National Conference on Synergetic Trends in engineering and Technology (STET-2014)
International Journal of Engineering and Technical Research ISSN: 2321-0869, Special Issue

Age Classification via Facial Images Based On SVM

Aparna Asthana, Prof. S.K. Singh, Hitesh Madan

***Abstract*— A person's face provide a lot of information such as age ,gender and identity.Faces allows humans to estimate/classify the age of other persons just by looking at their face. Researchers who carried out work in studying the process of age classification by humans conclude that humans are not so accurate in age classification hence the possibility of developing facial age classification methods poses an attractive direction.In this research, we try to prove that computer can classify human age according to features extracted from human facial images using Support Vector Machine (SVM)..Many attempts towards age classification are tried and most of them give results for wide ranges of ages or classify the ages in groups such as child ,adult and old.We focus our research on more accurate age group.**

***Index Terms*— Support Vector Machine (SVM), AGE CLASSIFICATION,AGING**

I. Introduction

Human brain and mind has an extraordinary ability to identify different faces from knowledge of appearance. Human brain is not so precise when it comes to the classification of age. Researchers who perform work in studying the process of age classification by humans comes to a decision that humans are not so precise in age classification hence the possibility of developing facial age classification/estimation methods presents an arousing interest and exciting direction.
To achieve our goal, we have to find good databases that we can use to test and train our proposed method also we have to construct a proper SVM to model our problem. Our main goal is age classification so we care only about images of frontal faces.

II. 2.RELATED WORK

Finding an appropriate method for age classification for getting more specific categories of age ranges is still a challenging problem. Thus we focus our research on predicting a more accurate age group. In our paper we have decided to make a good database and construct a SVM model. Since the focus of the current research paper is on age Classification and not on other allied fields like face recognition, so we have narrowed down our image database to only the images of frontal faces. We have taken care to include only such images in our dataset which are as clear as possible without and images without any external additions viz. glasses, beards, scar marks, cosmetics etc. Also the faces should conform to general average faces of the populations. Age classification shares numerous problems encountered in other typical face image interpretation tasks [1] such as:
- face detection,
- face recognition,
- facial expression, and
- gender recognition

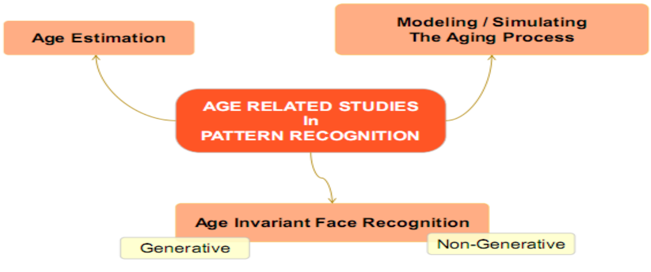

Fig 1- Age estimation model

Once our dataset is ready we will carry out training of our proposed method. After reasonable amount of training is completed we will proceed to the testing of the application by providing to it a set of new faces of known age, which are not part of the training dataset, and then to see with what accuracy the application is able to predict the age group.

A. Pattern Recognition

A definition of pattern recognition then is field whose objective is to assign an object or event to one of the number of groups, based on features derived to emphasize commonalities. The term pattern recognition can be misguiding. The patterns connected with pattern recognition are not single instances of patterns in a signal - not an area of stripes in an image or an interval of sinusoids in a sound clip Now let's focus on the word recognition. in a broad sense, recognition implies the act of associating a classification with a label. Using the Figure 1 below, The traditional aim of the feature extractor is to characterize an object to be recognized by measurements whose value is very similar for objects in the same group, and very different for objects in different groups. The work of the classifier stage of a full system is to use the feature vector provided by feature extractor to assign the object to the group. The degree of problem of the classification depends on the variability in the characteristic values for the objects in the same group relative to the difference between characteristic values for objects in different groups.

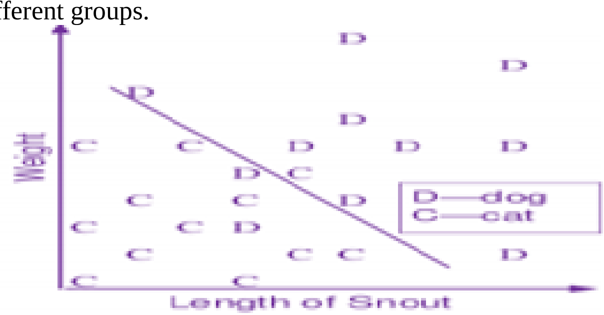

Fig 1- Word Recognition

B. Image

A digital image is composed of pixels which can be thought of as small dots on the screen. A digital image is an instruction of how to color each pixel. A typical size of an image is 512-by-512 pixels. In the general case we say that an image is of size m-by-n if it is composed of m pixels in the vertical direction and n pixels in the horizontal direction [***Gray scale***

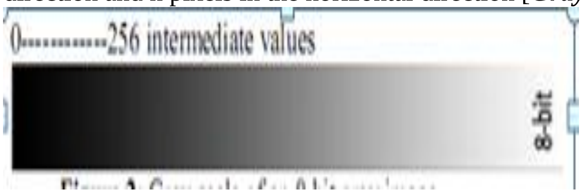

Figure 2: Grey scale of an 8 bit gray image.

