Impact of Ethnic Group on Human Emotion Recognition Using Backpropagation Neural Network

Nabil M. Hewahi
Computer Science Department
Bahrain University, Bahrain
nhewahi@gmail.com

Abdul Rhman M. Baraka
Computer Science Department
Islamic University of Gaza, Palestine
ambwafa@hotmail.com

Abstract

We claim that knowing the ethnic group of human would increase the accuracy of the emotion recognition. This is due to the difference between the face appearances and expressions of various ethnic groups. To test our claim, we developed an approach based on Artificial Neural Networks (ANN) using backpropgation algorithm to recognize the human emotion through facial expressions. Our approach has been tested by using MSDEF dataset, and we found that, there is a positive effect on the accuracy of the recognition of emotion if we consider the ethnic group as input factor while building the recognition model. We achieved 8% improvement rate.

Keywords: Emotion Recognition, Artificial Neural Networks, Human Computer Interaction, Features Extraction.

1. Introduction

Obviously, the emotions is an important aspect in the interaction and communication between people, and the human can express these emotions through facial elements, speech, body movement, or etc.. Since the science of Artificial Intelligence (AI) is concerned with the automation of intelligent behavior, we need to improve the interaction between computer and human by recognition the human emotions by using face expressions [9] [12] [15] [20], speech [4] [16] or etc. [11]. In additional, appearance of emotions are universal across individuals as well as human ethnics and cultures[1] [2], but the way of emotion expression is that vary from one ethnic group to another and the differences cross-cultural in the familiarity effect on recognition accuracy for different types of expressions.

There are many of researches that focused on emotion recognition but regardless the ethnic group as factor in their models.

In this paper, we propose an emotion recognition approach using facial expression considering ethnic group based on Backpropagation neural network, to study the impact of ethnic group on accuracy of emotion recognition models.

In next section we focus on some related works. Section three describes our proposed approach. Section four discusses the implementation. Experiments and results are presented in section five. Last section is the conclusion.

2. Related Works

In the early 1990's the engineering community started to construct automatic methods of recognizing emotion from facial expression in an image and videos [12], and many of computer studies that focused on emotions recognition. Some studies use facial expressions to build a model for emotion recognition, and others use another elements or factors to build the models like voice, pulse, body movement and etc.

Matthew and Patterson [15] discuss a framework for the classification of emotional states, based on still images of the face, using active appearance model (AAM) and get distance using n

Euclidean's as features. To train and test the classifier they chose to use the facial expression database known as "FEEDTUM" - Facial Expressions and Emotion Database [8] , and seven basic emotions are used, happy, sad, angry, surprise, fear, disgust and natural state. The best results they obtained are in happy, natural and disgust emotions at the rate 93.3%, fear at 90.0%, and 79.7% in surprise, angry and sad at rate 63.9%. This study does not consider the ethnicity.

Raheja and Kumar [14] presented architecture for human gesture recognition, considering color image with different gestures by using backpropagation Artificial Neural Networks (ANN or NN). Four stages applied in the approach, face detection, image reprocessing, training network and recognition model. The pre-processing stage contains three methods, histogram equalization, edge detection, thinning and token generation. The ethnic group was not considered in this model. The model was trained using the three different gesture images, happy, sad and thinking expressions of faces. The model was tests with 100 images of three gestures, the results were 94.28% for happy, 85.71% for sad and 83.33% for thinking.

Karthigayan et.al [12] used Genetic Algorithms (GAs) and Artificial Neural Networks to build human emotion classifier. This classifier detects six human emotions Neutral, Sad, Anger, Happy, Fear, Disgust (or Dislike) and Surprise. They depend on two facial elements in the classifier, eyes and lips. By applying some preprocessing methods and edge detection, they extracted the eyes and lip regions, then extracted the features from these regions. Three feature extraction methods are applied, projection profile, contour profile and moments. The GA is applied to get the optimized values of the minor axes of an irregular ellipse corresponding to the lips and the minor axis of a regular ellipse related to eye by using a set of new fitness functions. Finally, they apply the results from GA on the ANNs model. Two architectures of ANNs models are proposed with an average of 10 trials of testing. The achieved results of 3x20x7 and 3x20x3 of ANN architecture were 85.13% and 83.57% of success rate respectively. The successful classification even goes to the maximum of about 91.42% in the ANN model of 3x20x7 structure. A South East Asian (SEA) race is only considered in this work.

3. Proposed Approach

Our proposed approach uses the face expression to detect the emotions through five steps that shows in Figure 1.

