Comparing Accuracy of K-Nearest-Neighbor and Support-Vector-Machines for Age Estimation

Anchal Tomar *[1], Anshika Nagpal *[2]

*1 M.tech student at M.I.ET College, Dr. A.P.J. Abdul Kalam Technical University (APJAKTU) Uttar Pradesh, India

*2 Assistant Professor at M.I.E.T, APJAKT University, India

Abstract

In term for a robot machine or a computer to perform any task, it must recognize the problem and then act on it. Given a a picture a computer must have the capacity to group/classify what the picture represents exactly. While this is a genuinely simple for humans, it is not an ethereal task for computers. Computers must experience a progression of steps in order to classify a single picture. So, there are different techniques, tools and procedures are introduced. And here explained two algorithms: (SVM) Support Vector Machine and K Nearest Neighbor (KNN) classification. Support vector machine is a model for measurements and software engineering which perform supervised learning, methods that are used for analysing data and recognize various patterns. SVM is used for classification and regression analysis. Likely k-nearest neighbor algorithm is also a classification algorithm but it is used to classify data using training examples. In this paper SVM and KNN algorithm are explained and also evaluate which one has good accuracy and in which conditions.

Keywords —*Support Vector-Machine (SVM), supervised learning, regression analysis, K-Nearest Neighbor (KNN), classification.*

I. INTRODUCTION

People can gather a wide variety of information from a face image, including identity, sexual orientation, gender, and ethnicity [3]. Age estimation is one of the errands of facial image classification. It can be defined as characterized of a person's age or age group from face pictures. Facial characteristic focuses can be defined as a standard reference points on human face used by researchers as a part of identity to recognize a person's face, or for this situation, to estimate the age of human utilizing different classifiers [1, 2].

Before, numerous classifiers have been developed by different analysts. These strategies incorporate naïve Bayes classifier, k-nearest neighbours(KNN), Gaussian mixture model, support vector machines (SVM),decision tree and(RBF) radial basis function classifiers [4]. These classifiers are used as a part of algorithms that include object recognition. But object recognition is challenging for few reasons [5]. The first and very important reason is that there are around 15,000 to 30,000 different object classifications. Another reason is the perspective variety where numerous objects look unique in relation to various points. However we can say that different classifiers are good for different analysis, but here we will discuss about KNN and SVM results for age classification.

II. RELATED WORK

With the unstable development of information on the distinctive area like education, businesses and others required to extract knowledge from data in such manner to explore much knowledge from that data. For that purpose we identify the most frequently used data mining algorithm and we found that for the classification purpose researchers are go through the SVM and KNN [3].So, for age classification which classifier works best is the question to find out .As every algorithm consists its own set of rules or steps, it also involves some basic steps [5, 6].

The three fundamental steps for age classification include feature extraction, boosting and multi scale detection. With the end goal of classification, geometric and wrinkle features are used in the framework [1]. In another phase i.e. feature extraction phase, there happens estimation of two geometric features. Geometric features are as the ration of partitions between eyes, noses, and mouths. For assessing the degrees of facial wrinkles, it is important to characterize three different wrinkle features [4]. Classification process is mainly a use of K-means clustering algorithm. In the given diagram (figure 1) complete process of the system is explained.

