International Journal of Multidisciplinary Research and Publications
ISSN (Online): 2581-6187

Application of Deep Learning for Facial Recognition Obstructed by Face Masks and/or Glasses

Alexander Waworuntu[1], Ester Lumba[2]
[1]Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia
[2]Department of Informatics, Bunda Mulia University, Jakarta, Indonesia
Email address: [1]alex.wawo@umn.ac.id, [2]l0178@lecturer.ubm.ac.id

Abstract— *Deep learning allows a machine to learn autonomously using neural networks. This study applies deep learning for facial recognition even when using masks and/or glasses. The research uses Python and the pre-trained convolutional network library VGG16. The application environment requires a dataset where the photographs of the participants are stored, which amount to a total of 1000 images, essentially capturing the breadth of the face using a mask and / or lenses; and placed in three subfolders: Train, Test and Valid. In the first phase, the training is carried out, learning in 500 epochs to create trained model. In a second phase, facial recognition is performed with the presence of occlusions on the face; using a webcam. The accuracy or precision achieved in the training of the neural network in Google Collab is 0.1 and the percentage of success obtained from the application in the second phase is 84.68%. The project manages to recognize people when they use a mask, glasses, mask and glasses or without the use of these; with the success rate mentioned above.*

Keywords— *Artificial Intelligence, Deep Learning, Facial Recognition, Neural Networks, VGG16.*

I. INTRODUCTION

In recent years, the facial recognition system has been rapidly developed for uses within information security, public safety, civil economy and entertainment [1], as well as work attendance registration, as a safe option [2]. However, in the advancement of technology, there are various challenges for facial recognition. Among the challenges are situations such as lighting variation that causes dramatic changes in facial features [3]. The facial recognition process contemplates the detection of facial objects, extraction of facial features and the classification process [4], during the extraction of facial features challenges are included such as: facial variation, changes in facial expressions, change of scale, effect of blur and presence of occlusions, when using accessories such as glasses, scarves, caps, etc., which cause loss of important information [5][6].

The use of an accessory does not allow a good facial recognition, since it is considered an occlusion; but there are cases in which the use of glasses cannot be avoided, due to the visual problem that a person has; nor can the use of a mask be avoided due to the health condition, especially during the pandemic.

Artificial Intelligence (AI) is in charge of studying models capable of carrying out activities typical of human beings based on reasoning and pattern [7]. Within AI there are two more specific branches, Machine Learning (ML) refers to automatic learning that seeks to create programs to generalize pattern from information [8]. The branch of Deep Learning (DL) imitates the functioning of the human brain when processing data and creating patterns, composed of networks capable of learning from unstructured data [9].

The main advantage of DL is the possibility of using a robust number of samples for training, allowing facial representation to be learned resilient to changes in the training data [10]. For which, if greater precision is desired, it is necessary to expand the dataset and familiarize the systems with the characteristics of each face [11]. What is sought is to reduce the number of useless features of the facial data to find the unique and effective features for classification [12].

This study applies the tools and process mentioned to achieve effective facial recognition, even when the subject has the mask and lens accessories that cause occlusions within the recognition process that affect the result in the identification. The Dataset used is 1,000 images at different angles and with different accessories that could create occlusions of ten people, forming subfolders, which are: Train, Valid and Test of 800, 100 and 100 images respectively for training, validation and test.

Using the images entered into the dataset of the test subjects, the training phase begins to generate knowledge, which once completed will be known as the trained neural network. Using this network, we will be able to identify the test subjects based on the pattern of images.

II. METHODOLOGY

This study occurs in two phases: the training phase (stage 1 and 2) and the recognition phase (stage 3 and 4), as shown in Figure 1.

