International Journal of Mathematics and
Computer Science, **19**(2024), no. 1, 163–169

An analysis of a hybrid algorithm for face detection and emotion recognition

Dina Mahdi Abdulhussein, Laith Jasim Saud

Control and Systems Engineering Department
University of Technology-Iraq
Baghdad, Iraq

email: Cse.19.10@grad.uotechnology.edu.iq,
laith.j.saud@yotechnology.edu.iq

(Received July 1, 2023, Accepted August 1, 2023,
Published August 31, 2023)

Abstract

Detecting and analyzing human face represent a necessity in many areas such as medicine, security, education, human-computer interaction, and much more. Face emotions recognition (FER) is one of the important aspects of face image analysis that is needed in many applications and technologies. In general, the research work in this area concentrates on enhancing the detection accuracy and decreasing the time needed for detection. In this work, the FER hybrid algorithm based on V-J (Viola-Jones) and CNN algorithms is tested, analyzed, and evaluated considering different aspects among which are the principle of work and factors involved and affecting it, the efficiency, the prospect of improving, and the possibility of exchanging the role between the V-J and CNN algorithms.

1 Introduction

Face detection used to find the human face location and size in a digital image [1]. Essentially, face detection are used by all facial analysis algorithms such as, face enhancement [2], face liveliness detection [3], gender and age

Key words and phrases: face emotion recognition, face detection, CNN, Viola-Jones algorithm.
AMS Subject Classifications: 68U10.
ISSN 1814-0432, 2024, http://ijmcs.future-in-tech.net

detection [4], face emotion recognition, face identification [5], face recognition [6] and [7], face alignment, face feature extraction [8]-[11] and many more. In this work, face detection is used as an essential step to accomplish facial emotion recognition (FER), which recently has attracted growing attention. FER is an important technology that is used to analyze facial expressions from both images and videos [12] to reveal information on human emotional state. Emotional aspects is essential for social communications and decision making. It can be widely applied to various areas, such as mental diseases diagnosis and Human-Machine Interaction [13]. Facial emotion recognition technology still face and produce challenges when classifying people's emotions, such as the variations in age, gender, and ethnicity, in addition to the variations in image quality [14]. Moreover, the system always links facial expressions automatically to certain emotions and so it cannot distinguish which ones are genuine and which are not.

2 Tests

In this work, we concentrate on testing and analyzing both human face detection and emotion recognition. The facial emotion recognition process mainly consists of three phases: detecting the face, extracting face features, and classifying face expression. In this work, face is detected by using Viola-Jones algorithm, while facial feature extraction and expression classification is obtained by the convolution neural network.

In 2001, Viola and Jones [15] were the first to introduce their Viola-Jones algorithm for frontal face detection. It mainly consist of four steps: Haar-like features, integral image, Adaboost, and cascade classifier. It is implemented using Python software with the aid of OpenCV and Numpy libraries under PyCharm IDE.

In this work, we obtain tests implemented on face detection by adopting one of the trained classifier models of (OpenCV-HaarCascade) called ("haarcascade_frontalface_alt2.xml"). This classifier is already trained on two sets of images (positive and negative) and it is ready to be used for face detection [16] and [15].

The first test is a preliminary test implemented on detecting faces in different single images with different characteristics. Some images are in GRAY color and others are in RGB color and they are of different resolutions. All input images contain frontal faces of various genders, ages, ethnicities, and expressions. Some images contain only single face while others contain sev-

