# Efficient technique to estimate age using PCA & multi SVM classification

**Laiphrakpam Jibanpriya Devi [1] *, Dr. J L Mazher Iqbal [1]**

[1] *School of Electronics & Communication Engineering, Vel Tech Dr.RR & Dr..SR University, Chennai Avadi-600062*
*Corresponding author E-mail: mjivanpriya@gmail.com*

## Abstract

Human interaction with computer is recent trend in computer technology. In order to obtain age information, image-based age estimation systems have been developed using information from the human facial images. We develop a new technology which identify the characteristic of human being like age. Facial information study will lead us to identify age. While generic growth patterns that are characteristics of different age groups can be identified. In order to create an accurate algorithm for age classification, we build an appropriate datasets for training is build using SVM classification method. We build an application base on MATLAB software to estimate age based on the trained data. Feature of face is extracted using PCA method and stored the data in array matrix. The accuracy of the trained data is 95.65%. We have an average matching percentage of 92%. We have Euclidean distance calculation method to verify the matched data and we found 100% verified.

*Keywords*: *Facial Images; SVM; PCA Algorithm; Age Estimation; Multi SVM.*

## 1. Introduction

As the advent of technology is developing, biometrics is emerging as an interesting and necessary field of research in today's scenario. Research in Biometrics involves the process of identification and authentication of a person's identity. For identification purpose there are various mechanisms in biometrics via fingerprints, face recognition, iris detection, palm print etc. Face recognition in biometrics is widely used in the area of security for confirming the identification. There are different system build using Face Images are Authentication, Race, ethnic and gender [4]. One of the components of identification is age estimation. Age estimation involves the process of estimating the age of a person from his/her biological characteristics. Age estimation is an attracting research interests as recent trend. Mainly because of its wide potential applications including developing group specific human computer interfaces for human-machine system. Age estimation from face can be fruitful if any region can depict this change in age. Various regions located on the face are the area below the eyes, cheek, chin where the changes are proportional to the age of the person. Website which children do not allow to visit can be blocked easily. Online Education system can be authenticated using face recognition. Any criminal caught in action or after the action can be easily identified. Any unidentified dead or alive found can be tracked from their records. Age of unknown dead person found can easily be identified User can be discriminated accordingly when visiting a web page. Cloud user will be authenticated using Facial recognition and hence data can be secured. To find out how to build a web service that will perform authentication process in server. We have to build a web service that also accesses the webcam i.e. the hardware present in the user computer. This web service will perform preprocessing of image and send the data as a test data to the server and receive by a service API running in server.

## 2. Existing methodologies

Implementing algorithms that enable the estimation of a person's age, Based on features derived from his / her face image. Due to some disease like Parkinson, or other common diseases etc., age or gender identification is difficult. Image acquisition with proper Orientation [1]. Design for Image processing with Memory Management. To maintain Image quality during acquisition and pre-processing. Feature Extraction from Noisy Image which went through noisy network. Geometry of Face and Feature as wrinkle extraction is challenging, noise present in image can be consider as wrinkle, differentiating noise from the wrinkle is a challenging task. Problem in intrinsic age ambiguity [5]. The face recognition accuracy may vary with young faces when compared to old aged faces. Automatic Identifying Facial Landmarks is crucial. Hardware algorithm for image processing is crucial part as data size is big, huge memory require for real time image processing.

1) This algorithm can only detect ages from 20 years onwards. So test image's age is assumed to be 20+ by this algorithm.
2) The person's image is expression less so that no artificial skin folds and wrinkles are present on the face else its age predicted would be more than the actual.
3) The person is not wearing any kind of mask covering the region under the eyes (the test area).
4) The person is not wearing spectacles and also his eye lashes are not big enough to cover the test area. Also there should not be injuries, moles or marks present below the eyes since then the change in the intensity of consecutive pixels will be more and as such age will not be predicted correctly.

# 3. Proposed technique

To Motivate the Age- based access control - prevention of minors to access some internet pages, Age-specific human-computer interaction – such as adjusting text size for different age groups, Age-based indexing of face images – photo albums, Age-invariant person identification, Detecting child pornography[6].