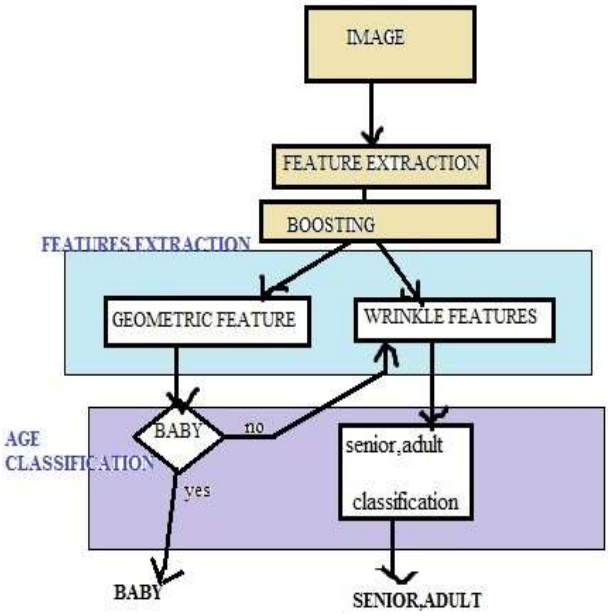

Figure 1:Process of the system

III.PROPOSED WORK

For proposed work to achieve given solution we will use the KNN and SVM implementation is used on different age group images. These algorithms implementation is done through MATLAB. Based on their accuracy, processing speed and other factors, compare the best results. Whole system process is explained in figure1.

Face detection algorithm is isolated into three essential steps.

A. Feature Extraction

It is clear that feature is critical to any substance detection algorithm [5]. With the end goal of face detection, a considerable measure of elements can be used like as the topology of eye and nose and eyes, nose [8]. Important main key issue of any characterization frameworks is to find an arrangement of reliable features as the premise for classification [9]. By and large these elements cam be classified into two classifications. These are wrinkle features and geometric features.

.

1) ***Wrinkle features:*** Critical property of wrinkle features is that it decides the age of a human being. As the age continues on increasing, wrinkles on face end up being clearer [6]. Matured people consistently have clear wrinkles on the face in the following areas as mentioned here:
   a) The temple has horizontal furrows [1].
   b) The eye corners have crow's feet.
   c) The cheeks have sickle shaped pockets, clear cheekbones and significant lines between the cheeks and the upper lips [11].

2) ***Geometric features:*** As showed by the examinations of facial representation and emotional cosmetics, there happens a considerable measure of change in the facial features as the age keeps on growing with person [7,6]. In this stage, worldwide features in combination with the grid features are extracted from the facial pictures [12]. The global features include the distance between chin to eye, two eye balls, eye to lip and nose tip to eye.

B. Boosting

This boosting procedure make learning to be effective and all around sorted out. Specifically the boosting works as follow as [1, 5]:

a) Firstly From a given dataset, take a unique and straightforward classifier and after that find out the errors it make.
b) The second step is to reweight the dataset and after that provide the data where it made errors.
c) Take the second straight forward classifier into consideration based on the reweighted dataset.
d) Consolidate the first and the second classifier, reweight the whole data and check where the data make errors.

IV. ANAYSIS

A. Support Vector Machine (SVM)

A SVM is a linear or non-linear classifier, which is a mathematical function that can distinguish two different kinds of objects [13]. These objects fall into classes; this is not to be confused for an implementation. Consequently the objective of SVM Classification is to deliver a model, based on the training data, which will be capacity to predict class labels of the test data precisely [12]. In this algorithm, we plot each data item as a point in n-dimensional space (where n is number of features you have) with the value of each feature being the value of a particular coordinate [9]. Then, we perform classification by finding the hyper-plane that differentiates the two classes very well as shown in below figure2.

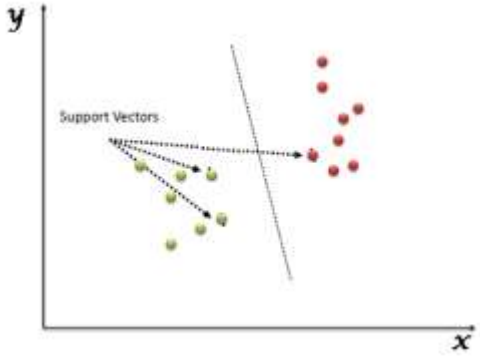

Figure2: SVM classifier in n-dimensional space

Advantages and disadvantages of SVM

A primary preferred standpoint of SVM order is that SVM performs well on datasets that have numerous attributes, even when there are just a couple cases that are accessible for the training process [12]. Be that as it may, a few disadvantages of SVM classification include limitations in speed and size during both training and testing phase of the algorithm and the selection of the kernel function parameters.