eral faces. In addition, faces are in different orientations such as tilted faces and upside down faces.
The second test is implemented on a folder of multiple images. Here, the source folder path is specified once at first and then the samples are fed one by one. The used images are in different resolutions and they are either in GRAY or RGB color. An outcome destination folder is used to save all detected faces after cropping them.
The third test is implemented using samples of animal faces. Moreover, tests are implemented on samples of faces but after removing one or more features like eyes, eyebrows, lips, nose, and face border. Later, the algorithm is fed by samples of faces with abnormal (12) face expressions. Finally, the algorithm is implemented on thermal images of faces.
Convolutional Neural Networks serves the purpose of feature extraction and emotion recognition.
In this work, the CNN model is trained on emotion classes at first and then used for facial emotions classification. The CNN model is trained on commonly used real expression dataset, FER-2013. The model used in this work consists of (11) layers. The first four layers are similar in their structure. Each one of them consists of two convolution layers and one max pooling layer. The fifth and sixth layers are implemented by using flattening. Finally, the output layer is dense layer. The CNN model is implemented using Python software and with the help of Keras library and using GPU.
At first, the CNN model is trained on five classes (emotions) of FER 2013 and then the CNN model is trained on seven classes of FER2013. The aim of this test is to see the impact of the class's number on the CNN classification performance. After that, the CNN model is trained on face detection; i.e., on only two classes (faces and not faces). The faces' images consist of only one human face while the non-faces' images consist of different topics such as cars and nature (but does not contain any human face). The training samples are (15300) and the testing samples are (14000). Then the last tests implemented on the hybrid algorithm (Viola-Jones + CNN) using different face images from the World Wide Web to detect faces and their emotions.

3 Results

The results of Viola-Jones tests show that face border is an important feature in face detection. As shown in the figure (1), faces without borders are not detected. In addition, face detection becomes difficult when two features are

Figure 1: Faces without borders.

Figure 2: Faces with abnormal expression.

missing at the same time such as, eyes and eyebrows, nose and eyebrows, or nose and lips, and of course the case worsen with three features missing.
Other test results show that sometimes Viola-Jones is unable to detect faces with abnormal expressions as shown in figure (2).
Moreover, from the implemented tests, we noticed that some images of animal faces and specifically monkey faces are detected and/or located falsely.
Consequently, results on thermal images show that sometimes colored or gray thermal images are detected as shown in figure (4).
Some advantages and disadvantages of Viola-Jones algorithm are listed in Table 1:

Table 1: Advantages and disadvantages of Viola-Jones Algorithm.

Regarding emotion recognition by CNN, tests results show that CNN can be used in detecting and classifying objects. The CNN model structure is divided into two main parts: the feature extracting part and the classification part. This indeed gives more flexibility to design different models. Moreover, the efficiency of emotion classification by CNN is decreased when categories (classes) number increase and they overlap. For instance, the used CNN

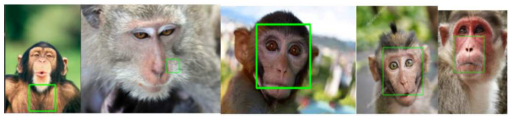

Figure 3: Faces of animals (monkeys).

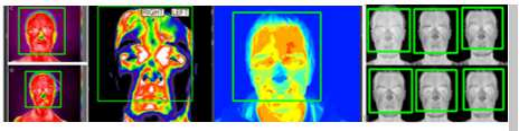

Figure 4: Tthermal images of faces.

No.	Advantages	Disadvantages
1	Simple and fast in real time.	Specific for frontal face detection.
2	Low computation requirements.	Training time is very long.
3	Can be used to any object detection.	Restricted to binary classification.
4	It is an ensemble classifier.	Highly prone to false-positive detection.

model is implemented in three cases: two, five, and seven classes. We noticed that the obtained accuracy is 98%, 73%, and 63%, respectively. The test results also show that the efficiency affected by other conditions such as the size and richness of the training dataset, face emotions with low expression intensity, image with low resolution and high variation in illumination, and some images of faces that contain mixed emotions. Some advantages and disadvantages of CNN are listed in Table 2:

Table 2: Advantages and disadvantages of CNN.