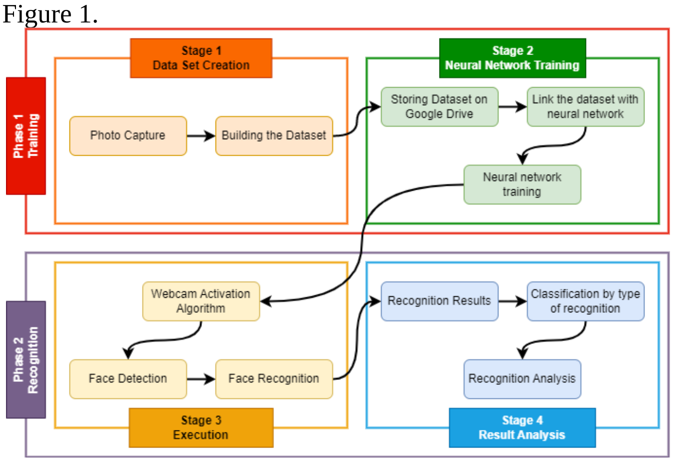


Fig. 1. Face Recognition Phase and Stage

A. Stage 1: creating the dataset

For the creation of the dataset there was a total of 2000 color photographs as a sample, manually cropped, stored as shown in Figure 2. Within each personal folder there are photographs focused on the face at different angles of the person without body movement without losing the features that characterize each person, such as the eyes, nose and mouth, with accessories such as masks and glasses.

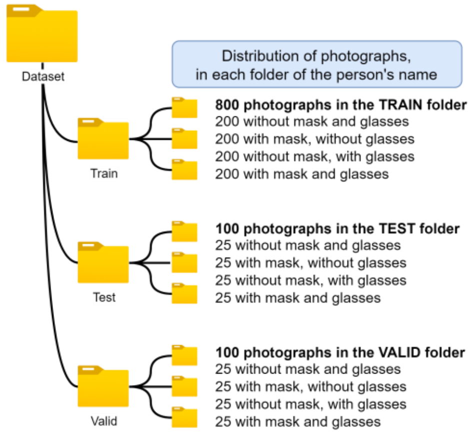


Fig. 2. Dataset distribution

B. Stage 2: training the neural network

For neural network training, it is necessary to upload the compressed dataset to the cloud, because the collected images are in bulk and become too heavy, the solution is to upload all the images to the cloud. The hosting used for dataset storage is Google Drive. As shown in Figure 3, the process of facilitating the use of datasets for neural network training takes place, which is done by transferring machine learning, with the use of the pre-trained VGG16 network, which facilitates image processing.

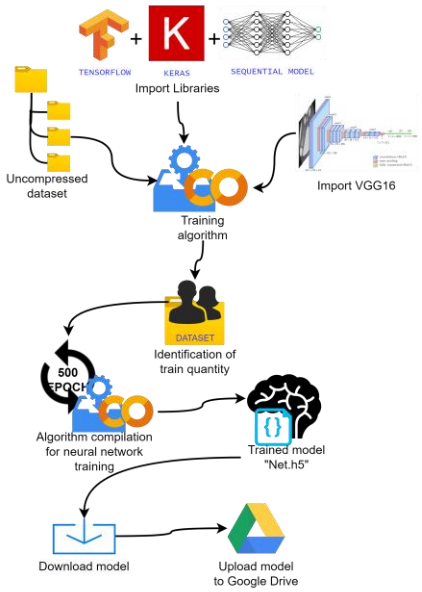


Fig. 2. Neural network training

The training was carried out in 500 epochs, at the end of the learning the trained neural network was obtained, which was downloaded with the name "Net.h5".

C. Stage 3: execution

For recognition, the trained network is previously hosted in google drive, and the same process of decompressing the dataset is followed.

The use of the Open CV library and *haardcascade frontalface* is essential, together with the Numpy and Python libraries for the identification of the image pattern and its various characteristics using the vector space for real-time recognition at the time of webcam activation.

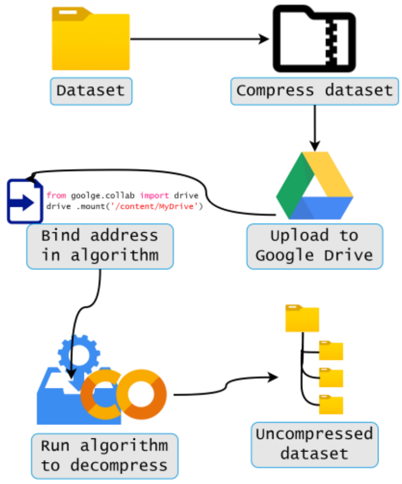


Fig. 3. Decompress dataset

D. Stage 4: analysis of results

Carrying out different tests in relation to the case proposed that are face without a mask and without glasses, face with a mask without glasses, face with glasses and a mask, face with glasses without a mask, these tests are classified by type to analyze the percentage of recognition of each one of them, so that through the application of Deep Learning that it meets an acceptable percentage of accuracy when recognizing and identifying the face of people, including the restrictions that are caused by masks or glasses that reduce facial features for a better recognition.