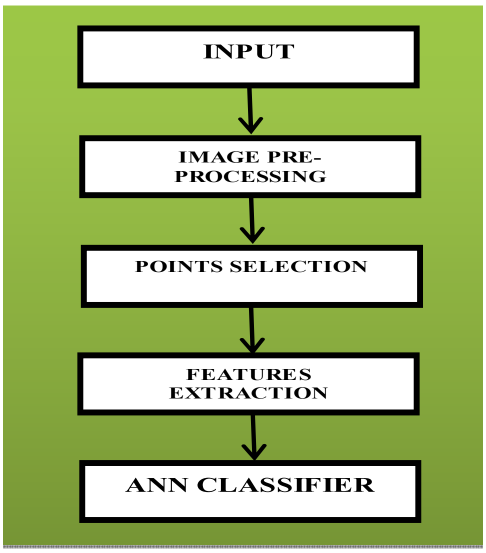


Figure 1. A proposed approach.

A. The Input

The input is an image of frontal human face which means that no bending the face to any angle and no rotated face to any side.

B. Image Reprocessing

In this stage we apply certain sub processes on the image to achieve a good image with one standard in terms of contrast and size. Because we depend on Euclidian distances to extract our features, the image size must be standardized. Also we apply a contrast process for each image to make the elements of the face distinguishable from other objects and also to be able to distinguish between background and the face. This will help the user to select all the points clearly. Simple algorithms are applied for resizing and contrast processes [5].

C. Point Selection

We chose 46 points which are distributed over human face image and use these points for features extraction. The choice of these points is to determine the shape of each element of the face (eyes, eyebrows and mouth), because the shape of these elements is changeable for each emotion, but these changes are different for each race as shown in Figure 2.

Figure. 2. Examples of Angry from different races. (A,B) African. (C,D) Asian. (E,F) Caucasian.

The number of points and the position of points are not standardized, but it is depending on the features that will be extracted, and used for the classifier. Many researches use various number of points and positions based on their view about the feature to be considered [13] [18] [19]. Figure 3 shows the points we used.

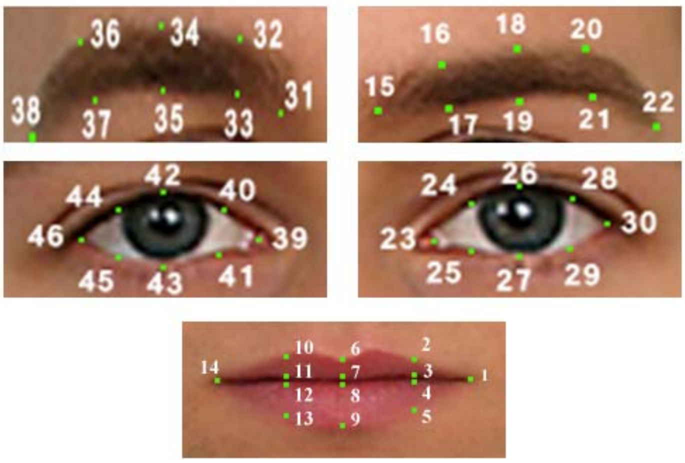


Figure 3. 46 points are selected on face elements to describe the emotions.

D. Features Extraction

While human emotion is changing, face expression is also changing. This means that, properties of face elements do change when face expression is being changed. As a consequence of this, the distances between points are changing. The distances between certain points as will be shown later describe the human emotion.

Based on what we stated above, we need to extract 28 features, which describe the distances between certain points explained in the previous stage, these features are classified into six groups, and each group describes the features of one face element. All features are a vertical distances between two points. Group one contains seven features for mouth, groups two and three contains 14 features for eyes, groups four and five contain six features for eyebrows, and the last group has one feature only which is the distance between the beginning of the eyebrow and the beginning of the eye in same side, this is significant (from point 23 to 15) because it is used to measure the distance of eyebrow from the eye. This feature is shown in Figure 4 by a line.

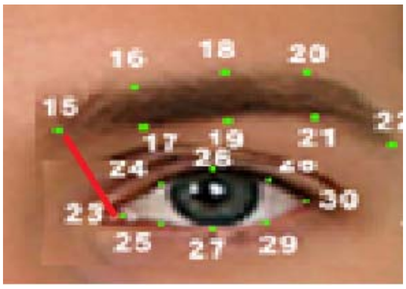


Figure 4. Distance between eyebrow and eye.

