PAPER • OPEN ACCESS

Face Recognition and Identification using Deep Learning Approach

To cite this article: KH Teoh *et al* 2021 *J. Phys.: Conf. Ser.* **1755** 012006

View the article online for updates and enhancements.

Face Recognition and Identification using Deep Learning Approach

KH Teoh[2], RC Ismail[1,2], SZM Naziri[2], R Hussin[2], MNM Isa[2] and MSSM Basir[3]

[1]School of Computing and Informatics, Albukhary International University, Kedah Malaysia.
[2]School of Microelectronic Engineering, Universiti Malaysia Perlis, Perlis, Malaysia.
[3]Department of Electrical Engineering, Politeknik Mukah, Sarawak, Malaysia.

rizalafande.ismail@aiu.edu.my

Abstract. Human face is the significant characteristic to identify a person. Everyone has their own unique face even for twins. Thus, a face recognition and identification are required to distinguish each other. A face recognition system is the verification system to find a person's identity through biometric method. Face recognition has become a popular method nowadays in many applications such as phone unlock system, criminal identification and even home security system. This system is more secure as it does not need any dependencies such as key and card but only facial image is needed. Generally, human recognition system involves 2 phases which are face detection and face identification. This paper describes the concept on how to design and develop a face recognition system through deep learning using OpenCV in python. Deep learning is an approach to perform the face recognition and seems to be an adequate method to carry out face recognition due to its high accuracy. Experimental results are provided to demonstrate the accuracy of the proposed face recognition system.

1. Introduction

In recent years, artificial intelligence [1] is developing in a rapid manner. Nowadays, it can be seen that the invention of self-driving car [2] or self-service supermarket have been introduced. Artificial intelligence is closely linked to computer vision. Humans, by utilizing vision to adapt and to understand the environments where they are surrounded with, whilst computer vision is working on duplication of human vision but in electronically to perceive and interpret an image. Computer vision is not only working as an eye to see but it needs to react. It needs to be empowered with the ability to detect, identify and process the images perceived like what human vision done. For example, when a person walks to the path of a driving car, driver must react quickly and take an action. From the view of driver's vision, his brain undergoes three major steps, identifying, processing and deciding and this is the aim of computer vision to perform same operations efficiently. However, vision is a fundamental element of intelligence. Vision consists of many components, including coordination, memory, retrieval, reasoning, estimation, recognition and more. System with only one of this ability is not qualified as a vision. Computer vision is actually mimicking human systems. Since our world is in three dimension but our visual sensor usually provide only two dimensional images which it increases the difficulty for computer to analyze an object in 3D [3].

2. Overview

2.1. Face Recognition

Face recognition is part of computer vision. Face recognition [4] is used to identifying a person in biometric method based on image on their face. A person is identified through biological traits. Human eyes can easily recognize people by simply looking at them but the concentration span for human eyes has its limit. Hence, a computerized method is invented to perform face recognition. Face recognition [5] includes the operations of automatically detecting followed by verifying a person from either picture or video. Although face recognition has been researched extensively [6-10] but there are still challenges to overcome several issues such as:

- Misalignment
- Pose Variation
- Illumination Variation
- Expression Variation

Multiple approaches have to be tested to improve the accuracy and degree of precision of the face recognition.

2.2. Deep Learning

Deep learning [11] is a product of development of artificial neural network. At the beginning, practice of training MLPs (Multi-layers Perceptron) where a linear layer is added from input of network connection to that of output [12]. Subsequently, G. Thomson [13] had proposed a new idea known as deep learning, where it is a new model training as shown in Figure 1.
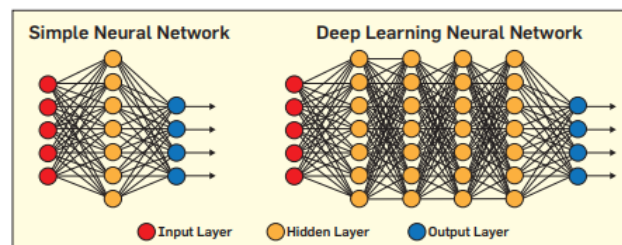

Figure 1. Simple neural network vs deep learning neural network [14].