III. Result and Discussion

The training was given with 500 epochs or cycles, which the neural network was learning in approximately 12 hours, giving a result of recognition precision of 0.1, which shows a certain approximation to the exact precision of 0.1. The precision shows the number of errors that the model makes, after comparing with the main objective, this is due to the amount of epoch that is used for training, since the sample is of a large magnitude. But if it increases, the increase in training hours occurs proportionally, since the machinery used to carry out the training influences.

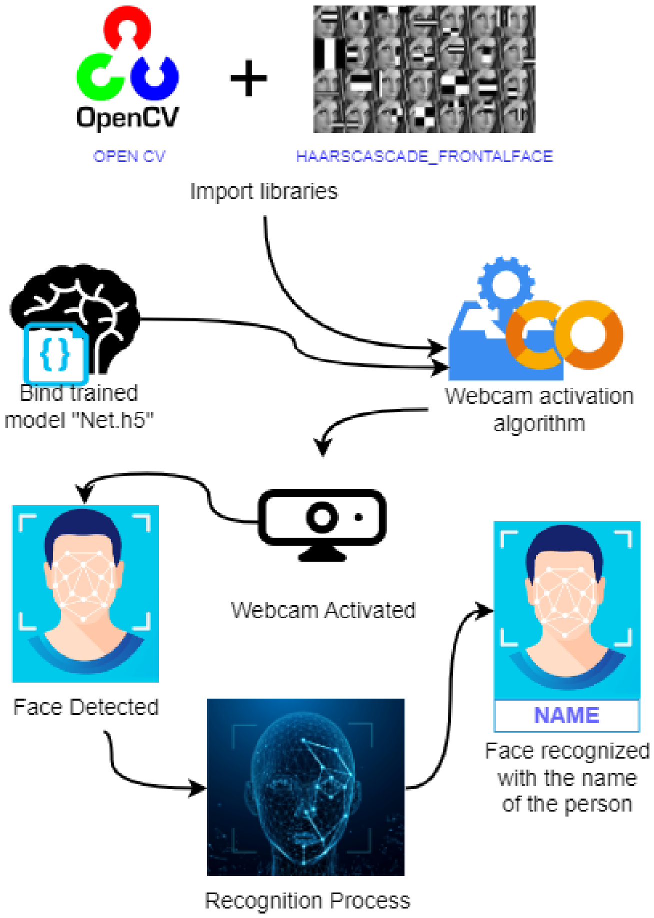


Fig. 4. Facial recognition

The training application tests were carried out, of a total of 10 people as test subject, with a recognition time of 10 seconds by classifying the following cases: without glasses-without mask (OO), with glasses-without mask (GO), without glasses-with mask (OM) and with glasses-with mask (GM). Table I shows the facial recognition results of each test subject based on each case.

TABLE I. Test by subjects and cases

Test Subject	**Case**				**Percentage**
	OO	**GO**	**OM**	**GM**	
1	1.00	0.98	0.78	0.65	85.25%
2	1.00	1.00	0.87	0.85	93.00%
3	0.98	0.88	0.89	0.88	90.75%
4	0.65	0.52	0.43	0.27	46.75%
5	0.78	0.87	0.91	0.77	83.25%
6	0.83	1.00	0.93	0.81	89.25%
7	1.00	0.89	0.78	0.89	89.00%
8	1.00	0.94	0.89	0.84	91.75%
9	0.99	1.00	0.94	0.67	90.00%
10	0.97	0.89	0.87	0.78	87.75%
Avg. %	92.00%	89.70%	82.90%	74.10%	**84.68%**

Tables II and III show examples of the recognition results of the observed test subjects, describing test success and recognition according to the classification of each case.

TABLE II. Test Sample – Ravelino (TC-01)