Non generative approaches concentrate on deriving "age invariant signatures" from faces, Face Database, Age Derivation using PCA algorithm[2] .Without generating the model of face we can directly estimate the age of a person by looking at his/her face like features, facial textures for example hair on the face, wrinkles , dark spots etc. And here we will discuss about non-generative approach only regarding age related studies [6]-[8].
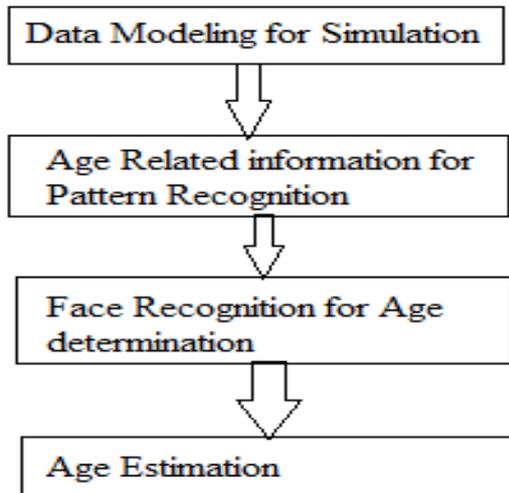


**Fig. 1:** Age Related Studies.

## 3.1. Block diagram

The Input image is human Face of different age, and different sex. These image are stored in a dataset in the form of feature extracted using PCA. The Block diagram representation is shown below:
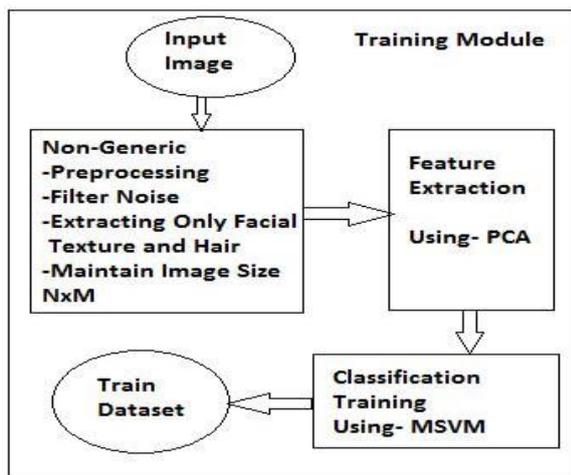


**Fig. 2:** Block Diagram of Proposed Technique.

In the above given block diagram train image and test image are the two database software's in which first the image is given to non –generative pre-processing which filters noise, extracts only facial texture –No hair, maintain image size to a define N x M. From here feature extraction is done by PCA and then image classification according to age group is done by multi class SVM and finally age is estimated [3].

## 3.2. Histogram equalization

The Histogram Equalization is an algorithm use for enhancing noisy image by improving the contrast of images and transforming the values of the image intensity so that the histogram of the output image approximately matches a specified histogram of the given image.

Result = A*prod (size (A))/n where n is the matrix size



**Fig. 3:** A) Input Test Image B) Output Histogram Equalization.

## 3.3. PCA algorithm

Principal Component Analysis (PCA) is a statistical procedure that uses an orthogonal to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components. The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. Following are steps involved in this algorithm:

Step 1: Column or row vector of size $N^2$ represents the set of M images (B1, B2, B3…BM) with size N*N.

Step 2: the average image by vector (W) is different for each trainee image

$$Wi = Bi - \mu \tag{1}$$

Step 3: Total Scatter Matrix or Covariance Matrix is calculated from matrix of W and result normalizes by $N – 1$, if $N > 1$, where N is the number of observations.

$$Covariance \ (x1, x2) = E \ [(x1−\mu1) ∗ (x2−\mu2)] \tag{2}$$

Step 4: Measure the eigenvectors UL and Eigen value λL of the covariance matrix C.