Histogram Equalization of Gray Scale Images

In our current research we have used histogram equalization to enhance the image. Histogram equalization spreads the intensity values of the images over a larger range. It is used because it decreases the effect of different imaging conditions, for example different camera gains and it may also increase image contrast [6]. The function that maps image pixel intensity to a histogram equation

$$S_k = \sum_{j=0}^{k} \frac{n_j}{k}$$

k = 0, 1, 2… L−1.

Where –

Sk is histogram equalized intensity value for kth intensity value in the range L of total number of possible intensity values in the original and target image. n is the number of pixels in the original and target image, and njis the number of image pixels that have intensity value i in the original image.

C. Classification

Classification goals to build an efficient and effective model for predictiing the class labels for new samples/observations. The model is built on the training set of samples/observations and their class labels.

D. Statistical Classification

In machine learning, statistical classification is the problem of identifying the sub-group to which new samples belong, where identity of sub-group is unknown. On the basis of a training set of data containing samples whose sub-group is known. Therefore these classifications will point out a variable behavior which can be studied by statistics[7][8].

E. Binary and Multi-class Classification

Age classification can be thought of as composed of two separate problems

Binary classification

Multi-class classification

In the binary classification problem as name indicate only two classes are involved, whereas in multi-class classification involves assigning an object to one of several classes. Since many classification methods have been developed for the binary classification, multi-class classification often requires the combined use of multiple binary classifiers.

F. Overfitting

The concept of over fitting is most important in machine learning algorithms. Over fitting means fitting the training data too much which may allow perfect classification of only the training data that may increase the performance of classification, but it is unlikely that it will perform well on new patterns. It degrades the generalization of performance

G. Review

Age estimation techniques fall within two main approaches.

Estimation on the basis of a set of facial features According to this approach, the problem is treated as a standard classification problem. And we solve using standard classifiers where age classification is performed by assigning a set of facial features to an age group. Estimation on the basis of aging process

According to this approach, as an option, age estimation techniques that depend on the modeling of the aging process have been developed.

III. MATERIALS AND METHODS

A. Database Selection

Among many available face databases around the world, two of them include significant sets for aging individuals but not fulfill all the requirements used for this experiment

In age classification we use the following steps to perform the classification – convert the color and gray scale images to gray scale images, preprocessing, face detection, extract the features to train the classifiers and at last we perform the classification on test faces.

B. General Approach for Age Classification using SVM

First we perform the pre-processing operations on the input gray scale image. These operations are histogram equalization, intensity normalization. Histogram equalization and intensity normalization are operated on the image due to varying lighting A training dataset is created by extracting the features from the faces. Then a classifier is trained with feature and the labeled data set pair and formed the model. For a test image, the features are extracted in the same way. The model uses these features to classify the age group of the person.

Testing Set 24 24 24 72

Convert Color image to Gray scale image

.Gray value for pixel i in an image is linear combination of three intensity values of three primary colors (Red, green, blue i.e. RGB) corresponding to pixel i.

Gray value (i) = 0.2989 * R(i) + 0.5870 * G(i) + 0.1140 * B(i) [Source: MATLAB help] eq. (3.1)

Where

Gray value (i) =gray level value for pixel i.

R(i) = Intensity of red color in pixel i .

G(i) = Intensity of green color in pixel i .

C. Face Detection

As is described before face detection is done with the help of OpenCV library [17]. Face detection techniques are used to determine the location of human faces in a gray scale images while ignoring other objects in the image. There are wide variety of implementations and we use the OpenCV implementation that uses haar features. we used face detection algorithm for the facial feature extraction. The description of the algorithm is as follow- Viola-Jones proposed the use of Haar-like features which can be computed efficiently with integral image. Figure 5 shows four types of Haar-like features that are used to encode the horizontal, vertical and diagonal intensity information of face images at different position and scale.

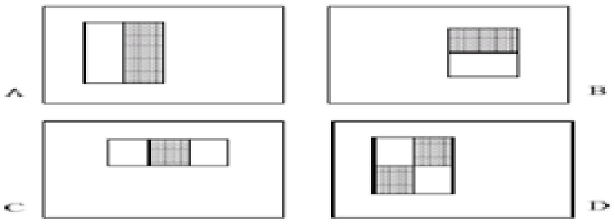

Figure 4- Sum of the pixels within rectangle

The Haar-like features are computed as the difference of dark and light regions. They can be considered as features that collect local edge information at different orientation and scale. The set of Haar-like features is large, and only a small amount of them are learned from positive and negative examples for face detection.