Deep learning can achieve a nice approximation of a complex function through increment of hidden layers, hence, it is capable to achieve astonish result in the face recognition. It is a part of machine language that teaches computer to do as what human does naturally. Thus, deep learning is chosen to be implemented in this paper.

2.3. Interface

S.Matuska et al. [15] did the comparison of the speed between OpenCV and Matlab. Basic algorithm of image processing is presented and the time consumption in OpenCV and Matlab is the main focus point as presented in Figure 2. It is experimented that OpenCV is much faster than Matlab up to 30 times and can be up to 100 times for Erosion algorithm.
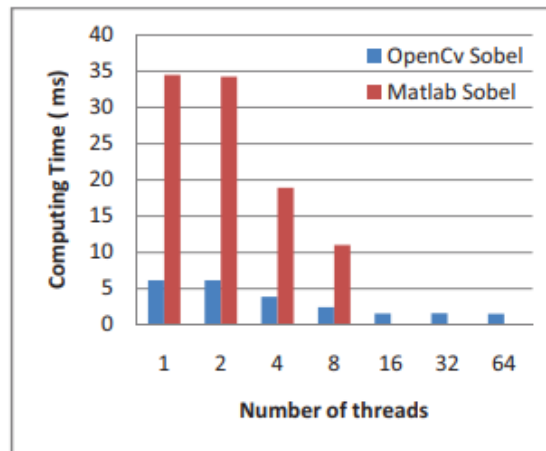
Figure 2. Time consumption between OpenCV and Matlab [16].

However, Matlab environment is simpler and more user-friendly as it provides various function and algorithm. Memory allocation and memory leak are not the issues in Matlab but they are significant task for OpenCV.

Matlab is written in high level language and it is built on Java while Java is built upon C language. Hence, when a Matlab program is ran, the computer takes time to interpret the coding. It then turns them into Java Code and lastly executes the code. In contrary, OpenCV is generally a library written in C language. Operations taken are major in processing but not interpreting. Thus, program in OpenCV can runs faster than that in Matlab.

Furthermore, OpenCV can be said more efficient than Matlab. Matlab wasted system resource as it uses them excessively as to ensure that memory allocation and memory leak will not be problem. However, in modern computer, the RAM element is not a point to be concerned. Therefore, in general, OpenCV runs faster than Matlab and it is the most comprehensive open source library for computer vision. Besides, it has large user community which means that more guidance can be obtained from various inputs. In contrary, Matlab is not open source and quite expensive to be obtained. Therefore, OpenCV is chosen to be applied in this paper.

3. Design Methodology

3.1. Face Recognition

Figure 3 shows the flowchart of the face recognition steps. To do face recognition, there must be an input to be detected and verified. Hence, an image sensor or typically a camera has to be set up for recording or capturing images. The camera should be compatible with the software used. The next step is the input image. The input can be images and recorded video or real-time video. After the input is provided, faces in the images or videos are to be detected. When the classifier is trained, it can be utilized to start to recognition work. It can be used in either video or image to recognize one or more person. Different set of python scripts are provided to run the different type of recognition. The python script will import the classifier that is trained in previous step in order to carry out the recognition for the person from the camera or from an image.

Figure 3. Flow chart of face recognition.

In face detection, Haar feature-based cascade classifiers is used and the classifier used is Haar Cascade for frontal face. A Haar Cascade is basically a classifier which is used to detect the object for which it has been trained for, from the source. The Haar Cascade is performed by superimposing the positive image over a set of negative images. The training is generally done on a server and on various stages. Better results are obtained by using high quality images and increasing the amount of stages for which the classifier is trained.

TensorFlow is the framework that is being used in the system classifier section. Classifier is trained and used in the recognition process. The training process takes a long time to achieve a better classifier. The longer the time of the training runs, the better the classifier is. In the proposed face recognition system, the training period taken is 3 days. If the training is allowed to run longer, the loss can be reduced further and hence the accuracy can be increased.