Step 5: For image classification, this feature space can be utilized. Measure the vectors of weight

$\Omega T= [w1, w2… wM']$ whereby,

$$Hk = UkT \ (B - \mu), \ k = 1, \ 2…M' \tag{3}$$

## 3.4. Multi class SVM

SVMs are inherently two-class classifiers. The traditional way to do multiclass classification with SVMs is to use one of the methods [8]. In particular, the most common technique in practice has been to build one-versus-rest classifiers and to choose the class which classifies the test datum with greatest margin. Another strategy is to build a set of one-versus-one classifiers, and to choose the class that is selected by the most classifiers. While this involves building classifiers, the time for training classifiers may actually decrease, since the training data set for each classifier is much smaller. However, these are not very elegant approaches to solving multiclass problems. A better alternative is provided by the construction of multiclass SVMs, where we build a two-class classifier over a feature vector derived from the pair consisting of

the input features and the class of the data [4]. At test time, the classifier chooses the class [3]. Multiclass SVM aims to assign labels to instances by using support vector machines, where the labels are drawn from a finite set of several elements the dominant approach for doing so is to reduce the single multiclass problem into multiple binary classification problems. Common methods for such reduction include: Building binary classifiers which distinguish between one of the labels and the rest or between every pair of classes. Classification of new instances for the one-versus-all case is done by a winner-takes-all strategy, in which the classifier with the highest output function assigns the class.

## 4. Results and discussion

This feature are collected from all different age human and stored as train dataset. The process of training the data SVM using PCA algorithm. The test dataset is also collected as same process of training dataset. So finally we have done PCA based age estimation using SVM classification and improved the performance and accuracy compare to existing design and identification rate is close to 100%.

Figure 4. Age Estimation Application Result Page

The result is based on the ROC curve analysis. The ROC curve is derived based on true positive rate with respect to false positive rate, where the values are calculated from the algorithm by testing several test images and calculated false acceptance ratio and false rejected ratio.
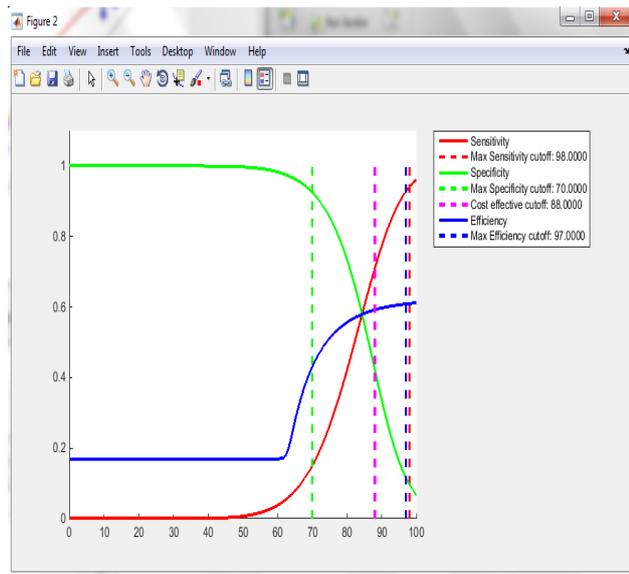
$$FRR = 1 - 1/FAR \qquad (4)$$



**Fig. 5:** Roc Curve Analysis.

a) Max Sensitivity Cut-off point= 101.00 b) Max Specificity Cut-off point= 60.00 c) Cost effective Cut-off point= 80.00 d) Max Efficiency Cut-off point= 100.00.
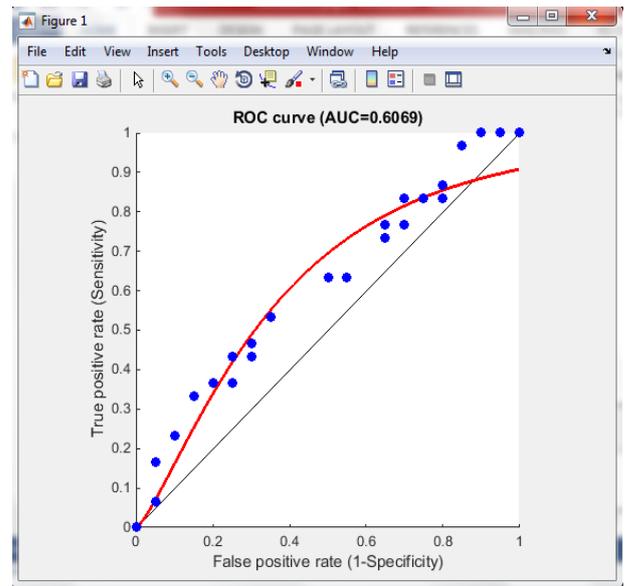


