A Comparative Analysis of Machine Learning Algorithms for Autonomous Face Mask Detection

Destin Joanny, Erik Sanjaya, Ronaldo, and Derwin Suhartono
Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia
Email: {destin.joanny, erik.sanjaya, ronaldo008}@binus.ac.id, dsuhartono@binus.edu

***Abstract*—Covid-19 pandemic is a global disease caused by severe acute respiratory syndrome. The rising number of infected individuals and death cases remain a major problem in 2021. Health protocols such as wearing a face mask is taken as prevention method to suppress the significantly growing numbers. Popular machine learning techniques have been addressed to assist in the global issue. This study intends to compare several popular classification algorithms namely K-Nearest Neighbors (K-Nearest Neighbors), Support Vector Machine (Support Vector Machine), Convolutional Neural Network (CNN), Decision Tree, and Naive Bayes for autonomous face mask detection as there are still limited sources of studies that does performance comparison of the related field. Experimental results are analyzed and evaluated using various measures such as precision, recall, accuracy, and F1 Score. Convolutional Neural Network proves to have the most promising performance than the other classification techniques to identify whether a person is wearing a mask or not with over 97% of accuracy.**

***Index Terms*—Covid-19, autonomous face mask detection, classification algorithm**

I. Introduction

Covid-19 outbreak was first detected in Wuhan, China in late 2019 [1]. Since then, it has continued to spread across the globe and as of 3rd September 2021, the virus has infected over 218 million lives with 4.5 million confirmed death cases [2]. Being a highly contagious virus, it spreads via respiratory droplets when an infected individual sneezes, coughs, or even talks and sings [3]. A symptomatic patient commonly deals with fever, cough, fatigue, shortness of breath, dizziness, as well as anosmia and dysgeusia [4]. A study claims that the elderly and individuals with underlying diseases such as asthma, diabetics, and heart disease are the most vulnerable to infection [5].

Fig. 1 shows that the number of infected individuals is consistently increasing every month. To squeeze out the steadily growing numbers of transmission due to the unknown specific cure for the outbreak, the prevention method is taken. Research has indicated that wearing mask could prevent infection of coronaviruses from infected individuals [7]. People are obligated to wear a mask in public place to limit the spread of the disease, currently making masks one of the essential objects for humans. Nevertheless, there are still many reckless people who neglect the importance of masks and are not aware of the health protocols. To help the global issue, there have been some recent studies of an autonomous system addressing face mask detection. The system helps to detect whether a person is wearing a face mask or not in real time.

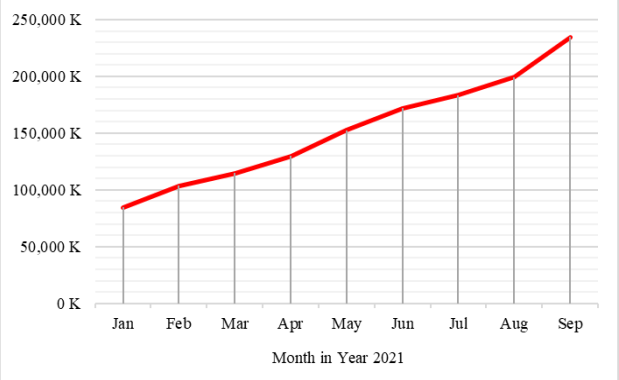

Figure 1. Cumulative cases of Covid-19 infected individuals worldwide since 2021 [6].

Various machine learning algorithms have been used before to develop the system. In this paper, we conducted research to compare and analyze the performance of the applicable machine learning algorithms. Many algorithms with high accuracy are introduced in developing face mask detection, yet there is little information about which algorithm is truly the best of all. The analysis will be conducted by comparing the algorithms head on with the same dataset to provide unbiased outcome. The algorithms to be compared are: Support Vector Machine, K-Nearest Neighbors, Convolutional Neural Network, Decision Tree, and Naive Bayes. Through this research, the best algorithm for autonomous face mask detection could be revealed. The result is expected to be helpful for developers or researchers in choosing the finest algorithm.

The remaining part of this paper is organized as follows. Section II briefly describes the literature review used to support the research. The research methodology is

Manuscript received January 1, 2022; revised June 10, 2022.

doi: 10.18178/joig.10.3.122-126

introduced in Section III. Section IV presents the result and discussion, and Section V concludes the research.