E. Classifier

For classification purpose of the emotions, we use ANN of supervised learning based on backpropagation algorithm. Backpropagation neural network architecture is used with its standards learning function with 28 inputs representing the extracted features and 6 outputs representing 6 emotions, happy, sad, angry, fear, shame and disgust. the emotions. We have also a hidden layer with 16 nodes selected after various trails to obtain the best results. The used ANN is depicted in Figure 5.

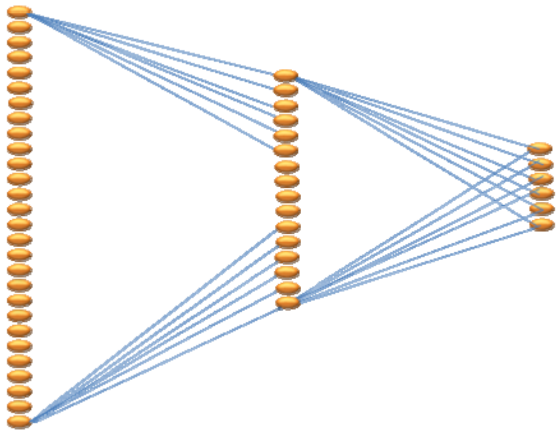


Figure 5. ANN Structure

4. Implementation

A. Face Features Extraction (FFE)

We used Microsoft Visual C#.net to develop a desktop application to apply the first of the four stages of our approach. This application is called Face Features Extraction (FFE), and it is used to perform four methods or functions as follow: Apply some pre-processing method, Selection of

Points, Features Extraction and save all input information for the classifier (Features, Ethnic group, Gender and emotion). Figure 6 shows the interface of FFE program.

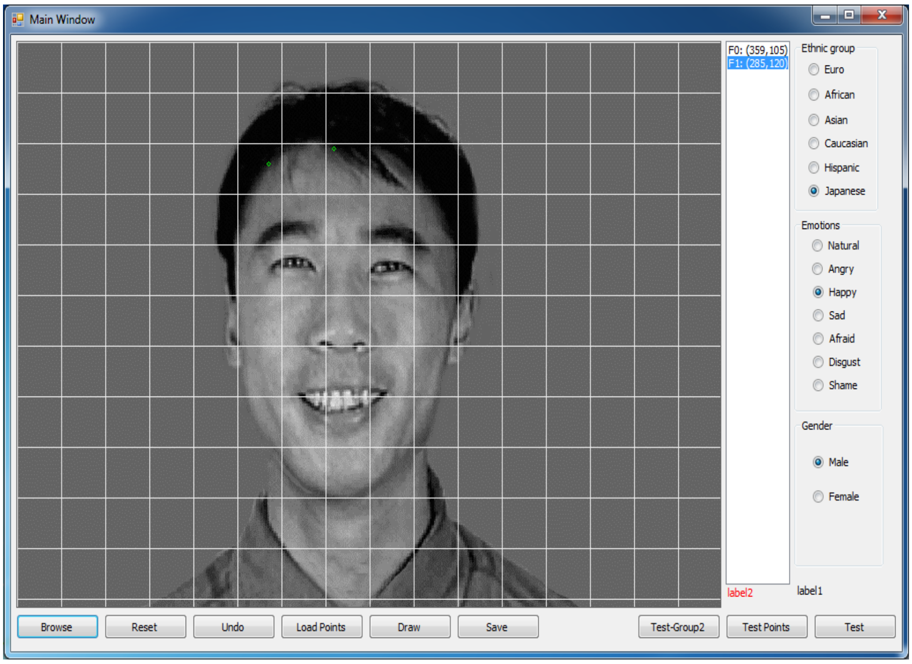


Figure 6. Interface of FFE program.

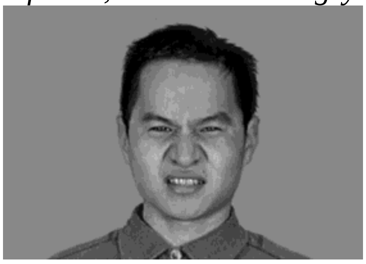
B. Classifier

To build our classifier we use Mathworks Matlab using NNTool tool. This tool allows building our specific ANN classifier, and use backpropagation method, and one hidden layers with 16 neurons.