**Fig. 6:** Roc Curve.

**Table 1:** Roc Curve Train Data Table

| Cut -off | Sensi- tivity | Speci- ficity | Effi- cien- cy | Wri nkle Ra- tio | Smoot hness Ratio | Posi- tive Pre- dic- tion | Nega- tive Pre- dic- tion |
|---|---|---|---|---|---|---|---|
| 70 | 0 | 1 | 0.4 | 1.33 33 | 0.9825 | 98.49 | 2 |
| 71 | 0.066 7 | 0.95 | 0.42 | 1.33 33 | 0.9825 | 98.49 | 2.03 |
| 72 | 0.166 7 | 0.95 | 0.48 | 3.33 33 | 0.8772 | 99.39 | 2.27 |
| 73 | 0.233 3 | 0.9 | 0.5 | 2.33 33 | 0.8519 | 99.13 | 2.34 |
| 74 | 0.333 3 | 0.85 | 0.54 | 2.22 22 | 0.7843 | 99.09 | 2.54 |
| 75 | 0.366 7 | 0.8 | 0.54 | 1.83 33 | 0.7917 | 98.9 | 2.51 |
| 76 | 0.366 7 | 0.75 | 0.52 | 1.46 67 | 0.8444 | 98.63 | 2.36 |
| 78 | 0.433 3 | 0.75 | 0.56 | 1.73 33 | 0.7556 | 98.84 | 2.63 |
| 81 | 0.433 3 | 0.7 | 0.54 | 1.44 44 | 0.8095 | 98.61 | 2.46 |
| 83 | 0.466 7 | 0.7 | 0.56 | 1.55 56 | 0.7619 | 98.71 | 2.61 |
| 85 | 0.533 3 | 0.65 | 0.58 | 1.52 38 | 0.7179 | 98.68 | 2.76 |
| 86 | 0.633 3 | 0.5 | 0.58 | 1.26 67 | 0.7333 | 98.41 | 2.71 |
| 87 | 0.633 3 | 0.45 | 0.56 | 1.15 15 | 0.8148 | 98.26 | 2.44 |
| 88 | 0.733 3 | 0.35 | 0.58 | 1.12 82 | 0.7619 | 98.22 | 2.61 |
| 89 | 0.766 7 | 0.35 | 0.6 | 1.17 95 | 0.6667 | 98.3 | 2.97 |
| 90 | 0.766 7 | 0.3 | 0.58 | 1.09 52 | 0.7778 | 98.17 | 2.56 |
| 92 | 0.833 3 | 0.3 | 0.62 | 1.19 05 | 0.5556 | 98.31 | 3.54 |
| 93 | 0.833 3 | 0.25 | 0.6 | 1.11 11 | 0.6667 | 98.2 | 2.97 |
| 94 | 0.833 3 | 0.2 | 0.58 | 1.04 17 | 0.8333 | 98.08 | 2.39 |
| 95 | 0.866 7 | 0.2 | 0.6 | 1.08 33 | 0.6667 | 98.15 | 2.97 |
| 97 | 0.966 7 | 0.15 | 0.64 | 1.13 73 | 0.2222 | 98.24 | 8.41 |
| 98 | 1 | 0.1 | 0.64 | 1.11 11 | 0 | 98.2 | 100 |
| 99 | 1 | 0.05 | 0.62 | 1.05 26 | 0 | 98.1 | 100 |
| 100 | 1 | 0 | 0.6 | 1 | 0 | 98 | 100 |

# 5. Conclusion

In this paper we have done PCA based age estimation using Multi SVM classification. In future work we will work on the generative approach for the age estimation in which we have to generate the model of the face of the person to estimate the age. This can be used for ethnic group of facial images.

# 6. Limitation of this algorithm

Ageing of human relates to other factor like use of intoxicant. Expose to working environment. Condition of the family and area of the person living. Person face may look more ageing due the environmental factor which leads to error of estimation. But we have achieved the estimation accuracy of 90%. If the face is coated with cosmetic then it is difficult to estimate in such case our error of estimation is of min of 2 years to max of 5 years.

# References

[1] Richang Hong,and Shuicheng Yan, Zhenzhen Hu, Yonggang Wen Jianfeng Wang, Meng Wang," Facial Age Estimation With Age Difference", IEEE Transactions On Image Processing, Vol. 26, No. 7, July 2017.