II. Literature Review

A study proposed a face recognition system to identify faces with the help of Naive Bayes algorithm [8]. The experiment results in 73% of accuracy. Another face mask detection study which uses decision tree algorithm gives accuracy of 83.35% with 33.1s of running time. The authors also proposed a Convolutional Neural Network method that claimed to minimize empirical training as well as generalization error by finding the largest margin between hyperplanes. The proposed method gives 91.11% of accuracy with 7.24s of running time which is faster than decision tree [9].

Comparison of Convolutional Neural Network, K-Nearest Neighbors, and Support Vector Machine algorithms for face mask detection applies 5 cross-validations on 3.886 image data. The study reveals that Convolutional Neural Network requires longer execution time yet provides higher accuracy. On the other hand, K-Nearest Neighbors and Support Vector Machine execute faster yet has lower accuracy. The final accuracy gathered for Convolutional Neural Network, K-Nearest Neighbors, and Support Vector Machine are 96%, 81.15%, 87.21%, respectively. For the execution time for Convolutional Neural Network, K-Nearest Neighbors, and Support Vector Machine are 2,507.802; 185.27; 352.968 seconds, respectively [10]. Another simplified Convolutional Neural Network approach used two datasets. The authors deal with image data using Numpy and OpenCV. In the proposed model, TensorFlow and Keras help build the desired model based on Convolutional Neural Network architecture. The experiment obtains 95.77% and 94.58% respectively for both datasets. The authors claim that MaxPooling is one of the main reasons behind the accuracy obtained as it provides fundamental translation invariance to the internal representation along with the reduction in the number of parameters the model must learn [11].

The other work found which focuses on implementing a face mask and social distancing detection model as an embedded vision system uses the pretrained models such as the MobileNet, ResNet Classifier, and VGG on the face mask dataset. The result of the paper presents VGG-19 model is the best model among the others [12]. Another algorithm proposed a YOLOv3 algorithm which implements Haar Cascade Classifier, feature extraction by Haar Wavelet technique and AdaBoost to eliminate redundant features. The system is able to distinguish whether the person is wearing a mask or not with 90.1% of accuracy [13]. The improved YOLO algorithm is introduced to YOLOv4. The authors use CSPDarkNet53 into the backbone feature extraction network. Furthermore, adaptive scaling algorithm can reduce the model's reasoning time. Lastly, the upgraded PANet structure is proposed so that the network has more semantic information in the feature layer. The result of this experiment reaches 98.3% of accuracy [14].

III. Research Methodology

The research methodology comprises of 5 main stages: gathering dataset, data preprocessing, classification model, training and testing, and performance analyzation as shown in Fig. 2.

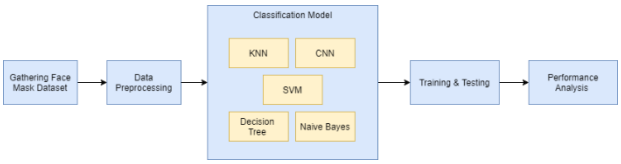

Figure 2. Research workflow.

A. Dataset

The dataset used in this study is a face mask detection dataset available on Github [15]. It consists of 2 folders of people wearing mask and people not wearing mask with the distribution of 690 with mask images and 671 without mask images in RGB colors. These images are collected for training and testing purposes. Fig. 3 shows images of an individual with mask and without mask.

Figure 3. Dataset images sample of an individual (a) with mask (b) without mask.

B. Data Preprocessing

As the dataset images are not equal in size, the first preprocessing action taken is to resize those images in a minimum of 64×64 pixels. Next action is to implement image enhancement. Auto white balance and unsharp masking will be used to enhance the image quality and edges. Auto white balance estimates the color of the overall scene illumination accurately to ensure color consistency [16]. Meanwhile, unsharp masking is utilized for contrast enhancement of the images [17]. The preprocessing phase helps the classification model to perform better with the training and testing process. For calculation, images are converted into new format of numpy array.

C. Classification Model Algorithm

1) K-Nearest Neighbors (KNN)

The model uses Manhattan as its distance measurement as it outperforms both Euclidean and Minkowski. The distance is tested with the number of neighbors from 1 to 20. The result shows 15 is the best parameters for the model. The experiment also implements K-Nearest Neighbors algorithm with Hyperparameter Tuning alongside RandomizedSearchCV and the number of cross validations is 10. Furthermore, the weights parameter is the distance itself.

