AI technique for resolving security in highly sensitive domain in automated face recognition system using Artificial Neural Network

Dr. Himani Mittal Gupta[1], Stuti Aggarwal[2]
Raj Kumar Goel Institute of Technology, Ghaziabad, U.P., India[1]
CDAC NOIDA[2]

Abstract— In this project a novel method based on facial skin aging features and Artificial Neural Network (ANN) is proposed to classify the human face images into three age groups –child©, adult(a) and old(o). Face images are being increasingly used as additional means of authentication in applications of high security zone. The facial skin aging features are extracted by using HAAR transform. And explicit use of PCA and HOG techniques is involved. In our project, effective age group estimation is done using face features like texture and shape of human face. For better performance, it uses two types of features namely, geometrical features and wrinkle features and it includes three main iterations: Pre processing, Feature extraction (using HAAR technique) and Classification(using ANN classifier).As the age of an individual increases, the facial features changes and the database of the security system needs to be updated at regular intervals of time which is a cumbersome task. So, in this project we have addressed the issue of facial aging and have come up with a solution that identifies a person in spite of aging. The geometrical features of facial image includes wrinkle geography, face angle, left to right eye distance, eye to nose distance, eye to chin distance and eye to lip distance. And Viola Jones technique is used for face detection

I. INTRODUCTION

There has been a lot much growth in the research areas of personal monitoring, identification and verification using facial images in many computer vision applications. Some applications in these fields include face recognition, face action classification, poses recognition, age estimation, gender recognition and ethnicity recognition. We have covered age estimation in this report.

Most powerful and accurate classifier in pattern recognition is the human brain. It has got intellectual power since it is a natural classifier involved with training and learning for a specific period of time. Our main attempt in this citation is to convert this biological and behavioural characteristic of human brain into artificial neural network in order to attain the same or better results.

Importance of age classification comes into picture when we see the decorum and rules laid down by our society, government and different bodies of our country. These things are sometimes different for different age groups and genders. Even in our languages sometimes very different vocabularies are used for addressing elders. Whole implementation for age classification is performed in MATLAB software. Advantages and drawbacks of using neural networks and how to overcome these drawbacks in network design is also covered in subsequent chapters. Following flowchart describes the complete algorithm utilized for age classification in this project

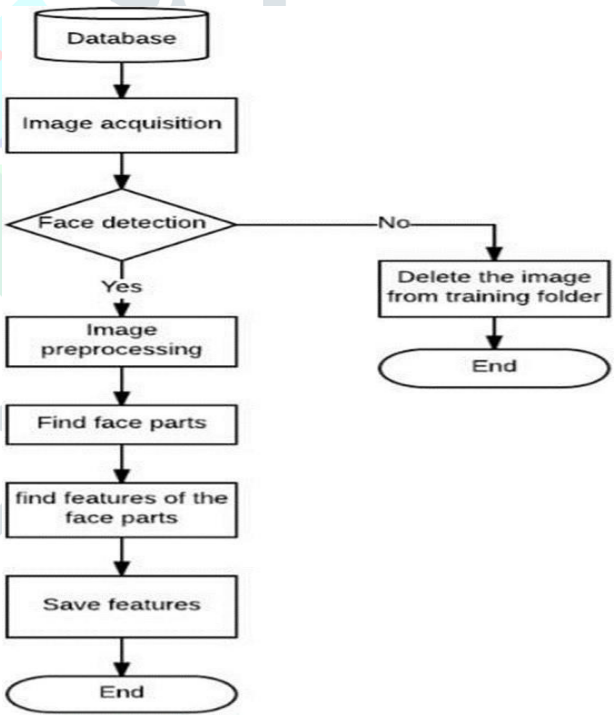

Fig1.1: flow chart training

i) TRAINING

To obtain the desired set of outputs from a set of inputs, a network has to be trained. Each input set is referred to as a vector. Training is accomplished by sequentially applying input vectors, while adjusting network weights according to a predetermined procedure. During training, the network weights gradually converge to values so that each input vector produces the desired output vector. Training algorithms are categorized into two types. They are

- Supervised training
- Unsupervised training

Supervised training: Supervised learning incorporates an external teacher, so that each output unit knows what its desired response to input signals ought to be. Both the inputs and the outputs are provided in supervised training. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually changed. The set of data which enables the training is called the "training set." During the training of a network, the same set of data is processed many times as the connection weights are ever refined.