3.2. Accuracy

The accuracy of the system will be tested via recognition of three peoples with multiple times at different locations, mainly to test how light intensity affect the accuracy of the system. The accuracy is verified using confusion matrix. The calculation is based on (1).

$$((TN + TP) / Total) \times 100\% \tag{1}$$

where TN is true negative while TP is true positive.

4. Results

4.1. Distance of Face Detection

Figure 4 and 5 show the distance between face and camera is affecting the recognition process. When the distance is close or less than 60cm, the proposed system can barely detect the face. In the other hand, when the distance is extended to more than 60cm, the recognition takes place.
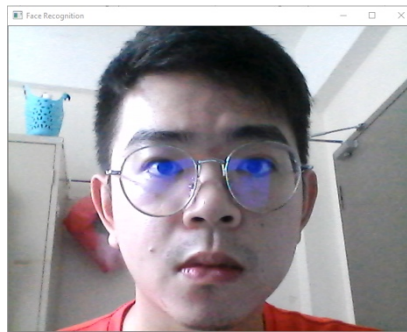
Figure 4. Distance between face and camera <60cm.
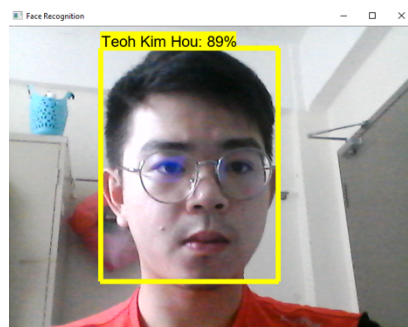

Figure 5. Distance between face and camera >60cm.

4.2. Lighting Condition

Figure 6 and 7 show the accuracy of the system is different depending on light intensity. It is clearly shown that high lighting intensity provides better accuracy compared to that of low lighting intensity. However, the proposed system shows true recognition even in darker surroundings.
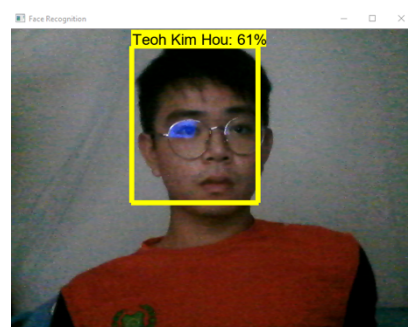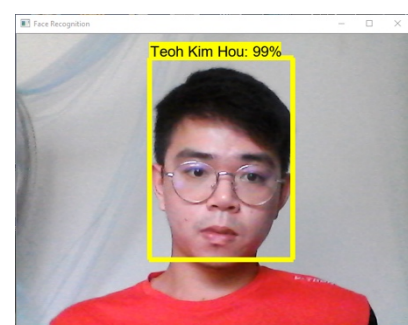

Figure 6. Low lighting intensity.

Figure 7. High lighting intensity.

4.3. Accuracy of Face Recognition based on Image

Multiple photos either in group or individual are loaded into system to verify the accuracy. A person should had appeared in those photos for 20 times. When the photos are all tested with the proposed face recognition system, the data is computed in confusion matrix to calculate the accuracy of the system.

From Table 1, it can be observed that the true and false recognition done by the proposed face recognition system. For the first person, 17 out of 20 recognitions are true. The true statement means that the identity of person in the photo that is recognized by the system is matched with real identity of the person. For second person, 18 of his photos are recognized correctly while for the last person, all photos are recognized correctly. Thus, the accuracy of the system can be calculated.

Table 1. Confusion matrix for image recognition.

Number of Face Recognition			Result
Person 1	Person 2	Person 3	
17	1	2	Person 1
1	18	1	Person 2
0	0	20	Person 3

Figure 8 shows the accuracy obtained with respect to each person. The overall accuracy of the system for face recognition from images is 91.7%.
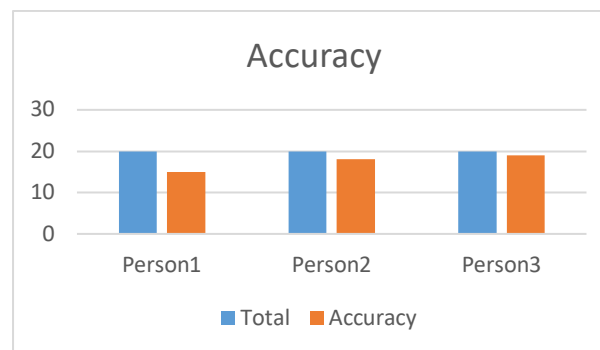

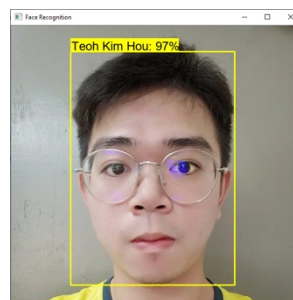**Figure 8.** Total accuracy per person based on image.

