Volume 3, Issue 3
June 2014

ITEE Journal
Information Technology & Electrical Engineering

ISSN: - 2306-708X

Detection and Classification of Facial Expressions using Artificial Neural Network

[1]Samin Iftikhar, [2] Rabia Younas, [3] Noshaba Nasir, [4] Kashif Zafar

Department of Computer Science, National University of Computer and Emerging Sciences Lahore, Pakistan

E-mail: [1]l125011@lhr.nu.edu.pk, [2]l125007@lhr.nu.edu.pk, [3]l135070@lhr.nu.edu.pk, [4]kashif.zafar@nu.edu.pk

ABSTRACT

Facial Expression gives important information about emotion of a person. Facial expressions are the most expressive way human being display emotions. This paper presents a technique based on multilayered artificial neural network for the detection and classification of frontal facial images. The process involves four stages pre-processing, face detection, features extraction and emotion detection. The emotion worked on is astonished, happy and neutral. The total average accuracy of classification is 86.1%.

Keywords: *Artificial neural network, facial expression, database, face detection.*

1. INTRODUCTION

The face of the person is regarded as the mirror of the mind. Facial expressions and the changes in facial expressions provide important information about affective state of the person, his temperament and personality, information related to stress levels, psychopathological diagnostic information, truthfulness, the authors present etc. [1]. For the past years, research is to correlate features of the face with emotional states. Suwa et. al introduced in 1978, facial expression classification system. The main building block of detection and classification of facial expression system is image processing, feature extraction, face detection and classification. Facial expressions are generated by contractions official muscles, which results in temporally deformed facial features such as eye lids, nose, lips and skin texture, eye brows, often revealed by wrinkles and bulges [2]. A number of facial expression detection from static image methods are proposed in literature. Ekman and Friesen [3] developed the Facial Action Coding System (FACS) for describing expressions such as disgust, fear, happy, neutral, sad, anger and surprise. Padgent [4], Hara and Kobayashi [5, 6], Zhang [7] and Zhao [8] used neural network approach for expression classification. In this research worked on to categorize face expressions into three classes with image processing feature extraction technique and multilayered artificial neural network as categorization. This article is framed as follows. Section 2 reviews Related Work, Section 3 describes the implementation details about image processing technique and neural network architecture process, Section 4 depicts the experimental setup, results and discusses the confusion matrix and accuracy levels of categorization.

2.RELATED WORK

In order to make computers perceive human emotions Ammar A. Alzaydi [9] presents a solution for facial expressions recognition using artificial neural networks.

Artificial neural network classifier takes grayscale images as input and classifies the facial expressions into three categories: astonished, smiling and calm. To make the recognition and matching easy, mouth information is used. 400 pixels of the mouth image are the nodes of input layer of classifier while three nodes are required to classify image into three emotional states in the output layer. The classifier is assembled on multilayered perceptron feed forward artificial neural network. A modified back-propagation training algorithm with acceleration, momentum, and a synthetically enlarged input dataset was implemented as the final "SYDE" method. The finally implemented activation functions were the tansig and purelin functions for the hidden layer and the output layer. Experimental results showed the overall recognition rate of 91.7% on 24 images.

Neeraj Shukla and Anuj Kumar [10] proposed a solution for facial expression recognition which uses facial features such as cheek wrinkle, forehead wrinkle, mouth length and mid forehead wrinkle in order to recognize the facial expression. Neural Network comprised of "input layer, 1-hidden layer and the output layer". Input layer contains four neurons where each neuron represents the data extracted from the cheek wrinkle, forehead wrinkle, mouth length and mid forehead wrinkle. "Hidden layer" contain number of neurons to confine the data into the nonlinear form. "Output layer" consists of 6 neurons that represents facial expressions: fear, anger, surprise, sadness, happiness and disgust. Network is trained using back-propagation algorithm. The experimental results showed that back-propagation algorithm with the method of face features extraction can distinguish the face expressions better than other methods. The expression like sadness and disgust are difficult to distinguish than the others.

