Human Age Estimation from Facial Images using Artificial Neural Network

Swapnil Bhalekar
Department of Computer Engineering,
Sinhgad College of Engineering,
Pune, India.

Ankush Bhandare
Department of Computer Engineering,
Sinhgad College of Engineering,
Pune, India.

Rohit Bhoite
Department of Computer Engineering,
Sinhgad College of Engineering,
Pune, India.

Aditya Dere
Department of Computer Engineering,
Sinhgad College of Engineering,
Pune, India.

Abstract—Due to various potential applications and uses, research related to age estimation using face images has become important. Estimating age from still face images by using facial features is trending research topic from past few years. A lot of approaches and models came forward as the research progressed. Using features like high accuracy, faster processing and ability to learn, we have used Neural Network for this process rather than conventional one. For this purpose, we have used multi-layer perceptron neural network (MLP) which uses backpropagation algorithm. We primarily classify the ages in four groups and further classify into eight groups in secondary stage. We have used FG-NET facial dataset for training and testing the network. Also we have collected our own real face dataset for the validation purpose.

Keywords—Age Estimation, Aging Pattern, Active Appearance Model, Artificial Neural Networks, Backpropagation, Facial Features, Feature extraction, Singular Value Decomposition.

I. INTRODUCTION

Face Images convey a significant amount of knowledge including information about identity, emotional state, ethnic origin, gender, age, and head orientation of a person shown in face image. This type of information plays a significant role during face-to-face communication between humans [1]. Above prospects of facial images can be used in emerging branch of Human Computer Interaction (HCI). Human age has following characteristics:

a. *Aging is uncontrollable process*: Aging cannot be delayed or advanced at will. It is slow and irreversible process.

b. *Personal Age Patterns*: The aging factor of a person is defined by his genetic structure as well as external factors like health, lifestyle, weather conditions, ethnicity, etc.

c. *Aging Pattern is temporal data*: Age and face patterns vary with time. Age pattern at an instance affects all future patterns [2].

Thus, automatic age estimation, being an important technique in real world applications, has become difficult due to these characteristics. Not only these factors but sex of the person also plays a vital role in this process. For this process we need a collection sufficient data of images for training purpose which is partly eased due to public availability of aging database FG-NET which contains the necessary features in co-ordinate form and the age of that individual. The age range covered in this database is 0-69. Fortunately, a complete aging face dataset is unnecessary since human beings also learn to perceive facial ages from incomplete aging patterns [2]. We use multi-layered Artificial Neural Networks (ANN) for classification of the age group of a person. The result is classified into total eight groups of age-ranges.

The motivation for our work lies in various daily life problems. The age estimation system can deal with following problems:

- Human computer interaction (HCI): System features such as text size, volume, display properties can be automatically adjusted depending upon age of the user.
- Multimedia: Content viewed by minors on the internet can be regulated. Also age based image retrieval and video retrieval systems can be benefited.
- Photo indexing: Automatic indexing of photos is possible based on the age of a person.
- Missing individuals: reliable prediction of one's appearance across ages has direct relevance in finding missing individuals.
- Age based access control: Developing systems which provide age specific access to an individual at sites like security offices, military areas, social networking, etc.
- Other common places: Age estimation system can be helpful at various locations like hospitals, police stations, banks, government offices, educational institutes, sport events, etc.

II. Related Work

Various image processing researches related to face have been of keen interest since a long time. From past decade, though the study related with respect to aging pattern and age estimation has become important, it is still very challenging. Mostly age estimation is done using shape patterns of face or using facial texture information such as wrinkles.

Existing methods for facial age estimation typically consists of 2 main steps: image representation and age prediction [3]. The general models used for representing images are Active Shape Model (ASM) [4], Active Appearance Model (AAM) [5], Craniofacial Growth Model [6], Aging Pattern Subspace [2], Manifold Learning [7] whereas for age estimation multiclass classification problem or regressing problem.

The ASM model was proposed by Cootes et al [8]. This was used for feature extraction by characterizing changes in process of facial appearance by face contours. The AAM was developed by Cootes, Edward and Taylor [5].Kwon, Lobo [9], first, proposed the concept of age classification by using face images. The approach was using the geometric ratio of local face features combined with wrinkle analysis. Using these, classification was done into 3 groups as babies, adults and seniors. Identifying the shortcomings of this method, Horng et al [10] used Sobel edge magnitudes for extracting facial features.

