Journal of Network Communications and Emerging Technologies (JNCET)
Volume 12, Issue 8, August (2022)

Face Emotion Recognition Using Machine Learning and Transfer Learning Method

A.A. Alfaridi
Computer Engineering-Fahd Bin Sultan University, KSA
alforaidialaa@yahoo.com

Nazar Elfadil
Computer Engineering-Fahd Bin Sultan University, KSA
nfadel@fbsu.edu.sa

ABSTRACT

The detection of human emotions is considered one of the important aspects of security and business promotions. In security matters, the identification of human emotions is very important, especially in vulnerable areas. In the business, emotion recognition is used in the promotion of the products and the analysis of customer satisfaction. This project aims to develop a face emotion system based on machine learning and transfer learning. The result of the implemented method shows that the prediction of the emotion performed successfully with a percentage equal to 77%.

Index Terms – Emotion, Transfer Learning, Tensor Flow, Machine Learning.

1. INTRODUCTION

Recently the detection of human emotions implemented in many areas that take care of the security issues or the business promotions. In the security aspects the identification of any persons is very important especially in the vulnerable areas such as the airport, the concerts and the public areas. On the other hand, the emotion detection plays an important role in the promotion of the products and in the analysis of the customer satisfaction [1, 2, 3, 4, and 5] regarding that product. The human face considered the most emotional portion of the human body, it allows you to understand the inner state of the human. Accordingly, the human face expressions play an important role in non-verbal communication. The first studies on the human face expressions are written by Darwin [6]. He said the facial expressions cannot be learned and it has a meaning for survival.

2. LITERATURE REVIEW

In general, the system architecture for the face emotion recognition is illustrated in Fig.1. This system architecture consists of input, pre-processing, feature extraction, faces database & classification and used by many authors in the literature. The model is trained for making the predictions for the new input images. The pre-processing is used before the feature extraction to do the contrast adjustment, scaling, and clarity of the image [7, 8, and 9].

2.1. Face Detection Techniques

Face detection technique is used to remove the unwanted area and eliminate it from further processing. In the literature Vila-Jones and Histogram of Oriented Gradients (HOG) are discussed. Viola-Jones is used to the facial region by using Haar features which is consist of two rectangles, three rectangles, and four rectangles [10, 11, and 12].

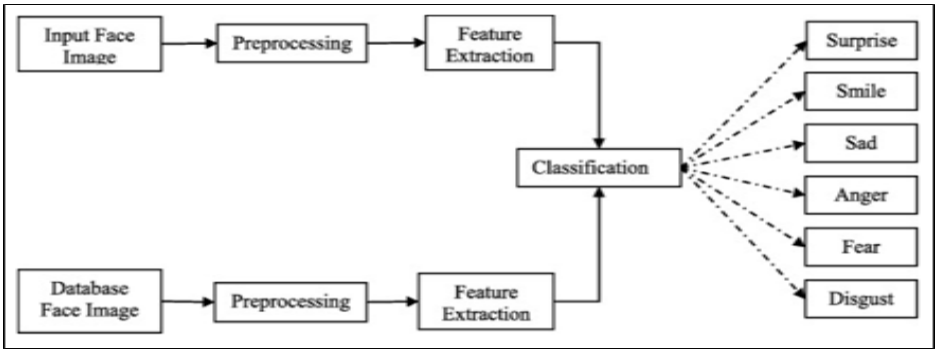


Fig. 1: face emotion recognition system architecture

These Haar values are used for image classification by subtracting the pixels in the white rectangle from the pixels in the black rectangle. Fig. 2 shows an example of Haar features [11, and 21].