[2] Petra Grd,Miroslav Bača,"Creating a Face Database for Age Estimation and Classification",MIPRO 2016,Opatija,Croatia,May 30 - June 3, 2016. https://doi.org/10.1109/MIPRO.2016.7522353.

[3] Julianson Berueco, Kim Lopena, Arby Moay, Mehdi Salemiseresht, and Chuchi Montenegro,"Age Estimation Using Support Vector Machine– Sequential Minimal Optimization",Journal of Image and Graphics, Volume 2, No.2, December 2014.

[4] G. Guo and X. Wang, "A study on human age estimation under facial expression changes. In Proc. IEEECVPR, pages 2547–2553, 2012.

[5] Yu Zhu* 1, Yan Li* 1,2, Guowang Mu1, and Guodong Guo1, "A Study on Apparent Age Estimation",IEEE International Conference on Computer Vision Workshops, 2015.

[6] Louis Quinn, Margaret Lech, "Multi-Stage Classification Network for Automatic Age Estimation from Facial Images", IEEE Transaction, 2015

[7] Imed Bouchrika, Nouzha,Harrati,Ammar Ladjailia and Sofiane Khedairia,"Age Estimation from facial Images based on Hierarchical Feature Selection",16th international conference on Sciences andTechniques of Automatic control & computer engineering STA'2015, Monastir, Tunisia, December 21-23, 2015. https://doi.org/10.1109/STA.2015.7505156.

[8] Young H. Kwon, Niels da Vitoria Lobo,"Age Classification from Facial Images, Computer Vision and Image Understanding",74(1),pp.1–21,1999. https://doi.org/10.1006/cviu.1997.0549.

[9] Y. Fu, G. Guo, and T. Huang, "Age synthesis and estimation via faces: A survey", IEEE Trans. PAMI, 32(11):1955 –1976, Nov. 2010. https://doi.org/10.1109/TPAMI.2010.36.

[10] XinGeng, Zhi-Hua Zhou, Senior Member, IEEE, and Kate Smith-Miles, Senior Member, IEEE,"Automatic Age Estimation Based on Facial Aging Patterns"–IEEE Transactions On Pattern Analysis and Machine Intelligence, Vol. 29, No. 12, December 2007.

[11] Chin-Teng Lin, Dong-Lin Li, Jian-Hao Lai, Ming-Feng Han and Jyh-Yeong Chang, "Automatic Age Estimation System for Face Images" - International Journal of Advanced Robotic Systems, 29 Aug 2012. https://doi.org/10.5772/52862.

[12] Dr.J L Mazher Iqbal, K. Iswarya Lakshmi, "Object Detection and Tracking Using Thermal Camera", International Journal of Pure and Applied Mathematics, Volume 114 No. 10 2017, 87-95.

[13] Dr.J L Mazher Iqbal, Shaik Shakir Basha, ,Real Time 3D Depth Estimation and Measurement of Un-calibrated Stereo and Thermal Images·, 2017 International Conference on Nascent Technologies in the Engineering Field (ICNTE-2017), 978-1-5090-2794-1/17/$31.00 ©2017 IEEE.

[14] J.L. Mazher Iqbal, M. Suriya Parveen, S. Arun, , Image stitching and 2D to 3D Image Reconstruction for Abnormal Activity Detection·, International Journal of Computer Applications (0975 – 8887) Volume 133 – No.17, January 2016.

[15] J.L.Mazher Iqbal, J.Lavanya and S.Arun, ,Abnormal Human Activity Recognition using Scale Invariant Feature Transform·, International Journal of Current Engineering and Technology, Vol.5, No.6, Dec 2015.

[16] Ravindraiah, Dr. J. L. Mazher Iqbal (2014), Hard Exudates Detection in Proliferative Diabetic Retinopathy using Gradient Controlled Fuzzy C MEANS Clustering Algorithm, GESJ: Computer Science and Telecommunications,Vol 4 (44), PP 27-31.

[17] R. Ravindraiah, Dr. J. L. Mazher Iqbal (2014), Segmentation of Diabetic Retinopathy Images through a Subjective Approach, GESJ: Computer Science and Telecommunications, VolNo.4 (44), PP 32-34, 2014.