Unsupervised training: Unsupervised learning uses no external teacher and is based upon only local information. It is also referred to as self organization, in the sense that it self-organizes data presented to the network and detects their emergent collective properties. In unsupervised training, the network is provided with inputs but not with desired outputs. The system itself must then decide what features it will use to group the input data.

ii) NETWORK ARCHITECTURE

The manner in which the neurons of a neural network are structured is intimately linked with the learning with the algorithm used to train the network. There are three fundamental classes of network architectures. They are

- Single layer feed forward networks
- Multi layer feed forward networks
- Recurrent networks

We have used multi layer feed forward networks architecture in our implementation which is explained below.

Multilayer feed forward networks: The second class of a feed forward neural network distinguishes itself by the presence of one or more hidden layers, whose computation nodes are correspondingly called hidden neurons or hidden units. The source nodes in the input layer of the network supply respective elements of the activation pattern (input vector), which constitute the input signals applied to the neurons (computation nodes) in the second layer (i.e. the first hidden layer).The output signals of the second layer are used as inputs to the third layer, and so on for the rest of the network. Typically, the neurons in each layer of the network have as their inputs the output signals of preceding layer only. The set of output signals of the neurons in the output layer of the network constitutes the overall response of the network to the activation pattern supplied by the source nodes in the input layer.

A feed forward network with m source nodes, h1 neurons in the first hidden layer, h2 neurons in the second hidden layer, and q neurons in the output layer is referred to as an m-h1-h2- q network. But, in our MATLAB implementation, we have used a feed forward network with 8 source nodes, 10 neurons in the hidden layer, and 3 neurons in the output layer. So this feed forward network is referred to as an 810-3 network.

Algorithm and technique used

- Viola Jones face detection algorithm
- Haar feature extraction using viola jones face co- ordinates.
- KNN used to classification the selected features
- Artificial neural network used to selected feature

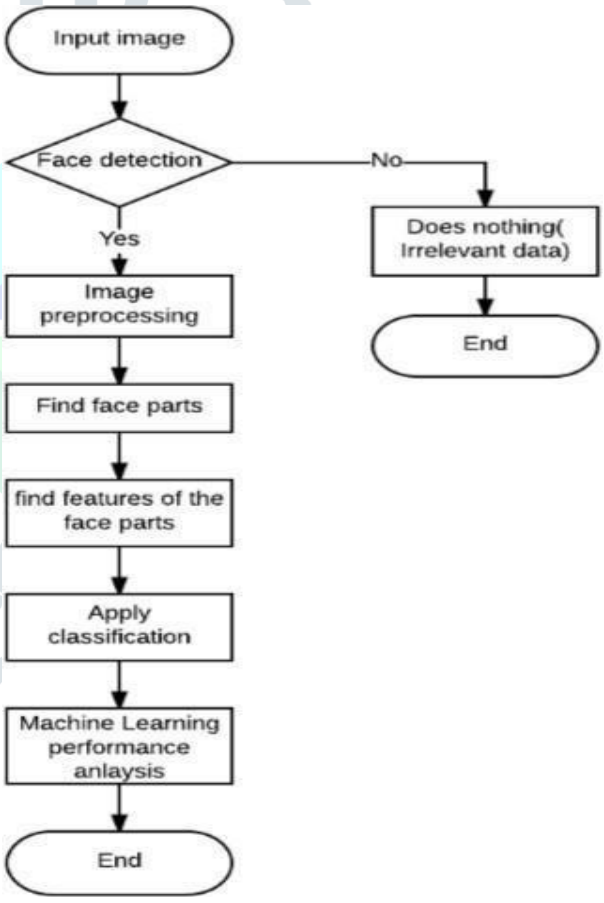

Fig1.2: flow chart testing

II. HAAR LIKE FEATURES AND VIOLA JONES DETECTION

Haar-like features are digital image features used in object recognition. They owe their name to their intuitive similarity with Haar wavelets and were used in the first real-time face detector. Working with only image intensities (i.e., the RGB pixel values at each and every pixel of image) made the task of feature calculation computationally expensive. So working with an alternate feature set based on Haar wavelets instead of the usual image intensities sounds a better option. Viola and Jones adapted the idea of using Haar wavelets and developed the so-called Haar-like features. A Haar-like feature considers adjacent rectangular regions at a specific location in a detection window, sums up the pixel intensities in each region and calculates the difference between these sums. This difference is then used to categorize subsections of an image. The position of these rectangles is defined relative to a detection window that acts like a bounding box to the target object (say, the face). Here comes the concept of summed area table. A summed-area table is a data structure and algorithm for quickly and efficiently generating the sum of values in a rectangular subset of a grid. In the image processing domain, it is also known as an integral image. As the name suggests, the value at any point (x, y) in the summed-area table is the sum of all the pixels above and to the left of (x, y), inclusive: (□, □) = ∑(□′, □′) □□□ □□□ □′ ≤ □, □′ ≤