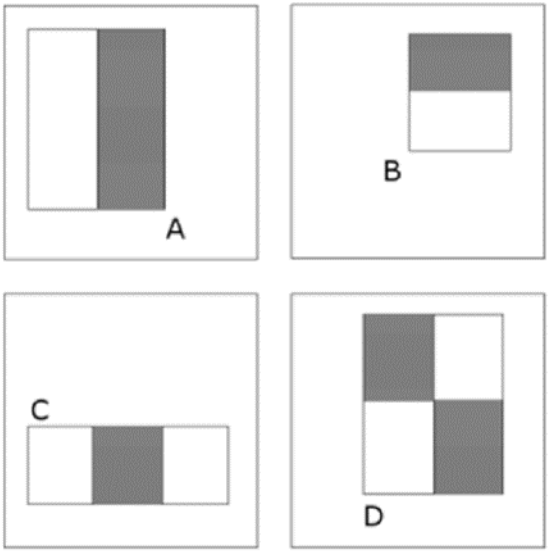


Fig. 2: Example of Haar feature [11]

On the other hand, HOG algorithm goes through multiple steps. The first one is the normalization of the color space. The second step is to find out the gradients for every pixel in x & y directions in n x n cell. In the third step, the algorithm performs the construction of the cell histogram. The next step is to group several n x n pixels in N x N blocks to find out the contrast normalization. These blocks are named Histogram of Oriented Gradient (HoG). After this step the vectors constructed by concatenating the descriptors of HoG in the detection window. In the last step the feature classified by the SVM [10].

2.2. Feature Extraction Methods

In machine learning [13] and image processing the feature extraction performed to build a derived informative value to facilitate the learning process and to reduce the dimension of the data [14]. However, the feature extraction helps to find the important points in the image. In the face emotion detection, the purpose of the feature extraction is to isolate the important point on the detected face. Ensembles of Regression Tree algorithm is used to find out the feature points on the face. Another algorithm called Local Binary Patterns (LBP) is also used to get the feature of the face [15].

2.3. The Classifier & Emotion Detection

In the literature the neural network classifier is used. The classifier consist of input layer, hidden layer and output layer consists of seven layers each one describes one of the emotions [16]. Fig. 3 shows the neural network model [17].

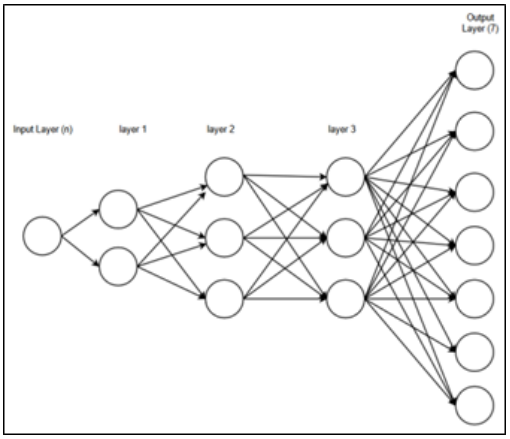


Fig. 3: Neural Network Model [17]

The author of the literature determines the classification rate during the testing phase as the training and the testing follow the same classification steps and they ran the convolutional neural network model three time with different iterations [9]. On the other hand, they also proposed another method by using the distances between facial landmark and support vector machine (SVM) to classify the emotions [18, 20, and 21].

3. PROJECT METHODOLOGY

In this work an experimental methodology used to perform the face emotion detection. In the beginning, the face detection algorithms, the feature extraction methods & the machine learning classifier were studied. The detection of face emotions implemented by using python with Dlib & OpenCV libraries [16, and 17]. Fig.4 shows the method in this project.

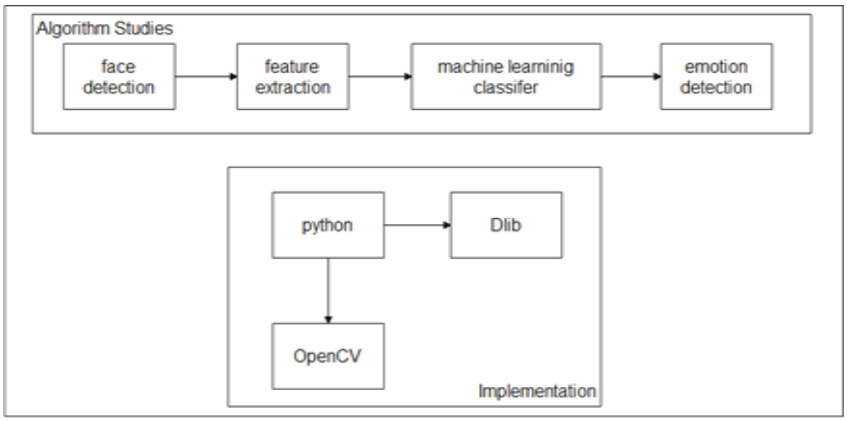


Fig. 4: Project Method

3.1. Python 3

Python is a high-level object-oriented scripting language Python is widely used in image processing and computer vision projects due to the libraries it provides in this field such as OpenCV & NumPy. Behind that, Python libraries support artificial intelligent, machine learning, and deep learning algorithms as well [16].