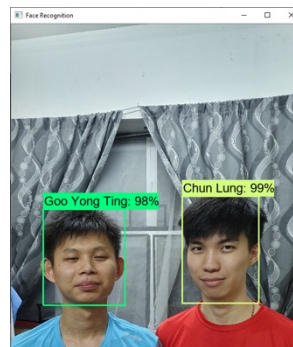**Figure 9.** Sample of testing Person 1 (true).

Figure 10. Sample of testing Person 2 (left) and 3 (right).

4.4. Accuracy of Face Recognition based on Real-time Video

Figure 11 shows sample of the screenshot of the face recognition done in real-time video. The candidates will show themselves to the webcam to undergo recognition where each of them has to do it 20 times at different location. When collection of data is done, the data is computed into confusion matrix as shown in Table 2.

Table 2. Confusion matrix for real-time video recognition.

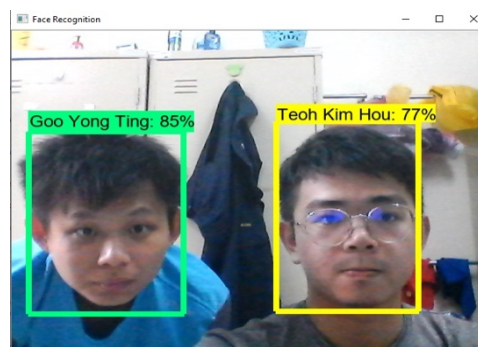| Number of Face Recognition			Result
Person 1	Person 2	Person 3	
15	1	4	Person 1
1	18	1	Person 2
1	0	19	Person 3

Figure 11. Sample of real-time video (true).

From Table 2, Person 1 has 15 true recognitions out of total of 20 recognition process. Person 3 has a quite high probability to be recognised as compared with Person 1 by the proposed system. For Person 2, the true recognitions are 18 out of 20. The overall accuracy of the face recognition on real-time video is 86.7%. Figure 12 shows the accuracy obtained with respect to each person.
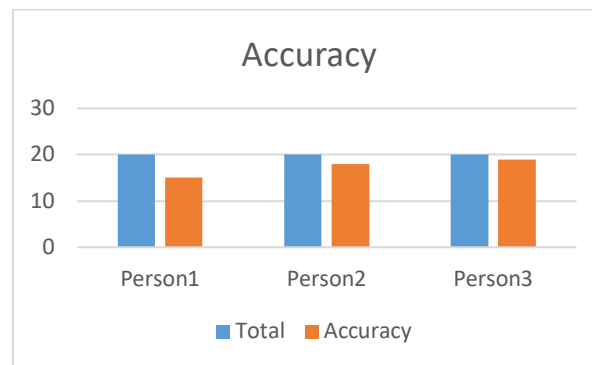
Figure 12. Total accuracy per person based on real-time video.

5. Discussion

From the results, it can be concluded that the accuracy of face recognition on image is higher than the accuracy on the real-time video. It can be observed that the resolution of image is a lot better than the real-time video. Due to the limitation of processing system, the fps of the video is low and causing the face captured is not as clear as in the photo. Hence, the classifier tends to assume that a particular person exists the characteristic of the other person. Therefore, the accuracy of the real-time video recognition has been reduced to 86.7%.

This model is trained using large number of images per candidate and using CNN approach. This led to huge dataset and improve the overall accuracy. By analysing the results, the light conditions can be seen as a factor that influences the recognition process. The recognition system tends to do false recognition when the light intensity is low. This could potentially be corrected by adding more training images which are captured in low light intensity in generating the face classifier.