Lanitis et al [11] constructed an aging function that used parametric model for human faces. He performed processes automatic age progression, age estimation, face recognition across age, etc. and compared it with neural network approach. The feature extraction was done using Principle Component Analysis (PCA) and Active Appearance Model(AAM). Guo et al [4] proposed a hybrid approach for estimating age using Active Shape Model (ASM) and Radon, DCT Transformation. Support Vector Regression (SVR) was used for learning based classification and regression.

A fine tuning age estimation scheme was proposed by Duong et al [12]. They combined global features and local features. Global features were extracted using AAM whereas local features were extracted using Local Binary Pattern (LBP). Age estimation is done in two steps: fine tuning method that requires an initial age guess followed by a refining step (fine prediction). Geng et al [2] proposed a subspace approach AGES (AGing pattErn Subspace) for automatic age estimation. Appearance Model was used as feature extractor that combined shape and intensity of face images. Further, Linear Discriminant Analysis (LDA) is applied to these features which try to find a subspace such that images of same age converge while images of different age scatter. The output is classified into seven groups of age groups of range ten years each.

III. DATA COLLECTION

We use FG-NET database to train and test our neural network. FG-NET dataset is preprocessed and the features are already extracted in form of (x,y) co-ordinate points. It contains 1002 grey and colored face images. About 250 images were chosen from this dataset for training. Clear images which comprised range of age values from different age-groups were selected. The age range of images in this dataset is 0-69 years. Images used for testing are combination of images from FG-NET dataset and images captured on our own using camera. Images from FG-NET dataset which were used for training are not allowed for testing purpose. About 75-80 images are used for validation.

IV. PROPOSED APPROACH

The steps followed in the process of estimating age are shown in Figure. 1.

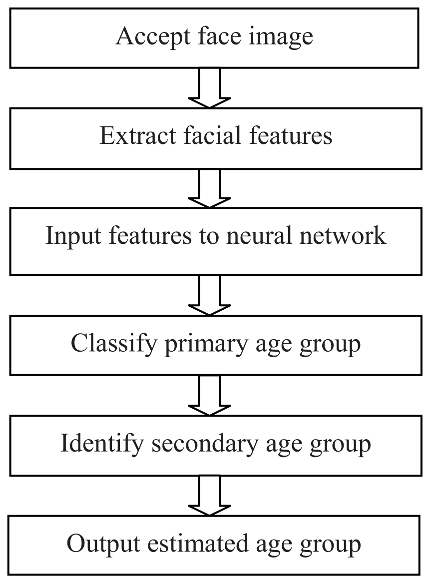

Figure. 1. Steps in age estimation process

We input face image to our system. The validation of this image needs to be done manually. Validation of image is done such that frontal view of face is visible so as to extract the features accurately. After validating image, we extract face features from it. Features are 68 pair of (x,y) co-ordinates. For extracting these features, we have used custom built am-markup tool. Features obtained are identical to the ones provided in FG-NET dataset. These features are stored into a text file in the form of (x,y) co-ordinates. The mapping of 68 points on face is done as given in Table 1.

Table 1.Facial Landmarks

Face region	Point id	No. of Points
Face edges	0-14	15
Right eyebrow	15-20	6
Left eyebrow	21-26	6
Left eye	27-30	4
Left iris	31	1
Right eye	32-35	4
Right iris	36	1
Nose	37-45,67	10
Nostrils	46,47	2
Lip outline	48,54,60-66	9
Top lip	49-53	5
Bottom lip	55-59	5

Above listed points are plotted on accepted face image as shown in Figure. 2.

Figure. 2. Landmarked face points

After the features are extracted, we input these features to the neural network. We use Multilayer Perceptron (MLP) neural network for our system. Multiple hidden layers are used in the applications where accuracy is the criteria and no limit for the training time is mentioned. Even the drawback of using multiple hidden layers in the neural network is that they are more prone to fall in bad local minima [13]. We have total 4 layers in the neural network: input layer, output layer and 2 hidden layers. Sigmoid function is used as activation function for activating the neurons.

Backpropagation algorithm is used for training the neural network. It is a type of supervised learning in which we know the expected result beforehand. When the expected result is not obtained, we propagate the error to previous layer until the input layer is reached. This process is continued iteratively until we get minimal error that can be tolerated. We calculate the mean squared error (MSE) over entire set of training pattern using formula,

$$E = \sum_p E_p = \frac{1}{2}\sum_p \sum_n (t_{j_n} - a_{j_n})^2$$

where, E is total training error and p represents all training patterns, E_p is training error over a single pattern, ½ is a value applied to simplify function’s derivative, n represents all output nodes, t_j sub n represents the target value for node n in output layer j, and a_j sub n represents the actual activation for the same node [14].