3.2. OpenCV

OpenCV is an abbreviation for (Open-Source Computer Vision). It is a free library that is used to process images & videos OpenCV was created to afford a global infrastructure for computer vision applications. Besides that, OpenCV introduces the use of machine learning in commercial applications. In this project, OpenCV along with Python is used mainly to detect the faces using haar cascade method and to printout on the images the type of the detected emotion [16].

3.3. Tensor Flow

Tensor Flow is an open-source library for creating machine learning applications. It was developed by the Google team to use internally, they released it in 2015 under Apache license 2.0. The architecture of Tensor Flow is flexible as it allows for the processing to be performed across multiple platforms on desktop or several clusters of servers [19].

3.4. Keras

Keras is an open-source library for the artificial neural network. However, Keras works as an interface for Tensor Flow. With Keras the implementation and the experiments of deep neural networks performed in a fast manner [20]. Keras covers the implementation of the most typically used neural network such as layers, activation function & optimizers. However, Keras has supported the convolutional neural network. So Keras allows users to design deep models [17].

3.5. MobileNet V2

MobileNetV2 is a convolutional neural network consist of 53 layers. Although it is intended for mobile devices, it can also be used with other types of devices. The bottleneck layers are connected to one another via the remaining connections, which are based on an inverted residual structure. The depth-wise convolutional features filters in the intermediary layers. Fig.5 shows the architecture of MobileNetV2 [18].

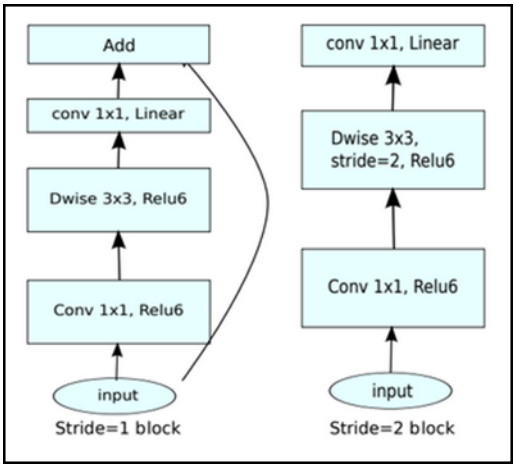


Fig. 5: MobileNet V2 architecture

3.6. Relu Activation Function

The purpose of the activation function is to transform the summation of the input weights to output. The rectified linear activation (ReLU) is a linear activation function that provides the output if the input is positive. However, ReLU is considered the default activation function in the neural network due to its high performance.

3.7. Transfer Learning

Machine learning model probably has a huge number of parameters or weights. Accordingly, to train the model from the beginning a large amount of data computing resources will be needed. Hence, transfer learning is a method used to take part in a model that has been trained and then reprocessing it to create a new model. This method will reduce the needed computing resources & time. in this method the improvement of learning is performed by a new learning task through the transfer of knowledge from a related task that has already been learned [14].

3.8. Dataset

A dataset is a collection of data related to a specific context. In machine learning projects the dataset is used to train the model to perform the needed tasks. Usually, three different datasets are used: training dataset, validation dataset & test dataset. In this project a dataset called FER 2013 is used, FER (Facial Expression Recognition) dataset was introduced in 2013. FER database was created using Google image search, the faces labelled with seven expressions (angry, disgust, fear, happy, neutral, sad, surprise). The training dataset contains 28709 images distributed between the expressions as illustrated in Fig.6.

Fig. 6: FER train dataset

4. IMPLEMENTATION

4.1. Preparing the Environment

The programming environment consists of Anaconda & Jupyter Notebook, both of them are installed. After that, tensor flow is installed through Anaconda Prompt & the pip (preferred installer program). The programming environment consists of Anaconda & Jupyter Notebook, both of them are installed. After that, tensor flow is installed through Anaconda Prompt & the pip (preferred installer program).

