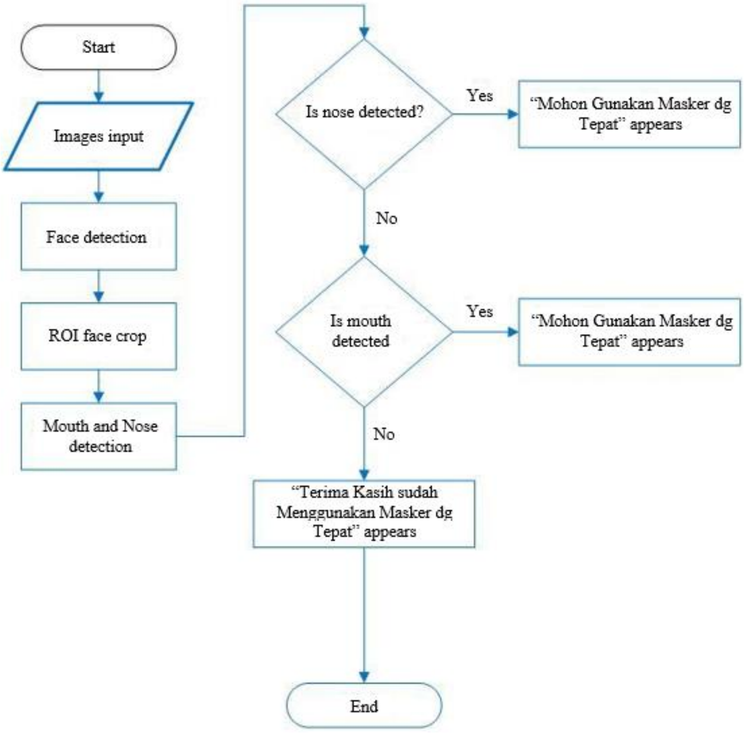

Fig. 3 System flowchart.

- Twenty mixed images (consisting of various conditions, such as various variants of masks and users wearing other attributes).

The image samples used in this study are shown in Fig. 2. Fig. 2(a) to Fig. 2(c) display a proper wearing of masks. Fig. 2(d) to Fig. 2(f) are images of faces wearing masks with visible noses. In Fig. 2(g), only the mouth is visible. Then, Fig. 2(h) shows the image in which a mask is placed in the neck/chin so that the mouth and nose are visible. Last, Fig. 2(i) shows an unmasked face.

B. Research Algorithms

1) Developed Algorithm: The algorithm that is built or embedded in the identification system is described in Fig. 3. The input images were sixty testing image data that had been grouped into five categories. The first stage was for images that underwent the face detection stage. If there was a face object in an input image, then the region of interest (ROI) was sought and the face object was cropped, allowing the object to be detected could focus more on the face.

The next stage was the nose detection. If a nose object was detected, the system marked the nose object with a yellow box and displayed a warning "Mohon Gunakan Masker dg Tepat" (please use the mask properly). However, if no nose object was detected, the stage proceeded to mouth detection. If the system detected the presence of a mouth, the system marked the mouth object and displayed a similar warning as the previous, "Mohon Gunakan Masker dg Tepat". When the mouth and nose were not detected, the system displayed the "Terima Kasih sudah Menggunakan Masker dg Tepat" (thank you for using the mask properly). All these stages were carried out sequentially, starting from face detection, nose detection, and mouth detection, so as to minimize detection errors.

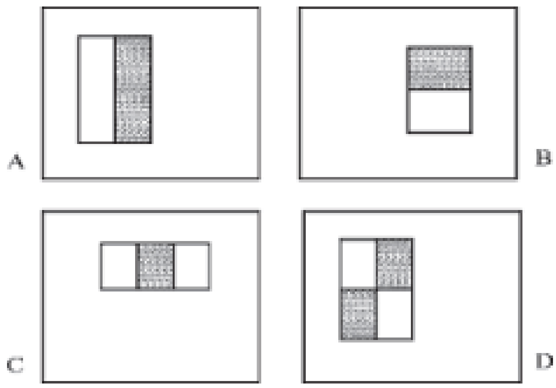

Fig. 4 Rectangular features.