Example: For A, The value of a two-rectangle feature is the difference between the sums of the pixels within two rectangular regions.

D. Integral Image

Preprocess: normalize each image by dividing each pixel value by the standard deviation of the image. **The value of the integral image at point (x; y) is the sum of all the pixels above and to the left.** Where, is the integral image and is the original image. Using the following pair of recurrences:

$$ii(x,y) = \sum_{x' \le x, y' \le y} i(x',y')$$

Where ii(x,y) is the integral image and i(x,y) is the original image.
Using the following pair of recurrences-:

$$s(x,y) = s(x,y-1) + i(x,y)$$
$$ii(x,y) = ii(x-1,y) + s(x,y)$$

The integral image can be computed in one pass over the original image.

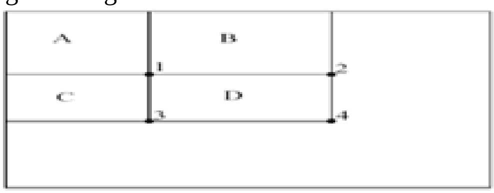

Figure 5: Sum of the pixels within rectangle

In the Figure 6 above, the sum of the pixels within rectangle D can be computed with four array references. The value of the integral image at location 1 is the sum of the pixels in rectangle A. The value at location 2 is A + B, at location 3 is A + C, and at
location 4 is A + B + C + D. The sum within D can
be computed as 4 + 1 - (2 + 3).
Using the integral image, the rectangular features can be calculated more efficiently.

E. Adaboosting

Boosting is a general method for improving the accuracy of any given learning algorithm. One can use it to combine simple "rules" (or weak learner), each performing only slightly better than random guess, to form an arbitrarily good hypothesis. and Jones employed Adaboost (an adaptive boosting method) for object detection and got good performance when applying to human face detection Classification using Support Vector Machine

F. SVM (Support Vector Machine)

are a useful The aim if SVM is to train the classifier using training data, and to produce a model which predicts the class labels for the test data when the input is only test data.
For the training set is linearly non-separable then it is mapped to high-dimensional feature space. This projection into high dimensional feature space is efficiently performed by using kernels.
The kernel K is defined as –
For this project, a polynomial kernel is used which is defined as –

$$K(x_i, x_j) = \varnothing(x_i)^T \varnothing(x_j)$$

For this project a polynomial kernel is used
Which is defined as-

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0.$$

Here , and are kernel parameters which controls the performance of the classifier. The following steps are taken for the classification in MatLab.
1.Given data are transformed in the format of SVM package (for example in the MatLab all data are taken as a separate matrix and corresponding group is taken as separate matrix).
A MatLab function libsvmread reads the file in LIBSVM format: [Class_labels, Data_Matrix] = libsvmread('data.txt').
Output of the libsvmread is Label
1:feature1
2:feature2
3:feature3 etc.
1.Perform scaling on the data
2.Polynomial function taken as a kernel.
3.10-cross validation and find the best parameter C and γ.
4.Whole training set is trained by best parameter of C
and γ.

IV. Test Scaling

Before using SVM, scaling of data is very important. There are advantages when we scale the data. First, the main advantage of scaling is to avoid attributes in greater numeric ranges dominating those in smaller numeric ranges. Second, another advantage is to avoid numerical difficulties during the calculation.
Kernel values depend on inner product of features vector, so polynomial kernel and linear kernel causes numerical problem if we do not perform scaling on data. The range of scaling each attribute is [−1, +1] or [0, 1].

A. Cross Validation

Cross validation is a statistical method of evaluating and comparing learning algorithms by dividing data into two partitions:

First is used to learn or train a model, and another is used to validate the model. The basic form of cross validation is v-fold cross validation. In v-fold cross validation the data is first partitioned into v equal sized folds. Then v iterations of training and validation is performed such that within each iteration a different fold of data is held-out for validation while remaining v-1 folds are used for learning. The advantage of this method over repeated random sub-sampling is that all observations are used for both training and validation, and each observation is used for validation exactly once. Using the cross validation procedure, over-fitting problem can be prevented.

B. Steps in Age Classification

Following steps are implemented for age classification –
Step 1: Convert color or gray scale images into gray scale images.
Step 2: And then pre-process the images using
Histogram Equalization and intensity normalization techniques.
Step 3: Extract the frontal face of the image using Viola-Jones Method with the help of OpenCV library.
Step 4: Pre-processing and feature extraction. Step 5: Create a Model using SVM classifier. Step 6: Predict the age group with the help Trained Model.