C. Dataset

Montreal Set of Facial Displays of Emotion (MSFDE) dataset is selected to test and evaluate our approach. This data set includes 224 gray images of females and males faces and four ethnic groups African, Asian, Hispanic and Caucasian, these images describe seven facial expressions nature, happy, sadness, angry, fear, shame and disgust. Each ethnic group contains two genders; each gender contains four persons, each person has seven emotions, but each person has a one image for nature emotion, so we ignored this emotion because we need more than one image for each emotion to be able to train well the ANN. This means we use only six emotions, and Figure 6 shows samples for images of MSFDE dataset.

African, Female and Angry



Asian, Male and Disgust

Caucasian, Male and Happy

Figure 7. Samples of MSFDE dataset.

5. Experiments and results

We performed groups of experiments to study the impact of ethnic group (race) in the accuracy of emotion recognition with three kinds of ethnic groups (Asian, Caucasian as African). So we have three experiments, each experiment has a neural network as a classifier, and each neural network has three layers where there are 16 neurons in the hidden layer except Asian network has 17 neurons (the best result with 17 neurons for Asians).

To study the accuracy of emotion recognition for our approach regardless the ethnic group we used 108 images for the training representing six emotions of six persons. For testing, we used 36 images representing six emotions of six persons.

On the other hand, to study the accuracy of emotion recognition for our approach considering the ethnic group we used 36 images for training for each ethnic group representing six emotions of six persons, and test the classifier by using 12 images representing six emotions of six persons.

Our experiments show that, the impact of ethnic group on the accuracy of emotions recognition is a positive where the accuracy of emotion recognition considering ethnic group is 83.3% as shown in Figure 8, and we got 75% of accuracy regardless ethnic group as shown in Figure 9.

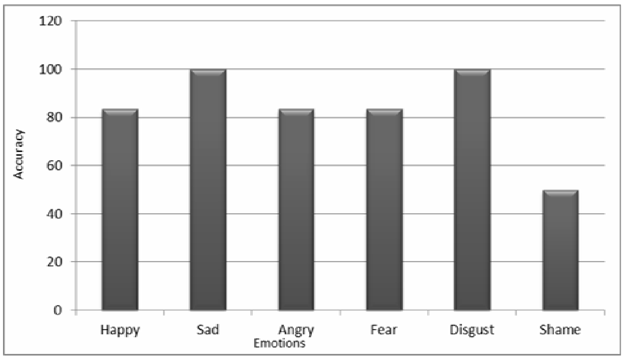


Figure 8. Classification accuracy of emotions considering the ethnic group (Total accuracy is 83.3%)

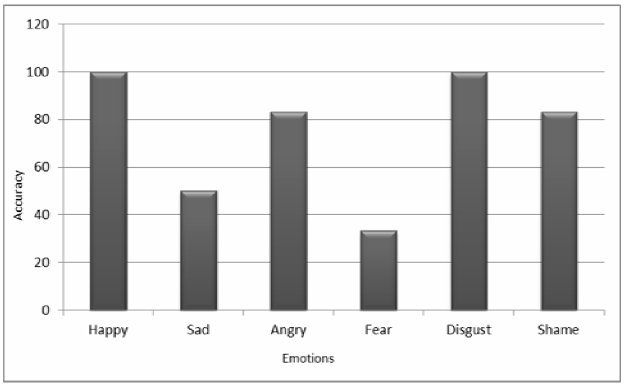


Figure 9. Classification accuracy regardless the ethnic group (Total accuracy 75%)

Also we got that, African race has the lowest accuracy, and the Caucasian race has the highest accuracy of emotions recognition. In additional, the disgust emotion has best accuracy of recognition and fear emotion has the worst accuracy of recognition.

6. Conclusion

Many psychological studies refer to that, the ethnic group is a considered factor in emotion recognition, and may play an important role to increasing the accuracy of emotion recognition. In this paper, we try to investigate the impact of ethnic group on accuracy of emotion recognition based on face expressions by proposing an emotion recognition approach using backpropagation neural networks. Our approach has five stages. To complete first four stages, we developed a desktop program called FFE system, and we used Matlab program to build an ANN classifier. We use MSFDE data set to test and evaluate our approach.

We demonstrated in this paper through our experiments positive results that the ethnic group has a positive impact on the accuracy of identifying human emotions based on facial features. We got 75% accuracy regardless ethnicity, and 83.3% accuracy considering ethnicity. However, this effect is not very large. In additional, the fear emotion has the worst recognition accuracy, and the disgust emotion has the best accuracy. Also, based on the race, the Caucasian has got the best accuracy of emotion recognition, and African has the worst accuracy.