4.2. Preparing the Dataset

FER 2013 dataset downloaded from Kaggle.com. The total size of the dataset is around 60 MB. Besides that, the dataset is divided into train & test. Each class on the train or the test is grouped in a separate folder. To let the programming became easier the dataset groups renamed from 0 to 6 for example angry image group renamed 0 & disgust image group renamed 1 and so on.

4.3. Data Preprocessing

The purpose of data preprocessing is to prepare the data and organize the raw data to make it suitable for the machine learning training process. In another word, this is the process of converting the raw data to an understandable format. The first step in the data preprocessing is to read all dataset images and save them into an array. The second step is to change the size of all images in the dataset to (224 x 224) that's because the transfer learning algorithm that is used can accept only this size of images. The third step is to shuffle the images on the array to avoid the learning of the sequence. The fourth step is to separate the image array and the label array. The fifth step is the learning convolutional network that used in this project (MobileNetV2) can accept only a 4D array. The sixth step is to normalize the data, although there are libraries that can do the normalization such as scikit, for simplicity we did it by dividing the values by 255 as it is considered the maximum value. Fig.7 shows the flow chart that describes the steps of data preprocessing.

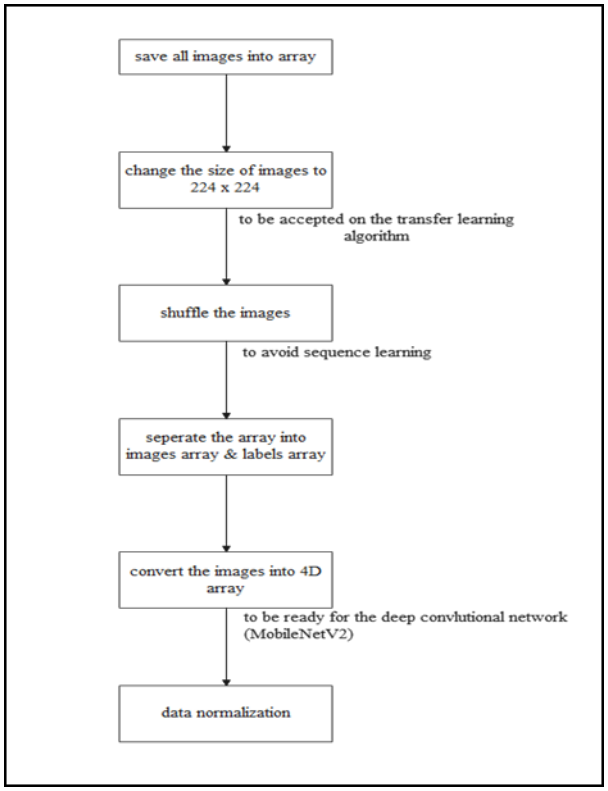


Fig. 7: data pre-processing steps

4.4. Transfer Learning Model

The input layer number 0 of MobileNetV2 is selected as input for the transfer learning model. The layer before the last one in MobileNetV2 was selected to be the output of the model, this layer is called (global_average_pooling2d). The output of this layer

taken into another final output layer with a dense of 128. The activation function that is used is (relu). However, as illustrated in Fig.8 another 64 layers & activation function used as well to get the final out which consists of 7 layers.