2) Convolutional Neural Network (CNN)

This study on Convolutional Neural Network model uses Keras library along with TensorFlow. The convolution model uses 2D Convolutional Layer. Meanwhile, the activation function used is the Rectified Linear Unit (ReLu).

The first and second layer is the Conv2D layer (2D Convolutional Layer) with 64 filters and kernel size equals to 3. The activation function used is ReLu with the initialization input shape as 64×64×1. The third layer is MaxPooling2D with the default pool size of 2×2. In the fourth and fifth layer, Conv2D with the same parameters used as the first and second layer is implemented. The next layer is followed by MaxPooling2D layer with pool size of 2×2. For the seventh and eighth layer of Conv2D uses the exact same parameters as the previous Conv2D layers followed by another MaxPooling2D layer with pool size of 2×2.

The next layer is Flatten. After all the layers are flattened into a single 1D layer, the dense layer with 512 units and ReLu activation function comes to play. Followed by Dropout with 0.3 rate to make the model less prone to overfitting. The last few layers are another 512-unit dense layer, 0.3-rate dropout layer, 256-unit dense layer, 128-unit dense layer, and 3-unit dense layer and the activation function used is the softmax function.

Finally, the model is compiled with Adam Optimizer and Sparse Categorical Cross Entropy as the loss function. Meanwhile for the training, the number of epochs used is 10 epochs.

3) Support Vector Machine (SVM)

From experiment, the Support Vector Machine model is built based on Radial-Basis Function (RBF) kernel as it gives better result than using linear kernel and poly kernel

4) Decision tree

The model of decision tree is built without Hyperparameter Tuning. The model has no specific max_depth and implements gini impurity as the criterion

5) Naïve bayes

Naive Bayes model is built using Gaussian Naive Bayes with no Hyperparameter Tuning.

D. Training & Testing

For training and testing purposes, data images are randomly split into two parts with the proportion of 80% (1088 images) for training and the remaining 20% (273 images) of the dataset images for testing. In training, the training set of images are loaded into the classification models for each classification algorithm. Testing stage is executed to validate the training model

E. Performance Analysis

The performance for each proposed algorithm is analyzed based on several metrics such as accuracy, precision, recall, and F1 score of the training and testing process [18].

1) Accuracy

Accuracy is the ratio of the correctly predicted instances to total instances observed.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

2) Precision

Precision is the ratio of true positive instances to total positive instances observed.

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

3) Recall

Recall is the ratio of true positive instances to total correctly predicted instance.

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

4) F1 score

F1 Score is the harmonic mean of precision and recall.

$$F_1\ Score = 2\left(\frac{Precision \times Recall}{Precision+Recall}\right) \quad (4)$$

IV. Result and Discussion

This section aims to deliver the result of the conducted work and analyze the performance of the classification models built on previous section. The performance is analyzed and evaluated based on several measurements such as precision, accuracy, recall, and F1-score. All these measurements are positively-oriented scores, meaning the higher outcome the better.

TABLE I. Performances Analysis Comparison

Metrics	Classification Algorithms				
	KNN	SVM	**CNN**	Decision Tree	Naïve Bayes
Accuracy	0.857	0.889	**0.970**	0.910	0.802
Recall	0.874	0.890	**0.967**	0.915	0.764
Precision	0.856	0.887	**0.965**	0.907	0.856
F1 Score	0.865	0.890	**0.965**	0.915	0.807

Table I depicts the result of the performance of different machine learning techniques computed on testing process. In order to make it easier to be evaluated, Fig. 4 and Fig. 5 are provided separately. Both figures show the results of classification algorithm such as K-Nearest Neighbors, Support Vector Machine, Convolutional Neural Network, and Decision Tree in bar diagram. The observation of Fig. 4 and Fig. 5 concludes that Convolutional Neural Network consistently produced the highest value of recall, precision, and accuracy among the other studied techniques which are 0.967, 0.965, and 0.970 respectively. Meanwhile, in Fig. 5, Convolutional Neural Network also outperforms the other techniques in terms of F1 Score.