2) Viola-Jones Algorithm: In this research, the detection algorithms of limbs such as nose and mouth were conducted using the Viola-Jones algorithm. The detection algorithm developed by Viola-Jones is an algorithm for classifying images based on simple feature values. Its fast feature processing compared to pixel image processing is the reason that this feature was used [6]. There are at least three feature types using black and white rectangular shapes.

There are three features used by Viola-Jones, viz., the two-rectangle features, the three-rectangle features, and the rectangular feature. The way to calculate the features in Fig. 4 above is to subtract the pixel value in the white area from the pixel value in the black/grey area. The basic resolution of the detector used was 24 x 24 [6], [22].

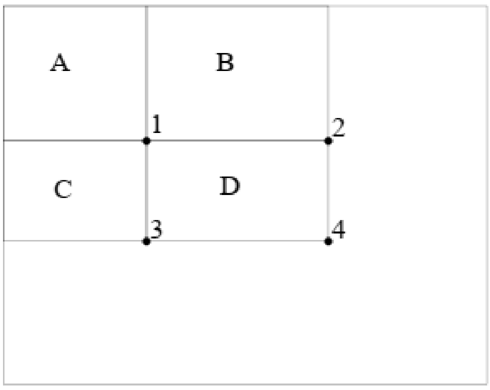

Fig. 5 Integral image.

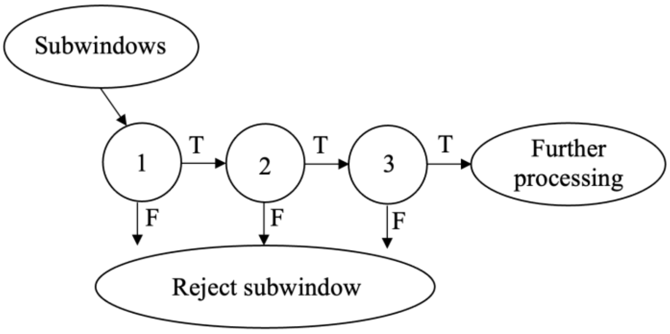

Fig. 6 Cascade classifier process.

These rectangular features can be calculated quickly using a representation for the image called an integral image. The integral image at location (x, y) contains the number of pixels above and to the left of (x, y) which can be calculated using (1).

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad (1)$$

with $ii(x,y)$ is the integral image and $i(x,y)$ is the original images.

$$s(x,y) = s(x, y-1) + i(x,y) \quad (2)$$

$$ii(x,y) = ii(x-1, y) + s(x,y) \quad (3)$$

with $s(x,y)$ is the number of cumulative rows, $s(x,-1) = 0$, and $ii(-1,y) = 0$. The integral image can be calculated in one pass through the original image [6].

Using the integral image, any number of squares can be counted in the four reference arrays (Fig. 5). The integral image value at location 1 is the number of pixels in rectangle A. The integral image value at location 2 is A + B, at location 3 is A + C, and at location 4 is A + B + C + D. The sum of D can be calculated as 4 + 1 - (2 + 3).

Eight references can be used to calculate the difference between two rectangular sums. Considering that two rectangle features comprise the adjacent rectangle sums, they can be calculated in the six array references, eight array references for the three-rectangle features, and nine array references for four-rectangle features [6].

Furthermore, for the classification process, Viola and Jones used an algorithm called cascade classifier. This cascade classification algorithm has the characteristics of stratified classification. This algorithm consists of several levels where the first level uses a simple classifier to reject most of the sub windows before more complex classifiers are called to achieve a low false positive rate. This makes this classifier set to be a powerful suite that can classify faces. The overall shape of the detection process is a degenerated decision tree, which is referred to as a "cascade", as shown in Fig. 6. A positive result from the first classifier triggers an evaluation of the second classifier adjusted to achieve a very high detection rate. A positive result from the second classifier triggers the third classifier, and so on. A negative result at any point leads to an outright rejection of the subwindow [6].