Input layer consists of 136 neurons which are (x,y) co-ordinates. Each (x,y) co-ordinate pair is correlated into a single unit using Singular Value Decomposition (SVD). Therefore, we have total 68 neurons at first hidden layer.

The singular value decomposition (SVD) of a matrix A is the factorization of matrix A into the product of three matrices,

$$A = UDV^T$$

where, the columns of U and V are orthonormal and the matrix D is diagonal with positive real entries. It have been used widely for dimensionality reduction where it can effectively reduce large dataset into smaller one which still contains a large fraction of the variability present in the original data [15].Singular Value Decomposition is applied successfully in many applications such as data analysis, signal processing, pattern recognition, image compression, weather prediction, and Latent Semantic Analysis [3].

Second hidden layer consists of 36 neurons. The number of neurons in hidden layer is calculated using various factors such as number of neurons in input layer and output layer, complexity of activation function, neural network architecture, training algorithm and training samples database on which the neural network is executed [16].

First stage classification consists of four age groups. Thus we have four output neurons. In primary stage classification the age groups are divided into 0-12, 13-25, 26-45, 46-69 years.

The neural network model for primary stage classification is shown in Figure. 3.

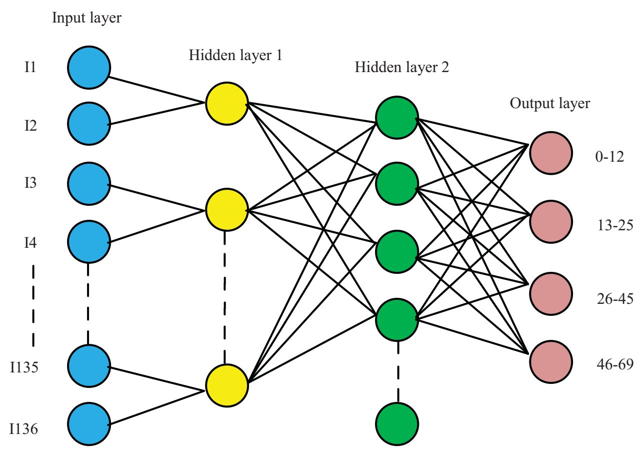

Figure.3.Neural Network model for 1st stage classification

After primary classification of age group, we classify the secondary age group of the accepted input image. The secondary neural network also consists of 136 input neurons and two hidden layers are used with 68 neurons in hidden layer one and 36 in hidden layer two. The output layer consists of 2 layers which are subcategorized from primary stage output layers. The secondary age group ranges are 0-8, 10-12, 13-19, 20-25, 26-35, 36-45, 46-53and 54-69. A part of this neural network is given in Figure. 4.

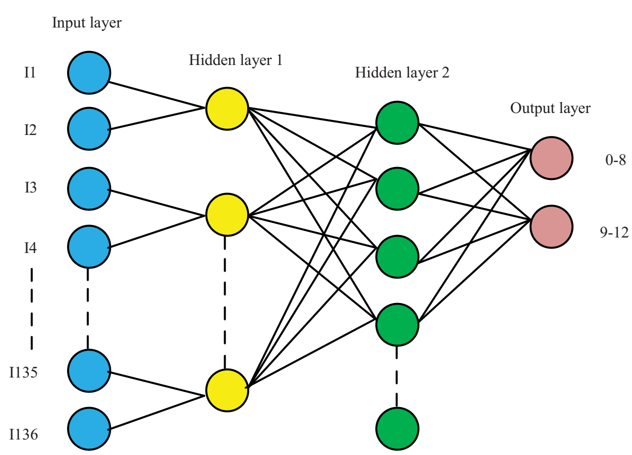

Figure.4. Neural Network model for 2nd stage classification

It is quite difficult to get the exact age as an output and hence the result is obtained in form of age group (bin) rather than a value.

About 250 clear and frontal images are selected from FG-NET dataset for training the neural network . The distribution of images across the age groups is shown in Figure. 5.

Figure. 5. No. of trained images for each group

V. Experimental Results

We carry out different experiments on our system to validate efficiency and accuracy of feature extraction and age estimation using FG-NET database and self-collected images.

The details of the neural network parameters set for evaluation of the system are given in Table 2.

Table 2. Details of neural network parameters

Neurons in input layer	**136**
Hidden layers	**2**
Neurons in hidden layer 1	**68**
Neurons in hidden layer 2	**36**
Learning rate	**0.01**
Momentum	**0.1**
Activation function	**Sigmoid**
Epochs	**2000**
Training Error	**0.0343**

.