For both recognitions either in image or real-time video, the accuracy for Person 3 is higher compared to the other two. The reason is probably related to the classifier. The classifier is more sensitive in recognizing Person 3. This might due to photos of Person 3 contain more variation such as different orientation of face and different background or lighting. Thus, the classifier can recognise Person 3 in more efficient manner.

6. Conclusion

In this paper, a face recognition and identification system is designed and developed using deep learning approach. The overall procedure of developing this face recognition system from training the data using CNN approach to face recognition is described. It is verified that with the large number of face images being trained into a classifier can achieve accuracy of 91.7% in recognising image and 86.7% in real-time video. There are few factors that can affect the accuracy of the system. When the light intensity is insufficient, the accuracy is relatively low compared to higher light intensity. Other than that, classifier is the main element in the recognition process. The longer the classifier is trained, the better the classifier is performed. The images that are used to train the classifier must be in variety of conditions in order to generate a robust classifier.

References

[1] J. Copeland, *Artificial Intelligence: A Philosophical Introduction*, 1st ed. Massachusetts: Blackwell Publishers, 1993.
[2] E.Lance and E.Michael, *Autonomous Vehicle Driverless Self-Driving Cars and Artificial Intelligence: Practical Advances in AI and Machine Learning*, 1st ed. LBE Press Publishing, 2014.

[3]   S.Milan, H.Vaclav, and B.Roger, *Image Processing, Analysis, and Machine Vision*, 4th ed. Cengage Learning, 2014.

[4]   G. Thomson, "Facial Recognition," *Encyclopedia*, 2005. [Online]. Available: https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/facial-recognition. [Accessed: 11-Oct-2018].

[5]   M. Kafai, L. An, and B. Bhanu, "Reference face graph for face recognition," *IEEE Trans. Inf. Forensics Secur.*, vol. 9, no. 12, pp. 2132–2143, 2014.

[6]   S. Zhaoqing, Z. Su, and L. I. Zhicheng, "Face Images Recognition Research Based on Smooth Filter and Support Vector Machine *," pp. 2760–2764, 2010.

[7]   N. K. Balcoh, M. H. Yousaf, W. Ahmad, and M. I. Baig, "Algorithm for Efficient Attendance Management: Face Recognition based approach," vol. 9, no. 4, pp. 146–150, 2012.

[8]   J. G. Roshantharanga, S. M. S. C. Samarakoon, T. A. P. Karunarathne, K. L. P. M. Liyanage, M. P. A. W. Gamage, and D. Perera, "Smart Attendance Using Real Time Face Recognition (Smart - Fr)," *SAITM Res. Symp. Eng. Adv. 2013*, vol. 2013, pp. 41–44, 2013.

[9]   M. Anggo and La Arapu, "Face Recognition Using Fisherface Method," *J. Phys. Conf. Ser.*, vol. 1028, no. 1, 2018.

[10] M. Arsenovic, S. Sladojevic, A. Anderla, and D. Stefanovic, "FaceTime - Deep learning based face recognition attendance system," *SISY 2017 - IEEE 15th Int. Symp. Intell. Syst. Informatics, Proc.*, pp. 53–57, 2017.

[11] Q. Wu, Y. Liu, Q. Li, S. Jin, and F. Li, "The application of deep learning in computer vision," *Proc. - 2017 Chinese Autom. Congr. CAC 2017*, vol. 2017–Janua, pp. 6522–6527, 2017.

[12] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Conf. Comput. Vis. Pattern Recognit.*, Dec. 2015.

[13] Stanko I. The Architectures of Geoffrey Hinton. In: Skansi S. (eds) Guide to Deep Learning Basics. Springer, Cham, 2020.

[14] C. Edwards, "Deep learning hunts for signals among the noise," *Commun. ACM*, vol. 61, no. 6, pp. 13–14, 2018.

[15] S. Matuska, R. Hudec, and M. Benco, "The Comparison of CPU Time Consumption for Image Processing Algorithm in Matlab and OpenCV," pp. 75–78, 2012

[16] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning Face Representation from Scratch," 2014.