Sreevatsa et al. [11] proposed a target oriented method using Neural Network for Facial expression detection. Target oriented in context that with the help of single image of face the detection of a facial expression is performed at the apex of the expression, as contrast to the Gesture Oriented approach where face secular information is extracted from a series of images. The preprocessing on image includes face detection and localization, facial feature extraction where only specific features such as mouth and eyes - the main most expressive features on human face -are extracted from image to be used as input for the Recognition network. The extracted features, mouth and eye block from image are resized to 28x20 and 20x32 resp. and then this 2-D representation is converted to 1-D vector of 1200x1 length for feeding as an input to neural

network. A 4 layered - 1 input, 2 hidden, 1 output layer- feed forward neural network is trained to classifying the emotions based on the input from image processing modules, using back propagation algorithm. The number of "hidden layers" and units in every layer are selected by trial and error based method until performance reaches the specific goal. The network classifies the given image into face expression of six emotional classes – "Happy, Sad, Anger, Fear, Disgust and Surprise". The Experiments were performed on frontal view of colored images, obtained from World Wide Web. The Neural Network toolbox for Matlab and image processing are used for simulations. The recognition rate is 83.33 % for Happy, Anger and Surprise and 66.7 % for Sad, Fear and Disgust.

Raheja et al. [12] describe architecture for human face expression detection. The system consists of 5 modules as; Image acquiring, face detection, Image pre-processing, Neural Network and Recognition. The image processing includes enhancing image, edge detection, thinning and tokenization (to convert image to appropriate input for Neural Network). The whole image is converted to neural network input (via tokenizing) after thinning, in contrast to the feature extraction in [11]. The Neural Network is implemented using Multi Layered Feed forward neural network in which back-propagation is used as learning algorithm. The system trained on 3 gestures; Happy, Sad and Thinking, using 100 images as training data. The accuracy for Happy, Sad and Thinking is 94.28%, 85.71%, 83.33% respectively.

Gaurav B. Vasani et al. presented method for face expression detection using Principal Component Analysis (PCA). In this paper [13], the proposed method starts initially class of expression from the database face images from each with 3 training images and 6 testing images. Subsequent to this the same method is repeated by decreasing the number of testing images and increasing the number of training images from each set of expression. A low dimensional face space is created by the train images. On the training image set is applied Principal Component Analysis (PCA) and reducing the Eigen space by taking the principal components. On the face space are also projected the test images. The input test images with these eigenvectors were categorized based on Euclidian distance. The database of 30 different persons with different expressions was used to test the proposed technique. With even the small number of training images the detection rate was greater which confirmed that it is simple, fast and works better in a constrained situation.

Surbhi and Vishal Arora [14], proposes the facial expression detection from human facial image by using neural network. In this method the database of images, the group of all the images includes numerous expressions of different poses like "neutral, anger, happiness, the authors present etc.". Read input image from database and localize face using expression. The database is used to recognize manually giving expression that are angry, happy, fear, sad, disgust, surprise and neutral. Load an image and click on facial expression recognition then insert the recognized number of facial expression and then match it to texture expression with train neural network which starts testing the facial expression. Recognize expression if match the facial expression to texture expression then image expression is successfully recognized. For the training of neural network the algorithm used is Back-propagation algorithm. The experimental results prove that the proposed technique can automatically recognize human facial features better than other methods for classification of facial expressions. The overall recognition rate and accuracy of the proposed algorithm is 100% on Japanese Female Facial Expression Database (JAFFE) which has 23 pictures of the same person.

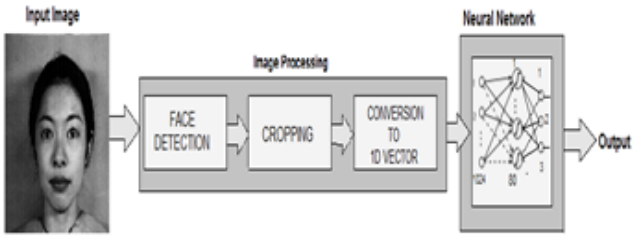
3.IMPLEMENTATION DETAILS

In this technique, Neural Network classifier takes grayscale image as input and gives an expression as output. This process is performed by first preprocessing the input image and then training or testing the image on neural network which in turn classifies the image into one of three categories: Astonished, Normal and Happy.

Figure 1: Flow of data from input to output. Where input is a 2D frontal face image, and output is classified expression of the image from neural network

Image Processing

This research uses the image processing technique in [15]. In this technique are taking 8-bit gray scale images of front view of face, and then performing skin detection based on which the face is detected, after that the face is cropped in such a way to get only the forehead to chine and ear to ear area, this



is done by converting the image to binary image – for the conversion of image to binary average value of RGB is calculated for each pixel, this technique replaces it by "black pixel" or else it replaces it by "white pixel" , if the average value is below 70. In this way, it transforms an input image into binary image. The binary image is then scanned from the middle, to locate continuous "white pixels" after a continuous "black pixel". Then this technique search vertically both left and right side to find the maximum width of the "white pixel". Then, this technique stop the search if the new width becomes smaller that is half of the previous maximum width because it will reach the eyebrow portion then this condition will take place. Then height will be 1.5 * width and from the previous stopping condition it will cut the face from the starting position of the forehead. Figure 2 shows input and final output image sample.