Table 3 gives us the number of images that are tested for each age group and the accuracy obtained by correctly estimating the age group in which the image belonged. Images used for training are not considered during the testing phase of the system.

Table 3. Classification results for different age groups

Age group	Samples tested (no. of images)	Accuracy
0-12	22	81.67%
13-25	26	86.54%
26-45	20	74.36%
46-69	12	65.92%

We get best accuracy when the number of age groups is less. As we increase number of groups the accuracy decreases. As a result, we get less accurate results for more specific age groups.

VI. CONCLUSION

This work demonstrates Artificial Neural Networks (ANN) can significantly increase the accuracy for estimating age. Extraction of the face features from accepted real time images is done using Active Appearance Model (AAM). To implement our neural network, we have used Multi-Layer Perceptron (MLP) which yielded better results than other neural networks. Singular Value Decomposition (SVD) is used for mapping of the features into a meaningful unit.

We classified ages primarily into four groups which are later sub-classified into eight groups. FG-NET dataset is used for training as well as testing the neural network. Also, we used real time dataset for testing the neural network. The results obtained are quite satisfactory even for the real time dataset.

VII. ACKNOWLEDGEMENT

We are grateful to Prof. Timothy Cootes for providing the pre-processed FG-NET dataset of facial images. Also, we are thankful to Prof. Pravin M. Kamde for his valuable contribution in developing the system and guiding us throughout the entire process.

REFERENCES

[1] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation", IEEE Transactions on Systems, Man, and Cybernetics, Part B, Vol. 34, No. 1, 2004, pp. 621-628.

[2] X. Geng, Z. Zhou, and K. Smith-Miles, "Automatic age estimation based on facial aging patterns", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, No. 12, pp. 2234–2240, 2007.

[3] Sarah N. Kohail, "Using Artificial Neural Networks for Human Age Estimation Based of Facial Images", 2012 International Conference on Innovations in Information Technology (IIT), pp. 215 -219, 2012.

[4] J. Guo, Y. Liou and H. Nguyen, "Human face age estimation with adaptive hybrid features", System Science and Engineering (ICSSE), 2011 International Conference, pp.55-58 , 2011.

[5] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models", Proceedings of ECCV, Vol. 2, pp. 484–498, 1998.

[6] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," Proceedings of IEEE Conf. CVPR, pp. 387–394, 2006.

[7] G. Guo, Y. Fu, C. Dyer, and T. Huang. "Image-based human age estimation by manifold learning and locally adjusted robust regression", IEEE Trans. Image Proc., Vol. 17, No.7, pp.1178–1188, 2008.

[8] Cootes , F., C. J. Taylor, D. H. Copper, and J. Graham, "Active Shape Models – Their training and Application", Computer Vision Graphics And Image Understanding, Vol. 61, No 1, pp. 38–59, 1995.

[9] Y. H. Kwon, and N. V. Lobo, "Age classification from facial images", Computer Vision and Image Understanding, Vol. 74, No. 1, 1999, pp.1-21.

[10] W. B. Horng, C. P. Lee and C. W. Chen, "Classification of Age Groups Based on Facial Features", Tamkang Journal of Science and Engineering, vol. 4, no. 3, pp. 183-192, 2001.

[11] Lanitis, C. Draganova, and C. Christodoulou. Comparing different classifiers for automatic age estimation. IEEE Transactions on Systems, Man and Cybernetics - Part B, 34(1):621–628, February 2004.

[12] N. Duong, K. G. Quach, K. Luu, H. B. Le and K. Ricanek, "Fine Tuning Age Estimation with Global and Local Facial Features", The 36th Intl. Conf. on Acoustics, Speech and Signal Processing, (ICASSP 2011), Prague, Czech Republic, May, 2011.

[13] Yinyin Liu, Janusz A. Starzyk, Zhen Zhu, "Optimizing Number Of Hidden Neurons in Neural Networks", http://www.ohio.edu/people/starzykj/network/Research/Papers/Recent%20conferences../Hidden%20Neurons%20AIA2007_549-204.pdf.

[14] http://www.webpages.ttu.edu/dleverin/neural_network/neural_networks.html.

[15] S. Ziabakhsh, M. Payravi, H. Alavi-Rad and S. Ziabakhsh, "Color Image Compression Using Singular Value Decomposition", Proceedings of the first regional conference on new approaches in computer engineering and information technology, 2011.

[16] Saurabh Karsoliya, Approximating Number of Hidden layer neurons in Multiple Hidden Layer BPNN Architecture, International Journal of Engineering Trends and Technology- Volume3 Issue6, pp. 714-717, 2012.