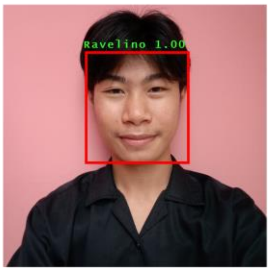
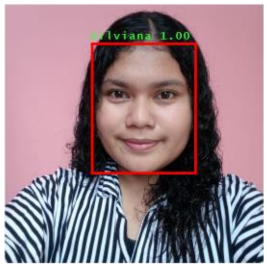
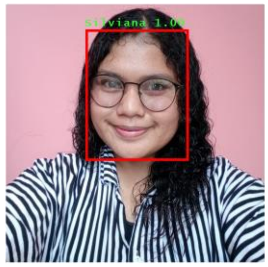
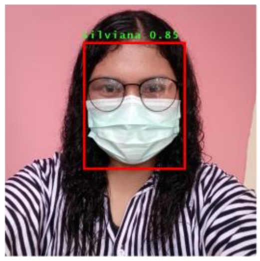
Name: Ravelino	Test: TC-01 Success: 85.25%
without mask | without Glasses	without mask | with glasses
	
with mask | without glasses	with mask | with glasses
	

TABLE III. Test Sample – Silviana (TC-02)

Name: Silviana	Test: TC-02 Success: 93.00%
without mask | without Glasses	without mask | with glasses
	
with mask | without glasses	with mask | with glasses
	

IV. CONCLUSIONS

The percentage of precision reached in the training phase of the neural network is 0.1, which demonstrates a 10% error in the recognition phase, thus obtaining a 84.68% success rate in carrying out the tests, observing a good result. regarding the percentage of error in the recognition of the person using a mask, glasses, even if the person is using both accessories or not, given that the created neural network allows facial recognition with the presence of occlusions such as the glasses and the mask.

The use of the VGG16 library allows the transfer of learning in training and the easy processing of the images found in the dataset. The delay time for training the neural network at 500 epochs is approximately 12 hours in Google Colab, working with 1000 images, so the processing time depends on the number of images to train and the training cycles.

The recognition percentages seen in the efficacy tests depend a lot on the training photos that are taken, the more photos and the more head rotation movements are captured, the better facial recognition will be with the use of a mask, glasses, both accessories at the same time or with none of them.

ACKNOWLEDGMENT

Thank you to Universitas Multimedia Nusantara, Indonesia for supporting this research. Hopefully this research can contribute to the advancement of Information Technology.

REFERENCES

[1] A. Kusnadi et al., “Multi-Instance Face Recognition System Using Pca And Ann”, *Nveo-Natural Volatiles & Essential Oils Journal| NVEO*, vol. 9, no. 1, pp. 710–718, 2022.

[2] N. Kar et al., “Study of Implementing Automated Attendance System Using Face Recognition Technique”, *International Journal of Computer and Communication Engineering*, vol. 1, no. 2, pp. 100–103, 2012.

[3] A. Archilles, A. Wicaksana, “Vision: a web service for face recognition using convolutional network”, *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 18, no. 3, pp. 1389, 2020, doi:10.12928/telkomnika.v18i3.14790.

[4] J.-F. Connolly, E. Granger, R. Sabourin, “An adaptive classification system for video-based face recognition”, *Information Sciences*, vol. 192, pp. 50–70, 2012, doi:10.1016/j.ins.2010.02.026.

[5] Z. Wang et al., “Masked Face Recognition Dataset and Application”, *arXiv preprint arXiv:2003.09093*, pp. 1–3, 2020.

[6] J. Deng et al., “Masked Face Recognition Challenge: The InsightFace Track Report”, *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 1437–1444, 2021.

[7] L. Chen, P. Chen, Z. Lin, “Artificial Intelligence in Education: A Review”, *IEEE Access*, vol. 8, pp. 75264–75278, 2020, doi:10.1109/ACCESS.2020.2988510.

[8] I. H. Sarker, “Machine Learning: Algorithms, Real-World Applications and Research Directions”, *SN Computer Science*, vol. 2, no. 3, pp. 160, 2021, doi:10.1007/s42979-021-00592-x.

[9] C. Tian et al., “Deep learning on image denoising: An overview”, *Neural Networks*, vol. 131, pp. 251–275, 2020, doi:10.1016/j.neunet.2020.07.025.

[10] G. Guo, N. Zhang, “A survey on deep learning based face recognition”, *Computer Vision and Image Understanding*, vol. 189, pp. 102805, 2019, doi:10.1016/j.cviu.2019.102805.

[11] D. Schofield et al., “Chimpanzee face recognition from videos in the wild using deep learning”, *Science Advances*, vol. 5, no. 9, 2019, doi:10.1126/sciadv.aaw0736.

[12] A. H. S. Ganidisastra, Y. Bandung, “An Incremental Training on Deep Learning Face Recognition for M-Learning Online Exam Proctoring”, *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, pp. 213–219, 2021, doi:10.1109/APWiMob51111.2021.9435232.