□ In the detection phase of the Viola–Jones object detection framework, a window of the target size is moved over the input image, and for each subsection of the image the Haar-like feature is calculated. This difference is then compared to a learned threshold that separates non-objects from objects. Because such a Haar-like feature is only a weak learner or classifier (its detection quality is slightly better than random guessing) a large number of Haar-like features are necessary to describe an object with sufficient accuracy. In the Viola–Jones object detection framework, the Haar-like features are therefore organized in something called a classifier cascade to form a strong learner or classifier. The key advantage of a Haar-like feature over most other features is its calculation speed. Due to the use of integral images, a Haar-like feature of any size can be calculated in constant time (approximately 60 microprocessor instructions for a 2- rectangle feature).

III BACK PROPAGATION ALGORITHM

For many years, there was no theoretically sound algorithm for training multilayer artificial neural networks. The invention of the back propagation algorithm has played a large part in the resurgence of interest in artificial neural networks. Back propagation is a systematic method for training multilayer artificial neural networks (Perceptrons). Each input is multiplied by corresponding weights, analogous to a synaptic strength, and all the weighted inputs are then summed to determine the activation level of the neuron. These summed (NET) signals are further processed by an activation function

(F) to produce the neuron's output signal (OUT). In back propagation, the function used for the activation is the logistic function or Sigmoid. The Sigmoid compresses the range of NET so that OUT lies between zero and one. Since the back propagation uses the derivative of the squashing function, it has to be everywhere differentiable. The Sigmoid has this property and the additional advantage of providing a form of automatic gain control (i.e. if the value of NET is large, the gain is small and if it is small the gain is large). The BPNN algorithm involves two phases, during the first phase, the input vector is presented and propagated forward through the network to compute the output values for each output unit. This output is compared with its desired value, resulting in an error signal for each output unit. The second phase involves a backward pass through the network during which the error signal is passed to each unit in the network, and appropriate weight changes are calculated. Learning process in back propagation requires providing pairs of input and target vectors. The output vector 'o' of each input vector is compared with the target vector 't'. In case of difference of these two, the weights are adjusted to minimize the difference. Initially, random weights and thresholds are assigned to the network. These weights are updated every iteration in order to minimize the cost function or the mean square error between the output vector and the target vector

IV RESULT

Feature extraction (using HAAR technique) and Classification(using ANN classifier)as the age of an individual increases, the facial features changes and the database of the security system needs to be updated at regular intervals of time which is a cumbersome task. So, in this project we have addressed the issue of facial aging and have come up with a solution that identifies a person in spite of aging. The geometrical features of facial image includes wrinkle geography face angle, left to right eye distance, eye to nose distance, eye to chin distance and eye to lip distance.