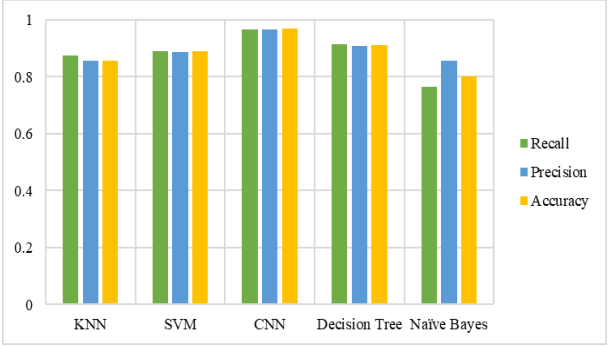

Figure 4. Prediction result of classification algorithm.

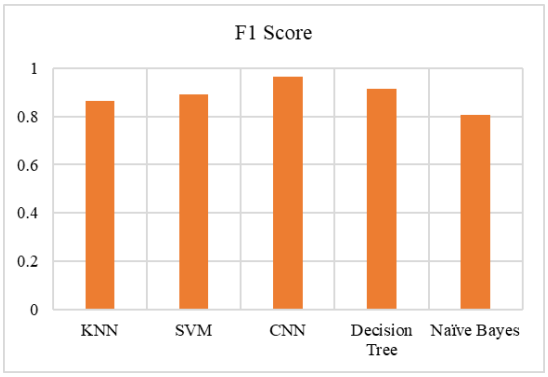

Figure 5. Accuracy result of classification algorithm.

Based on the evaluation criteria mentioned above, the model which delivers highest accuracy, precision, recall, and F1 score is considered to be the best approach for the system. This finding proves that Convolutional Neural Network is the best algorithm for face mask detection with an accuracy of 97% followed by Decision Tree with 91% accuracy, Support Vector Machine with 88.9% accuracy, K-Nearest Neighbors with 85.7% accuracy, and lastly Naive Bayes with 80.2% accuracy. Although Convolutional Neural Network might take longer time in the training process, but it performs better than the other popular classification algorithms.

V. Conclusion

The conclusion of the conducted research are as follows:

- Several popular classification machine learning algorithms such as K-Nearest Neighbors, Support Vector Machine, Convolutional Neural Network, Decision Tree, and Naïve Bayes are tested, compared, and analyzed to witness which algorithm is the most superior for face mask detection system.
- Dataset images of people wearing mask and not wearing mask first undergoes data preprocessing stage. The preprocessing stage involves image enhancement, image reshaping and lastly being converted into numpy array.
- Next, the images are split into train data and test data with the ratio of 1:4. Train data are then fit into each classification models with certain parameters to optimize the outcome.
- Experimental results revealed that Convolutional Neural Network (Convolutional Neural Network) model achieves better performance on the evaluated metrics such as precision, recall, F1 score, and accuracy.

The experimental results are expected to be useful for any researchers to develop or to enhance the face mask detection system. The paper has the potential to open up other new perspectives for future work. More findings could be done to optimize the model of K-Nearest Neighbors by using appropriate parameter tuning. Further examination on image enhancement and image perspective could also be done for more convincing outcome. Lastly, since Convolutional Neural Network is the most promising model, it is very feasible to integrate the model to hardware tools like camera for specific purposes.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

D. Joanny participated in the data analysis and drafted the manuscript. E. Sanjaya carried out the data analysis and participated in the technical issues. Ronaldo conceived the study and performed the experiment. D. Suhartono advised all process for this work. Regarding the manuscript, D. Joanny, E. Sanjaya, R, and D. Suhartono helped to revise the manuscript. All authors had read and approved the final manuscript.

References

[1] J. T. Wu, K. Leung, and G. M. Leung, "Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study," *Lancet (London, England)*, vol. 395, pp. 689-697, February 29, 2020.

[2] WHO Coronavirus (COVID-19) Dashboard. World Health Organization (WHO). [Online]. Available: https://covid19.who.int/

[3] M. Lotfi, M. R. Hamblin, and N. Rezaei, "COVID-19: Transmission, prevention, and potential therapeutic opportunities," *International Journal of Clinical Chemistry*, vol. 508, pp. 254-266, 2020.

[4] M. Cascella, *et al.*, "Features, evaluation, and treatment of Coronavirus (COVID-19)," in *StatPearls*, Treasure Island (FL): StatPearls Publishing, 2021.