The cascade stage is built by training classifiers using AdaBoost machine learning and adjusting the threshold value to minimize false negatives. Note that this AdaBoost threshold

Fig. 7 System can correctly detect masked faces.

Fig. 8 Detected nose on the face when a mask placed on chin.

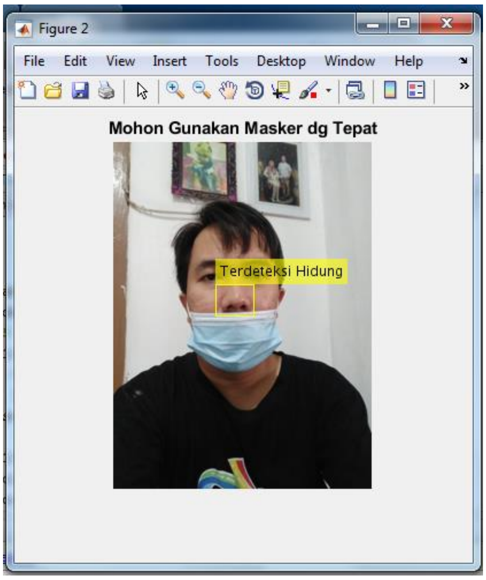

Fig. 9 Detected nose on the face when using a mask improperly.

Fig. 10 Detected mouth on the face when using a mask improperly.

is designed to produce a low error rate in training data. The description of the cascade classification is shown in Fig. 6 [6].

C. *Analysis Method*

At this stage, the analysis stage was carried out using the parameters of the accuracy, sensitivity, and specificity value. The accuracy parameter is a measure of the system's success in correctly identifying an object. The sensitivity parameter is the percentage of the number of correctly identified objects of interests in an image dataset/testing containing them. Next, the specificity parameter is the percentage of the number of objects other than objects of interest in all datasets that do not contain objects of interest. In calculating the values of accuracy, specificity, and sensitivity, this research employed true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. TP values indicate that the system has correctly detected objects of interest, TN values indicate that the system

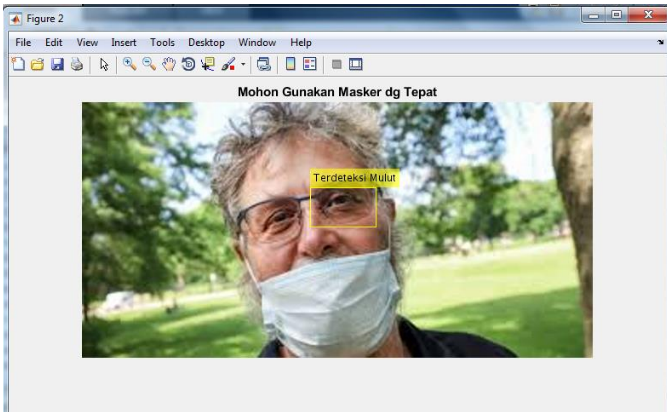

Fig. 11 Detection error on the eye object.

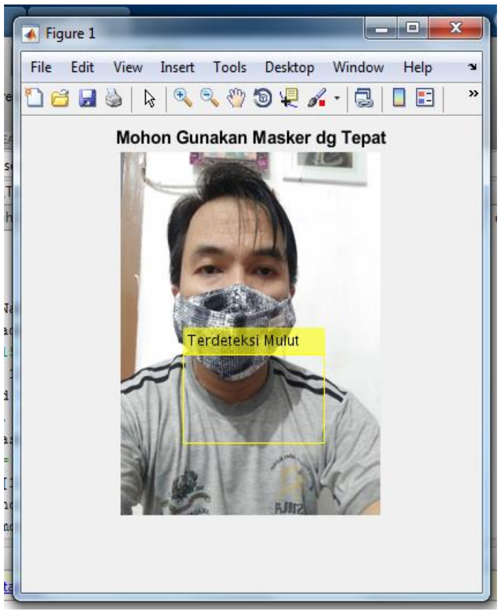

Fig. 12 Detection error on the mouth object.

can correctly distinguish objects and it will not detect other objects besides those that are of interest. FP values indicate errors in detecting objects, that is the system detects other objects as noses and mouths. FN values indicate detection errors, meaning that the system has failed to detect noses and mouths. Equations used in calculating the accuracy, sensitivity, and specificity values are respectively shown in (4), (5), and (6).