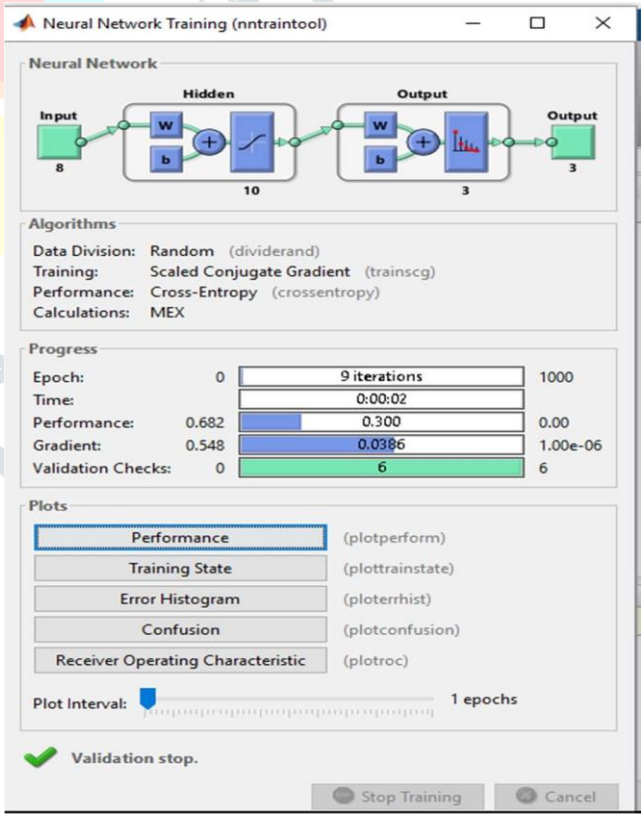

Fig4.1 : NN training tool

Fig4.2: An old man image

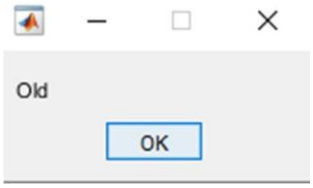

Fig4.3: Result of image

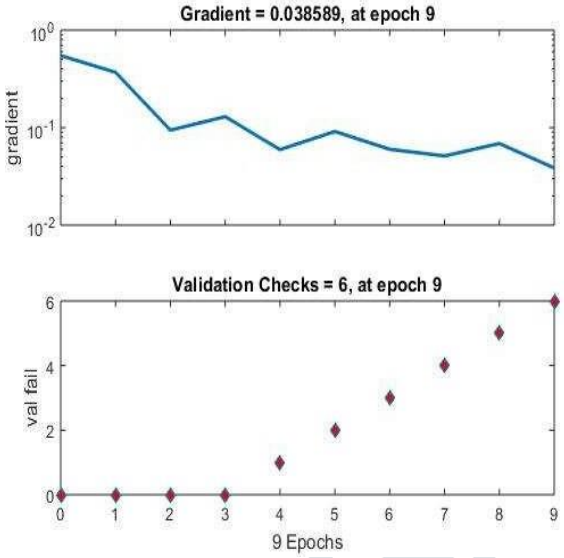

Fig4.4 : Gradient and Validation checks

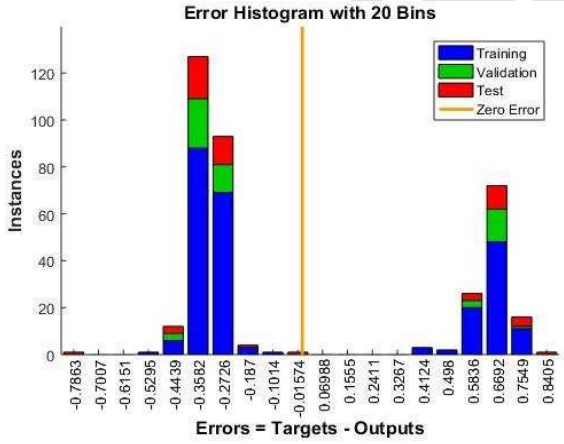

Fig4.5: Error histogram

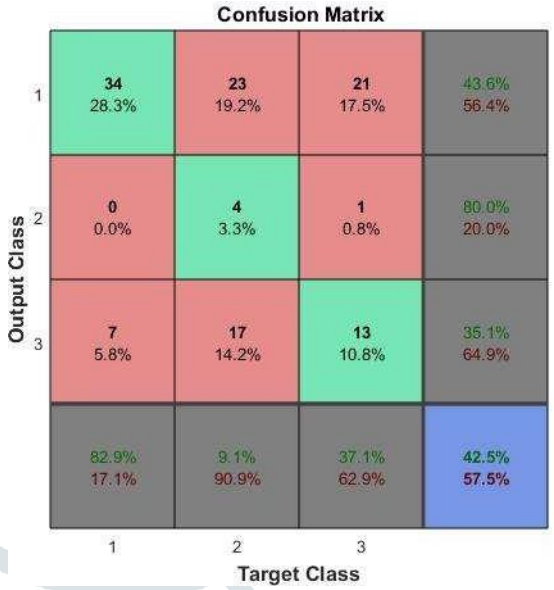

Fig4.6: Confusion Matrix

Confusion matrix :It is a table that is often use to describe the performance of a classification model on a set of test data for which the true values are known. It allows the visualization of the performance of an algorithm. A confusion matrix is a summary of prediction results on a classification problem. The number of correct and incorrect predictions are summarized with count values and broken down by each class. This is the key to the confusion matrix. The confusion matrix shows the ways in which your classification model is confused when it makes predictions .

V. REFERENCES

1. Gil Levi and Tal Hassner, Department of Mathematics and Computer Science, The Open University of Israel .Age and Gender Classification using Convolutional Neural Networks. Published in: Computer Vision and Pattern Recognition Workshops (CVPRW), 2015 IEEE Conference on Date of Conference: 7-12 June 2015 Date Added to IEEE Xplore: 26 October 2015 .

2. Jayant Jagtap, Manesh Kokare .Human age classification using facial skin aging features and artificial neural network. Published in: Cognitive Systems Research. Volume 40, Pages 116-128. Date of Publication: December 2016 .

3. Sneha Thakur, Ligendra Verma, CSE Department ,CSVTU, RIT, Raipur, Chhattisgarh INDIA. Age identification of Facial Images using Neural Network. International Journal of Computer Science and Information Technologies, Vol. 3 (3) , 2012,4244 – 4247

4. P. Viola, M. J. Jones. The Viola Jones Algorithm for face detection. “Robust Real-Time Face Detection”, International Journal of Computer Vision. 57(2), 137–154, 2004