It converts this final image to a 32x32 image in binary, and then converts it to 1024 length 1-D vector to present as an input to neural network, where black pixel is represented as 1 and white as 0 in this final vector

a. input image

b. output image

Figure 2 : Image (a) is the 8-bit gray scale input image to the image processing unit and image (b) is the binary image which is the output we get from image processing unit

Neural Network Architecture

In order to design this technique Neural Network classifier, this research used Encog Neural Network library [16]. Encog supports variety of advanced algorithms and is considered as an advanced machine learning structure. Encog supports many machine learning algorithms like Artificial Neural Networks (ANN), Support Vector Machines (SVM), Hidden Markov Models (HMM), Bayesian Networks (BNN), Genetic Programming, and Genetic Algorithms (GA).

Encog Neural Network Library supports adding multiple hidden layers with any number of neurons. For neural network training it provides numerous training techniques such as Manhattan Update Rule, Back-Propagation, Resilient Propagation, authors present etc. It also supports all the activation functions.

Using Encog Neural Network library, this research designed multilayered feed-forward Neural Network architecture. This technique architecture is consisted of 1024 neurons in input layer i.e. one dimensional vector of 32×32 binary image. Output layer has 3 neurons, one for each emotion category i.e. Astonished, Normal and Happy. Different combinations of hidden layers and neurons per layer were tried to train the network. On the basis of least error, one hidden layer with 80 neurons was chosen for this technique network. The parameters such as learning rate and momentum are also chosen by hit and trial method. This research used for training of network a Back-Propagation algorithm and for activation function a sigmoid function is used. The neurons in every layer and training parameters are shown in Table 1.

Layer	No of Neurons
Input layer	1024
Hidden layer 1	80
Output layer	3

Parameters: Learning rate 0.1 Momentum 0.3

Table 1: Architecture and training parameters

4.EXPERIMENTAL SETUP AND RESULTS

Data Set

This research uses Japanese Female Facial Expression Database (JAFFE) [17] for this technique experiments. The Database contains 256x256, eight bit gray scale, of frontal view of face, sample images shown in Figure 3. This technique only uses image of the following expressions: Happiness, Astonished and Neutral. The total number of subject it chose are 6 and 4 picture of each emotion is taken for each subject making 72 total images, This dataset is divided in half, one set for training and one for testing. The table 2 summarizes the number of pictures in each dataset per expression

Figure 3: Some Sample Images from Database used in this experiment

No if images in each data set per expression	Training Dataset	Testing Dataset
Happy	12	12
Astonished	12	12
Neutral	12	12

Table 2: Number of images in training data set and testing datasets for each expression

Experimental Results

This research conducted this technique experiments for the data set as mentioned in previous section. The experimental results are shown in table 3 in form of confusion matrix which gives the accuracy of the classification. The diagonal elements in the confusion matrix show the classified groups. For the training of the network gave 12 images of every subject. Resultant Confusion matrix (table 3) illustrates that, out of 12 samples of Astonished 11 samples fall in astonished class and 1 in Neutral while 0 is Happy, 1 in Astonished and 1 in Neutral. Similarly classification is done with all groups of subjects as shown in confusion matrix. As a result total average accuracy of classification is 86.1% and misclassification or error rate is 13.9%.

Confusion Matrix			
	Images classified as Astonished	**Images classified as Happy**	**Images classified as Neutral**
Astonished Total Test images : 12	11 91.7%		1 8.3%
Happy Total Test images : 12	2 16.7%	9 75%	1 8.3%
Neutral Total Test images : 12	1 8.3%		11 91.7%

Table 3: Confusion Matrix, diagonal values showed number of correctly classified images and recognition rate, while the non-diagonal values show the error rate

Provided more images in dataset the results sets will improve.

5.CONCLUSIONS

This research presents facial expression detection and classification technique from static gray scale images that contains the frontal view of the human face. The facial expression categorization is based in the given image into three emotional classes – "Happy, Neutral and Astonished". The method consists of three steps, namely face detection, facial feature extraction and emotion recognition. First, feature extraction and face detection is carried out by binary image conversion based segmentation and connected component analysis which is then followed by the exact face cropping approach. Then extraction of facial features and conversion of these features into one-dimensional vector is performed. Finally, emotion detection is performed by giving one-dimensional vector as inputs to a feed-forward neural network trained by back-propagation.