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} x\ 100\% \quad (4)$$

$$Sensitivity = \frac{TP}{TP+FN} x\ 100\% \quad (5)$$

$$Specificity = \frac{TN}{FP+TN} x\ 100\%. \quad (6)$$

IV. RESULTS AND DISCUSSION

As mentioned in the dataset section, there were five variants of testing images with different conditions. The system built using MATLAB 2017a software with a UNIPMA campus license, in which a nose and mouth detection algorithm has been implanted, yielded quite good results in detecting the state

TABLE I
IMAGE IDENTIFICATION RESULTS BASED ON CONFUSION MATRIX

Category	TP	TN	FP	FN
Faces wearing mask correctly (ten images)	8	0	2	0
Faces wearing masks with the visible noses (ten images)	10	0	0	0
Faces wearing masks with the visible mouths (ten images)	8	0	2	0
Ten mixed images of faces wearing masks placed on the neck/chin and not wearing masks (nose and mouth are visible)	10	0	0	0
Twenty mixed images (consisting of various conditions, such as various mask variants and users wearing other attributes)	8	10	2	0
Total	44	10	6	0

TABLE II
COMPARISON OF ACCURACY, SENSITIVITY, AND SPECIFICITY VALUES

Developed System	Accuracy	Sensitivity	Specificity
Face mask identification based on mouth and nose detection	90%	100%	62.5%

of a person's mask. The identification result display was directly displayed using the MATLAB software interface. Fig. 7 exhibits that the system was able to correctly detect the masked face so that a notification "Terima Kasih sudah Menggunakan Masker dg Tepat" appeared.

Fig. 8 shows that the system could correctly detect the nose object on a face with a mask worn on the neck/chin. Due to this inappropriate use, a notification "Mohon Gunakan Masker dg Tepat" appeared. The detected nose object was then labeled with a yellow box. Fig. 9 demonstrates how the system could correctly detect the nose object on a face when the mask was inappropriately used. As a result, a notification "Mohon Gunakan Masker dg Tepat" appeared. Fig. 10 displays the system could precisely detect the mouth object on a face when the mask was inappropriately used. Hence, a notification "Mohon Gunakan Masker dg Tepat" appeared.

Fig. 11 and Fig. 12 show errors detected by the system. In Fig. 11, the eye was incorrectly detected as the mouth object, resulting in the notification "Mohon Gunakan Masker dg Tepat" appearing. Due to the tilted position of the face and the similarity of the pattern/texture of the eyes to the mouth object, an FP value emerged in the detection system analysis. Fig. 12 shows an error in detecting the mouth on a masked face, hence a notification "Mohon Gunakan Masker dg Tepat" appeared. It was caused by the mask pattern, which interfered with the detection stages by the system. This issue led to the emergence of an FP value at the detection system analysis stage. Accessories such as variations in mask patterns, hijabs worn by women, and background images of objects can interfere with this detection system, resulting in less accurate detection.

Based on the research that has been done, the values of TP, TN, FP, and FN are presented in Table I. In Table I, there are many detection errors (FP) in the masked face image group with only mouths visible. It happened since the masks were inappropriately worn (only covered the noses) when image data were collected. As a result, the system faced difficulties in detecting mouth objects because of the narrow ROI region of the mouth areas. Detection errors were also found in the mixed images in the form of FP, meaning that the system still detects objects such as mouth and nose caused by other accessories such as hijab and mask patterns, lighting factors, and photo angles. Table II shows the value of accuracy, specificity value, and sensitivity value.