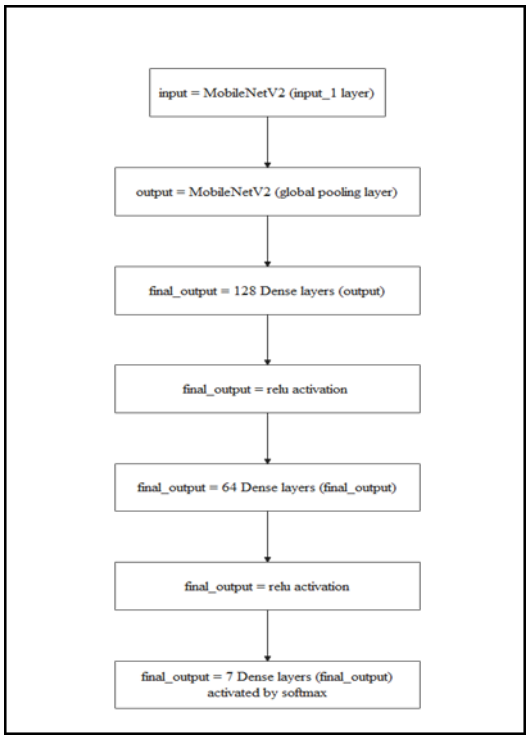


Fig. 8: Transfer learning steps

Finally, in order to create the model (keras.Model) function is used. The input in the flow chart is selected to be the input for function & the final output is selected to be the output in the function. The transfer learning started by implementing (fit) function with a number of epochs equals to 10. Each epoch needs around 43 minutes to finalize the needed calculations so the total need time to finalize the calculation is around 7 hours. Behind that, the transfer learning started with high accuracy such as 0.4388 which is not possible if the training started from the beginning. Once the training gets done it is saved to use in the process of the prediction.

5. RESULT

The performed test shows the success of the emotion detection by using the transfer learning method. The accuracy of the prediction model reached up to 77% after performing 10 epochs. The training was performed in seven hours on core-i7 with 16 GB RAMs. Fig.9 shows the output of the prediction model. Besides that, Fig.10 shows the epochs and its relation with the accuracy.

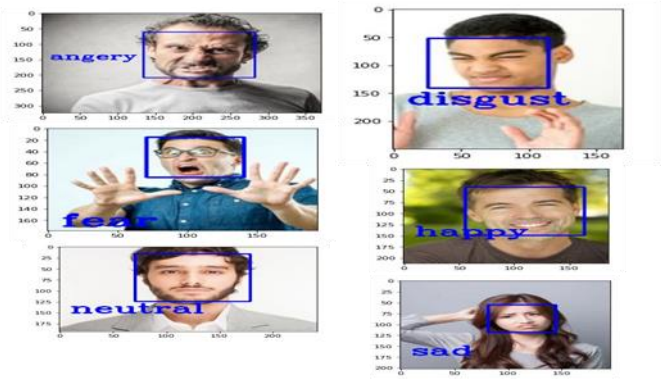


Fig. 9: detected faces based on the proposed system

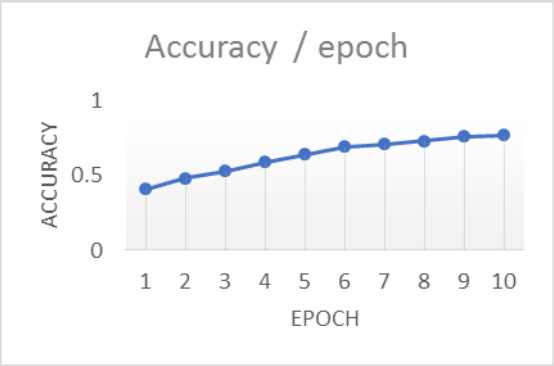


Fig. 10: the relation between the accuracy & epochs

The needed time to calculate & the number of epochs in this research is compared with the results in the literature. The researchers in [18] use the CNN and they reached up to 79% in terms of accuracy but after 500 epochs and in 16 hours which was needed too many iterations and a very long time when compared with the results in this research.

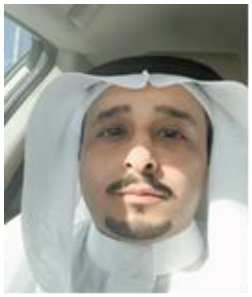
6. CONCLUSION

This paper aimed to recognize the different types of facial emotions. The recognition performed based on machine learning. The needed computing resources and time to learn the machine is reduced by using the transfer learning method which takes part of a model that is already trained and then reprocessing it to create a new model. Python 3 is the programming language that is used to perform this project and the OpenCV library with haar algorithm is used for the face detection. The application of machine learning is designed based on TensorFlow library to perform the data processing and to create and train the model. The neural network is performed in a faster manner by using Keras library which is working from the above layer based on TensorFlow to perform the optimization and the activation. However, the convolutional neural network works based on MobileNetV2. FER2013 is the dataset used to train the model. After 10 epochs and 7 hours of training, the accuracy of the prediction reached up to 77%.