For future work we recommend that; develop a system that can extract the features from human face image automatically, also finding more useful features that may lead to more accuracy on recognition process.

References

[1] Elfenbein Hillary A. and Ambady Nalini, "Universals and Cultural Differences in Recognizing Emotions". American Psychological Society. Volume 12, No 5, pp. 159-164. (2002).

[2] Beaupré Martin G. “Cross-Cultural Emotion Recognition among Canadian Ethnic Groups”, Journal of Cross-Cultural Psychology, Vol. 36, 355-370, (2005).

[3] Castelli Fulvia. "Understanding emotions from standardized facial expressions in autism and normal development". National Autistic Society, Vol 9. Issue: 4, pages 428-449 (2005).
[4] Douglas-Cowie E., Campbell N., Cowie R. and P. Roach, “Emotional Speech: Towards a New Generation of Database” Speech Communication, Vol.40 (1-2). pp 33-60 (2003).
[5] Efford N.,”Digital Image Processing”, Personal Education Limited, (2000).
[6] Ekman Paul;Friesen Wallace V.;O'Sullivan Maureen;Chan Anthony;Diacoyanni-Tarlatzis Irene;Heider Karl;Krause Rainer;LeCompte William Ayhan;Pitcairn Tom;Ricci-Bitti Pio E.;Scherer Klaus;Tomita Masatoshi;Tzavaras Athanase . "Universals and Cultural Differences in the Judgments of Facial Expressions of Emotion". Journal of Personality and Social Psychology. Vol.53, No.4 pp.712-717 (1987).
[7] Ekman, P., Friesen, W. V., & Hager, J. C.”Facial Action Coding System. Salt Lake City”. Research Nexus. (2002).
[8] FG-NET Database with Facial Expressions and Emotions Database (FEED), http://cotesys.mmk.e-technik.tu-muenchen.de/isg/content/feed-database. Resources and Evaluation Conference (LREC). (2006).
[9] Fischer Robert, "Automatic Facial Expression Analysis and Emotional Classification", University of Applied Science Darmstadt (FHD), (2004).
[10] Hewahi N.,Olwan A., Tubeel N., El-Asar S. and Abu-Sultan Z., ”Age Estimation based on Neural Networks using Face features”, Journal of Emerging Trends in Computing and Information Sciences(CIS), Vol.1. No.2. pp. 61-68 (2010).
[11] Horlings Robert. "Emotion Recognition Using Brain Activity". Man-Machine Interaction Group. TU Delft. (2008).
[12] Karthigayan M., Rizon M., Nagarajan R. and Yaacob S., “Genetic Algorithm and Neural Network for Face Emotion Recognition”, University Malaysia Perlis (UNIMAP), (2008).
[13] Kim Moon Hwan, Joo Young Hoon, and Park Jin Bae. "Emotion Detection Algorithm Using Frontal Face Image". Ministry of Information and Communication (MIC). (2005).
[14] Raheja J., Kumar U., “Human Facial Expression Detection from Detected in Captured Image Using Backprobagation Neural Network”, International Journal of Computer Science & Information technology (IJCSIT), Vol.2. No.1, (2010).
[15] Ratliff Matthew S. and Patterson Eric "Emotion Recognition Using Facial Expressions with Active Appearance Models". Third International Conference on Human Computer Interaction. ACTA Press Anaheim, CA, USA (2008).
[16] Rudra T., Kavakli, M. and Tien, D. “Emotion Detection from Male Speech in Computer Games”. TENCON, IEEE Region 10 Conference. 4428779. pp1-4 (2008).
[17] Soto Jose´ A. and Levenson Robert W. "Emotion Recognition across Cultures: The Influence of Ethnicity on Empathic Accuracy and Physiological Linkage". Department of Psychology, Pennsylvania State University, University Park, PA 16802, USA (2009).
[18] Suzuki Kenji, Yamada Hiroshi and Hashimoto Shuji “A similarity-based neural network for facial expression analysis”. Elsevier Science Inc. New York, NY, USA. (2007).
[19] Tsapatsoulis Nicolas, Karpouzis Kostas, Stamou George, Piat Frederic and Kollias Stefanos. “A Fuzzy System for Emotion Classification based on the MPEG-4 Facial Definition Parameter Set”. European Signal Processing Conference (EUSIPCO'00) (2000).
[20] Waller B., Cray .J. and Burrows A., “Selection for Universal Facial Emotion”, American Psychological Association, 1528-3542/08, (2008).