In the data presented in Table II, the accuracy value is 90%, the sensitivity value is 100%, and the specificity value is 62.5%. The measurement value of this parameter was obtained based on the values of TP, TN, FP, and FN in Table I and was the result of the confusion matrix of image testing. In Table II, the sensitivity parameter yielded a value of 100% since this parameter is a comparison of the correct detection result (TP) with a negative detection error (FN). The most influential factor on the sensitivity parameter was the negative error factor in which in this built system produced an FN value of 0; hence, based on the existing formula on the sensitivity parameter, a value of 100% was obtained. The specificity parameter obtained a value of 62.5% because there were still many errors detection caused by FP; in other words, the system detected other objects that were not objects of interest (i.e., objects other than the nose and mouth). This FP arise due to several factors, including the similarity of patterns from other objects in the image with the object of interest (the nose and mouth), the mask pattern used has a motif that is almost the same as the nose or mouth, the lighting that us too low or too high may affect the detection system. With the values of these parameters, it can be said that the system works well in detecting whether the mask is used properly or not.

Unlike previous studies, there have been no studies that have implemented a mask detection algorithm based on nose and mouth detection. It then becomes the gap between this research and previous research. This research is certainly very useful for the development of other applications related to the COVID-19 protection system during the pandemic.

V. CONCLUSION

From the research that has been completed, it is concluded that the nose and mouth detection algorithm can be used to identify the right use of masks on the face. By using sixty testing images, this study resulted in an accuracy value of 90%, a sensitivity value of 100%, and a specificity value of 62.5%.

The advice given for further research is to add a test dataset with various variations and accessories used on the face. In addition, this research can be used to develop other applications. Then, the researcher also suggests that this algorithm can be tested on images of many masked faces (multiface) in the same images.

Conflict of Interest

The authors declare that this research was conducted and written with no conflict of interest.

Author Contribution

Conceptualization, Denny Hardiyanto and Ihtiari Prastyaningrum; methodology, Denny Hardiyanto; software, Umi Kholifah; validation, Denny Hardiyanto, Ihtiari Prastyaningrum; formal analysis, Denny Hardiyanto; investigation, Ihtiari Prastyaningrum; resources, Denny Hardiyanto; data curation, Denny Hardiyanto; writing—original draft preparation, Ihtiari Prastyaningrum and Dyah Anggun Sartika; writing—review and editing, Umi Kholifah and Dyah Anggun Sartika.

Acknowledgment

Authors' gratitude goes to LP2M PGRI Madiun University for supporting this research, colleagues in the PTE UNIPMA study program who have fully supported. Many thanks should be given to all teams who have helped this research so that it could run smoothly and be completed.

References

[1] M.D.C. Pane (2021) "Virus Corona," [Online], https://www.alodokter.com/virus-corona, access date: 4-Apr-2021.

[2] (2020) "Kilas Balik Pandemi Covid-19 di Indonesia," [Online], https://www.cnnindonesia.com/nasional/20201110123516-25-568018/kilas-balik-pandemi-covid-19-di-indonesia, access date: 4-Apr-2021.

[3] T. Anastasia (2021) "Mau Pandemi Usai? Ketahui Pentingnya Gerakan 5M COVID-19," [Online], https://www.klikdokter.com/info-sehat/read/3644583/mau-pandemi-usai-ketahui-pentingnya-gerakan-5m-covid-19, access date: 4-Apr-2021.

[4] M.N. Afifah (2020) "Virus Corona Berpotensi Menyebar Melalui Percikan Ludah," [Online], https://health.kompas.com/read/2020/04/10/134800768/virus-corona-berpotensi-menyebar-melalui-percikan-ludah?page=all, access date: 26-Mar-2021.

[5] N. Azizah (2020) "Warga Bogor tak Pakai Masker Didenda Rp 50.000," [Online], https://www.republika.co.id/berita/qdmdn1463/warga-bogor-tak-pakai-masker-didenda-rp-50000, access date: 2-Apr-2021.

[6] P. Viola and M. Jones, "Rapid Object Detection Using a Boosted Cascade of Simple Features," *Proc. IEEE Comput. Soc. Conf. Comp. Vis., Pattern Recognit.*, 2001, pp. I-511-518.