REFERENCES

[1] Ekman P., Huang T. S., Sejnowski T. J., Hager J. C., "Final Report to NSF of the Planning Workshop on Facial Expression Understanding, 1992.

[2] T. Kanade, "Computer Recognition of Human Faces, Birkhauser, Basel, Switzerland and Stuttgart", Germany, 1973.

[3] Ioanna-Ourania and George A. Tsihrintzis, "An improved Neural Network based Face Detection and Facial Expression classification system," IEEE international conference on Systems Man and Cybernetics 2004.

[4] Padgett C., Cottrell G.W., "Representing Face Images for Emotion Classification," Proceedings of Conference on Advances in Neural Information Processing Systems, 1996, 894-900.

[5] Kobayashi H., Hara F., "Facial Interaction between Animated 3D Face Robot and Human Beings," Proceedings of International Conference on Systems, Man, Cybernetics, 1997, 3, 732-737.

[6] Kobayashi, H., Hara, F., "The Recognition of Basic Facial Expressions by Neural Network," Proceedings of International Joint Conference on Neural Network, 1991, 460-466.

[7] Zhang Z., Lyons M., Schuster M., Akamatsu S., "Comparison between Geometry- Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi- Layer Perceptron," Proceedings of International Conference on Automatic Face and Gesture Recognition, 1998, 454-459.

[8] Zhao J., Kearney G., "Classifying Facial Emotions by Backpropagation Neural Networks with Fuzzy Inputs," Proceedings of Conference on Neural Information Processing, 1996, 1, 454-457.

[9] Ammar A. Alzaydi, "Artificial Neural Network Design and Parameter Optimization for Facial Expressions Recognition," International Journal of Scientific & Engineering Research, Volume 2, Issue 12, December-2011.

[10] Neeraj Shukla and Anuj Kumar, "Using Back-Propagation Recognition of Facial Expression," Journal of Environmental Science, Computer Science and Engineering & Technology, 2013.

[11] Sreevatsa n. A. N, Sat hish Kumar . K. G, Rakeshsharma . S and Mohd. Mansoor Roomi. "Emotion Recognition from Facial Expressions: A Target Oriented Approach using Neural Network". ICVGIP, Proceedings of the Fourth Indian Conference on Computer Vision, Graphics & Image Processing, (2004).

[12] Raheja, Jagdish Lal, and Umesh Kumar. "Human Facial Eexpression Detection from Detected in Captured Image using Back Propagation Network." International Journal of Computer Science & Information Technology 2.1 (2010).

[13] Gaurav B. Vasani, Prof. R. S. Senjaliya, Prajesh V. Kathiriya, Alpesh J. Thesiya, Hardik H. Joshi, "Human Emotional State Recognition Using Facial Expression Detection" International Journal Of Engineering And Science, Vol. 2, No. 2, pp 42-44. January, 2013.

[14] Surbhi, Vishal Arora, "The Facial expression detection from Human Facial Image by using neural network", International Journal of Application or Innovation in Engineering & Management, Vol. 2, No. 6, June, 2013.

[15] "Human Emotion Detection from Image" http://www.codeproject.com/Articles/110805/Human-Emotion-Detection-from-Image

[16] "Encog Machine Learning Framework – Heaton Research" http://www.heatonresearch.com/encog

[17] "Japanese Female Facial Expression (JAFFE) Database" http://www.kasrl.org/jaffe.html

AUTHOR PROFILES

1. Samin Iftikhar is a student of Master in Computer Science in National University of Computer and Emerging Sciences. She received her BS degree in Computer Science from National University of Computer and Emerging Sciences in 2006.

2. Rabia Younas completed her BS(Hons.) degree from Preston University Ajman, United Arab Emirates, in 2011.

Currently, she is Graduate Student of Computer Science in National University of Computer and Emerging Sciences Lahore, Pakistan.

3. Noshaba Nasir received the degree in Electrical Engineering from University of Engineering and Technology Lahore, in 2011. Currently, she is doing MS in Computer Science from National University of Computer and Emerging Sciences Lahore, Pakistan

4. Kashif Zafar is associated with National University of Computer and Emerging Sciences since 2004 in the Department of Computer Science. He has done PhD in Computer Science from NUCES, Islamabad, Pakistan in 2010.