V. RESULTS AND DISCUSSION

The training set consists of 360 faces, 120 each of age group child, adult and old person faces. The test data have 72 faces, 24 each of age group child, adult and old person faces. fold cross validation of SVM classifier using polynomial kernel is 71.7361%. When the training set is directly given to the libSVM [16] routine, depending on the training set and testing set, the accuracy of the SVM classifier is varied. The average accuracy for the age group of child, adult and old using combination of different training and testing set are approximately 83.3333%, 68.75% and 61.4583% respectively. The result obtained from the first trial is shown in table 1 and results obtained from the other trials are shown in appendix H. The experimental result of first trial is shown in the Table 2.

Input Age Group	Child	Adult	Old
Child	18	3	75%
Adult	4	16	66.66667%
Old	2	5	70.83333%

The results obtained from the first trial are given below.
Training set: Total = 360 Child = 120 Adult = 120 Old = 120
A- Test: total = 24; 18 from class 0.
Output Predicted Label:
0 1 0 0 0 0 0 2 0 0 2 0 0 1 0 0 0 0 2 0 0 1
0 0
Accuracy: 18/24 = 75 %
B- Test: total = 24; 16 from class 1.
Output Predicted Label:
2 0 0 1 1 1 1 1 2 1 2 1 0 2 1 1 1 1 1 1 1
1 1 0
Accuracy: 16/24 = 66.6667 %
C- Test: total = 24; all 17 from class 2
Output class label:
2 1 1 2 2 2 2 1 2 1 2 2 2 2 2 2 2 2 0 2 0 2 2
1 2 Accuracy: 17/24 = 70.8333 %

VI. CONCLUSION

We proposed a method for age classification using facial features based on Support Vector Machine [SVM]. We classified the age into three age groups such as child, adult and old. The development process includes data collection, feature extraction and finally training and testing by the system Support Vector Machine. To train and test our system, we used the dataset taken from the websites. There are 432 gray-scale facial images of each size 100*100 used for experiment. 360 images are used to train the SVM classifier. And the rest of images are used to evaluate the performance of the system. For the test images, the correct rate for distinguishing child is 83.3333%, adult 68.75% and old 61.4583%. After systematic examination the factors that affect the performance of the classifier, some important are summarized as follows.
1. Facial wrinkles are usually removed by the photographer or due to some cosmetics i.e. men and women use cosmetic creams on their faces that cause problems to extract the exact features.
2. For some images light source are too strong so that some important features are lost.
3. Due to glasses and beards are also some important features are lost. The performance of the classifier is dependent strongly on the nature of the training and testing data sets.

REFERENCES

[1] .http://www.scholarpedia.org/article/Facial_Age_Estimationinternational association for pattern recognition Volume 25 Number 1 Winter 2003. http://amath.colorado.edu/computing/Matlab/Tutorial/ImageProcess.html
[2] Gonzalez R.C and Woods R.E (2008), Digital Image Processing Third edition, Pearson Education. Gonzalez R.C , Woods R.E and Eddins L. (2008), " Digital Image Processing Using MATLAB", Pearson Education.
[3] Erno Mäkinen. Introduction to Computer Vision from Automatic Face Analysis Viewpoint, Department of Computer Sciences University of Tampere, Finland.
[4] Duda R. and Hart P. Pattern Classification and Scene Analysis, Wiley, New York 1973.
[5] Nello Cristianini and John Shawe-Taylor, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, 2000.
[6] Tan PN, et al. Introduction to Data Mining (First Edn) 2005, Addison Wesley.
[7] Y.H. Kwon and N. da Vitoria Lobo. Age Classification from Facial Images, Computer Vision and Image Understanding Journal 74(1), pp. 1-21, 1999.
[8] Lanitis, C. J. Taylor and T.F. Cootes. Toward Automatic Simulation of Aging Effects on Face Images, IEEE Transactions of pattern analysis and machine intelligence, 24(4), pp 442-455, 2002.
[9] Lanitis, C. Draganova C. and C. Christodoulou C. Comparing Different classifiers for automatic Age Estimation,. IEEE Transactions on Systems Man and Cybernetics, Part B,34(1), pp 621-629, 2004.
[10] Geng, Z.H. Zhou and K. Smith-Miles. Automatic Age Estimation Based on Facial Aging Patterns, IEEE Transactions of Pattern Analysis and Machine Intelligence, 29(12), pp. 2234-2240, 2004.
[11] Y. Fu and T.S. Huang. Human Age Estimation with Regression on Discriminative Aging Manifold, IEEE Transactions on Multimedia, 10(4), pp. 578-584, 2008.