REFERENCES

[1] J. J. Cristiano, J. K. Liker, and C. C. White, "Key factors in the successful application of quality function deployment (QFD)," IEEE Trans. Eng. Manage. , vol. 48, no. 1, pp. 81-95, 2001.
[2] S. M. C. Loureiro, Consumer-Brand Relationship: Foundation and State-of-the-Art, 2013.
[3] Bouzakraoui, Sadiq, Alaoui, "Appreciation of Customer Satisfaction through Analysis Facial Expressions and Emotions Recognition," in IEEE 4th World Conference on Complex Systems (WCCS), 2019.
[4] Devi Arumugam, Dr. S. Purushothaman B.E, "Emotion Classification Using Facial Expression," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 2, no. 7, 2011.
[5] O. C. Hans mark, M. Albinsoon, "Customer satisfaction and retention: The experiences of individual employees," Manage. Service Quality, vol. 14, no. 1, pp. 40-57, 2004.
[6] Charles Darwin, Paul Ekman, and Phillip Prodger, The expression of the emotions in man and animals, John Murray, 1872.
[7] Nitisha Raut, "Facial Emotion Recognition Using Machine Learning," 2018.
[8] Phavish Babajee, Geerish Suddul, Sandhya Armoogum, Ravi Foogooa, "Identifying Human Emotions from Facial Expressions with Deep Learning," Zooming Innovation in Consumer Electronics International Conference (ZINC) , 2020.
[9] A Uçar, Y Demir, C Güzeliş, "A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering," Neural Computing and Applications, 2016.
[10] Wout Swinkels, Luc Claesen, Feng Xiao, Haibin Shen, "SVM Point-based Real-time Emotion Detection," IEEE, 2017.
[11] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object Detection with Discriminatively Trained Part-Based Model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, p. 1627–1645, 2010.
[12] Sarangi, Susanta; Sahidullah, Md; Saha, Goutam, "Optimization of data-driven filter bank for automatic speaker verification," Digital Signal Processing - ELSEVIER, vol. 104, 2020.
[13] T. Ojala, M. Pietikäinen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971 - 987, 2002.
[14] Revina, M and Emmanuel, "A Survey on Human Face Expression Recognition Techniques," Dept. of Computer Science, Christian College, Sunadaranar University, India, 2018.
[15] Deepak Ghimire, Sung Hwan Jeong, Joonwhoan Lee, "Facial expression recognition based on local region specific features and support vector machines," Multimedia Tools and Applications, vol. 76, 2017.
[16] "Opencv," [Online]. Available: https://opencv.org/about/. [Accessed 24 4 2021].
[17] "mathworks," [Online]. Available: https://www.mathworks.com/help/deeplearning/ref/mobilenetv2.html. [Accessed 25 4 2021].
[18] [Online]. Available: https://paperswithcode.com/method/mobilenetv2. [Accessed 25 4 2021].
[19] Cootes, T., Edwards, G. and Taylor, C, "Active appearance models," IEEE Trans. Patternnal.Mach. Intel, vol. 23, no. 6, pp. 681-685, 2001.
[20] Revina, M and Emmanuel, "A Survey on Human Face Expression Recognition Techniques," Sunadaranar University, India, 2018.
[21] P. Viola and M. Jones, "Robust Real-time Object Detection," International Journal of Computer Vision, vol. 57, no. 2, pp. 137-154, 2001.

Authors



Ala Alfaridi graduated from Fahad Bin Sultan University in 2019 with a bachelor of science in computer engineering and again in 2020 with a master of science in computer engineering. He currently holds a computer engineer position at the ministry of education**.**

Nazar Elfadil earned a BSc (Hon.) in Engineering Technology in 1993 from Geizra University in Sudan. From University Technology of Malaysia, he earned a Master's and a PhD in Computer Science (network security) and Computer Engineering in 1997 and 2002, respectively. Up until 2003, Dr. Nazar worked as an assistant professor at Nottingham University Malaysia. He started working at Sultan Qaboos University in Oman's Electrical and Computer Engineering department in 2003. Since 2008, he has worked as an associate professor at Fahad Bin Sultan University. His areas of interest in study include artificial intelligence, network security and performance, machine learning, knowledge engineering, wireless networks, and machine learning.