[7] P. Viola and M.J. Jones, "Robust Real-Time Face Detection," *Int. J. Comput. Vis.*, Vol. 57, No. 2, pp. 137–154, May 2004.

[8] D. Hardiyanto and D.A. Sartika, "Optimalisasi Metode Deteksi Wajah Berbasis Pengolahan Citra untuk Aplikasi Identifikasi Wajah pada Presensi Digital," *Setrum*, Vol. 7, No. 1, pp. 107–116, 2018.

[9] P. Indrawan, "Implementasi Sistem Pengenalan Wajah sebagai Penghubung Jejaring Sosial: Aplikasi Mobile untuk Deteksi Wajah dengan Android Face Detector API dan Komunikasi REST ke Komputer Awan," Undergraduate thesis, Universitas Indonesia, Jakarta, Indonesia, 2012.

[10] W. Ge, Z. Xu, C. Shi, and W. Zhan, "Recognition of Expression-Variant Faces Using SIFT Method," *2012 Fourth Int. Conf. Comput. Intell., Commun. Netw.*, 2012, pp. 398–401.

[11] P. Saikia, G. Janam, and M. Kathing, "Face Detection Using Skin Colour Model and Distance Between Eyes," *Int. J. Comput. Commun., Netw.*, Vol. 1, No. 3, pp. 102–105, Nov.-Dec. 2012.

[12] K. Vikram and S. Padmavathi, "Facial Parts Detection Using Viola Jones Algorithm," *2017 4th Int. Conf. Adv. Comput., Commun. Syst. (ICACCS)*, 2017, pp. 1–4.

[13] A. Tawari and M. Trivedi, "Audio-Visual Data Association for Face Expression Analysis," *Proc. 21st Int. Conf. Pattern Recognit. (ICPR2012)*, 2012, pp. 1120–1123.

[14] [G. Deore, R. Bodhula, V. Udpikar, and V. More, "Study of Masked Face Detection Approach in Video Analytics," *Conf. Adv. Signal Process. (CASP)*, 2016, pp. 196–200.

[15] R. Liu and Z. Ren, "Application of Yolo on Mask Detection Task," *13th Int. Conf. Comput. Res. Develop. (ICCRD)*, 2021, pp. 130–136.

[16] N.U. Din, K. Javed, S. Bae, and J. Yi, "A Novel GAN-Based Network for Unmasking of Masked Face," *IEEE Access*, Vol. 8, pp. 44276–44287, Mar. 2020.

[17] J. Zhang, F. Han, Y. Chun, and W. Chen, "A Novel Detection Framework About Conditions of Wearing Face Mask for Helping Control the Spread of COVID-19," *IEEE Access*, Vol. 9, pp. 42975–42984, Mar. 2021.

[18] M.S. Ejaz and M.R. Islam, "Masked Face Recognition Using Convolutional Neural Network," *2019 Int. Conf. Sustain. Technol. Industry 4.0 (STI)*, 2019, pp. 1–6.

[19] I.B. Venkateswarlu, J. Kakarla, and S. Prakash, "Face Mask Detection Using MobileNet and Global Pooling Block," *2020 IEEE 4th Conf. Inf., Commun. Technol. (CICT)*, 2020, pp. 1–5.

[20] B. Yilmazer and S. Solak, "Cloud Computing Based Masked Face Recognition Application," *2020 Innov. Intell. Syst., Appl. Conf. (ASYU)*, 2020, pp. 1–5.

[21] M.S. Ejaz, M.R. Islam, M. Sifatullah, and A. Sarker, "Implementation of Principal Component Analysis on Masked and Non-masked Face Recognition," *2019 1st Int. Conf. Adv. Sci., Eng., Robot. Technol. (ICASERT)*, 2019, pp. 1–5.

[22] M.D. Puto, T.B. Adji, and B. Winduratna, "Sistem Deteksi Wajah dengan Menggunakan Metode Viola-Jones," *Semin. Nas. Science, Eng. Technol.*, pp. 1–5, 2012.