4 Conclusions

This work was based on tests and analyzed a hybrid algorithm which combined two important algorithms: One was used for face detecting (Viola-Jones) and the other was used for emotion recognition (CNN). This hybrid

No.	Advantages	Disadvantages
1	Extracts features automatically.	Requires large and rich dataset for training.
2	High accuracy in image recognition.	Unable to realize the object position and orientation.
3	Less time consumed for training comparing to Viola-Jones.	Experience and effort are required to improve specific cases of CNN models.

algorithm considered was a highly efficient algorithm that can be used for a system intended for detection and classification purposes together.
The implemented tests shows that V-J algorithm excelled in the object detecting purposes, while convolutional neural network excelled in the analysis purposes such as classification and recognition of human emotions. The Viola-Jones algorithm was not suitable for classifying human emotions because it was a binary classifier; i.e., it was used to detect the existence and non-existence of an object. On the other hand, CNN was not suitable for detecting human face because it was incapable of realizing the location and the orientation of the objects within images.

References

[1] K. Hasan, S. Ahsan, Abdullah Al-Mamun, S. H. S. Newaz, and G. M. Lee, Human face detection techniques: A comprehensive review and future research directions, Electron., **10**, no. 19, (2021), 2354.

[2] I. Abdul-Jabbarr, Face Image Enhancement using Wavelet Denoising and Gabor Filters, Iraqi J. Comput. Commun. Control Syst. Engg., **16**, no. 1, (2016), 104–117.

[3] E. A. Raheem, S. M. S. Ahmad, W. A. W. Adnan, Insight on face liveness detection, A systematic literature review, Int. J. Electr. Comput. Engg., **9**, no. 6, (2019), 5165–5175.

[4] S. Hameed, Human Gender and Age Detection Based on Attributes of Face, Int. J. Interact. Mob. Tech., (2022).

[5] S. Hamandi, A. M. Rahma, R. Hassan, Multi-Spectral Hybrid Invariant Moments Fusion Technique for Face Identification, The International Arab Journal of Information Technology, Computer Science Department, University of Technology, Iraq, **18**, no. 3A, (2021), 9.

[6] F. F. A. Bashra Kadhim Oleiwi, Smart E-Attendance System Utilizing Eigenfaces Algorithm, Iraqi J. Comput. Commun. Control Syst. Engg., **18**, no. 1, (2018), 56–63.

[7] S. L. Galib, F. S. Tahir, A. A. Abdulrahman, Detection Face Parts in Image Using Neural Network Based on MATLAB, Engineering and Technology Journal, **39**, no. 1B. (2021), 159–164.

[8] A. Albakri, S. Almamory, H. Alfartosy, Feature-Based Face Detection: a Survey, Iraqi J. Comput. Informatics, **44**, no. 1, (2018), 18–23.

[9] A. T. Hashim, D. A. Noori, An Approach of Noisy Color Iris Segmentation Based on Hybrid Image Processing Techniques, Proc. Int. Conf. Cyberworlds, (2016), 183–188.

[10] H. Abdelwahed, A. Hashim, A. Hasan, Segmentation Approach for a Noisy Iris Images Based on Hybrid Techniques, Engg. Tech. J., **38**, no. 11, (2020), 1684–1691.

[11] S. I. Mohammed, N. A. Jaafar, K. M. Hussien, Face Recognition Based on Viola-Jones Face Detection Method and Principle Component Analysis, Iraqi J. Comput. Commun. Control Syst. Engg., **18**, no. 3, (2018), 52–59.

[12] W. H. Abdulsalam, R. S. Alhamdani, M. N. Abdullah, Facial emotion recognition from videos using deep convolutional neural networks, International Journal of Machine Learning and Computing, **9**, no. 1. (2019), 14–19.

[13] N. Jain, S. K. A, A. Kumar, Hybrid deep neural networks for face emotion recognition, Pattern Recognit. Lett., (2018).

[14] A. R. Khan, Facial Emotion Recognition Using Conventional Machine Learning and Deep Learning Methods: Current Achievements, Analysis and Remaining Challenges, Inf., **13**, no. 6, (2022), 268.

[15] P. Viola, M. J. Jones, Rapid Object Detection using a Boosted Cascade of Simple Features, Conf. Comput. Vis. Pattern Recognit., **18**, no. 10, (2001), 9.

[16] P. Viola, M. J. Jones, Robust Real-Time Face Detection Intro to Face Detection, Int. J. Comput. Vis., **57**, no. 2, (2004), 137–154.