[5] A. Sanyaolu, C. Okorie, Z. Hosein, R. Patidar, P. Desai, S. Prakash, U. Jaferi, J. Mangat, and A. Marinkovic, "Global pandemicity of COVID-19: Situation report as of June 9, 2020," *Infectious Diseases: Research and Treatment*, vol. 14, January 2021.

[6] COVID live update. (2021). Worldometer. [Online]. Available: https://www.worldometers.info/coronavirus/coronavirus-cases/#total-cases

[7] N. H. L. Leung, *et al.*, "Respiratory virus shedding in exhaled breath and efficacy of face masks," *Nat Med*, vol. 26, pp. 676-680, May 2020.

[8] K. Kaarthik, *et al.*, "Face detection and recognition using naïve bayes algorithm," *International Journal of Disaster Recovery and Business Continuity*, vol. 11, pp. 11-18, 2020.

[9] G. K. J. Hussain, *et al.*, "The face mask detection technology for image analysis in the Covid-19 surveillance system," *J. Phys.: Conf. Ser.*, 2021.

[10] M. F. Naufal, S. F. Kusuma, Z. A. Prayuska, A. A. Yoshua, Y. A. Lauwoto, N. S. Dinata, and D. Sugiarto, "Comparative analysis of image classification algorithms face mask detection," *Journal of Information Systems Engineering and Business Intelligence*, vol. 7, no. 1, pp. 56-66, 2021.

[11] A. Das, M. W. Ansari, and R. Basak, "Covid-19 face mask detection using TensorFlow, Keras and OpenCV," in *Proc. IEEE 17th India Council International Conference*, 2020, pp. 1-5.

[12] S. Teboulbi, S. Messaoud, M. A. Hajjaji, and A. Mtibaa, "Real-Time implementation of ai-based face mask detection and social distancing measuring system for Covid-19 prevention," *Scientific Programming*, vol. 2021, article ID 8340779, p. 21, 2021.

[13] T. Q. Vinh and N. T. N. Anh, "Real-Time face mask detector using YOLOv3 algorithm and Haar cascade classifier," in *Proc. International Conference on Advanced Computing and Applications*, 2020, pp. 146-149.

[14] J. Yu and W. Zhang, "Face mask wearing detection algorithm based on improved YOLO-v4," *Sensors*, vol. 21, no. 9, p. 3263, 2021.
[15] P. Bhandary. (2019). Face mask dataset. [Online]. Available: https://github.com/prajnasb/observations/tree/master/experiements/data.
[16] G. Zapryanov, D. Ivanova, and I. Nikolova, "Automatic white balance algorithms for digital still cameras - A comparative study," *Information Technologies and Control*, vol. 1, pp. 16-22, 2012.
[17] A. Polesel, G. Ramponi, and V. J. Matthews, "Image enhancement via adaptive unsharp masking," *IEEE Transactions on Image Processing*, vol. 9, no. 3, pp. 505-510, March 2000.
[18] H. Dalianis, "Evaluation metrics and evaluation," in *Clinical Text Mining*, Springer, Cham, 2018.

Destin Joanny was born in Medan on 3rd August 2001. She graduated from SMA Sutomo 1 Medan and is currently a student in Bina Nusantara University majoring in Computer Science at BINUS Kemanggisan, Jakarta, Indonesia.

Erik Sanjaya was born in Jakarta on 28th November 2001. He graduated from SMAN 2 Jakarta and is currently a student in Bina Nusantara University majoring in Computer Science at BINUS Kemanggisan, Jakarta, Indonesia.

Ronaldo was born in Tangerang on 31st October 2001. He graduated from SMA Mutiara Bangsa 1 and is currently a student in Bina Nusantara University majoring in Computer Science at BINUS Kemanggisan, Jakarta, Indonesia.

Derwin Suhartono is faculty member of Bina Nusantara University, Indonesia. He got his PhD degree in computer science from Universitas Indonesia in 2018. His research field is natural language processing. Recently, he is continually doing research in argumentation mining and personality recognition. He actively involves in Indonesia Association of Computational Linguistics (INACL), a national scientific association in Indonesia. He has his professional memberships in ACM, INSTICC, and IACT. He also takes role as reviewer in several international conferences and journals.