B. K-Nearest Neighbor (KNN)

For age classification, the k-nearest neighbor algorithm is a technique for classifying objects based on closest training examples in the feature space. Not only this, KNN is also a sort of instance-based learning, or lazy learning where the function is only approximated locally and all computation is conceded until classification [3].

Training process for this algorithm just comprises of storing feature vectors and labels of the training pictures [14]. In the classification procedure, the unlabelled question point is basically doled out to the label of its k nearest neighbors. Ordinarily the object is characterized based on the labels of its k nearest neighbors by majority share vote. In the event that k=1, the object is essentially classified as the class of the object nearest to it[5]. As when there are only two classes, k must be a odd integer. Be that as it may, there can at present be ties when k is an odd whole number when performing multiclass characterization.

The k-NN algorithm can likewise be adjusted for use in assessing consistent variables. Such implementation utilize an inverse distance weighted average of the (KNN) k-nearest multi-variate neighbors[16]. This calculation functions as follows:

a) Calculate Euclidean distance from main target plot to those that were sampled.
b) Order samples checking calculated distances.
c) Select heuristically optimal (knn)k nearest neighbor based on RMSE used by cross validation strategy..
d) Determine an inverse distance weighted average using the k-nearest multivariate neighbors.

Advantages and disadvantages of KNN

A principle favourable position of the KNN algorithm is that it works well with multi-modal2 classes in light of the fact that its decision is depend on a small neighbourhood of similar target [8]. Subsequently, regardless of the fact that the target class is multi-modal, the algorithm can in any case lead to great precision. However a major disadvantage of the KNN algorithm is that it utilizes every feature similarly in computing a part of processing for similitude [15]. This can prompt to classification errors, particularly when there is just a small subset of features that are helpful for classification [7].

V.IMPLEMENTATION AND RESULTS

A. Implementation

Here we have used MATLAB as our computer language. Here dataset contains images from different age group. The results are mainly shown in confusion matrix which is useful for comparison of classifiers performanance [16]. Results of KNN (In figure3) and SVM is explains in result section. Performance analysis is conduct under accuracy of result, search time for data set, memory used to process [8].

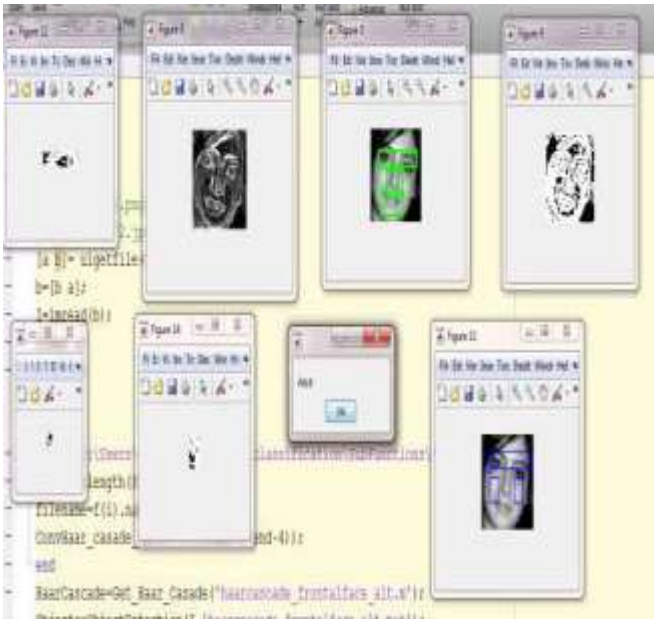

Figure3: Implementation and execution of KNN

1) Accuracy

Derived using the formula

Accuracy = (correct prediction/ total values (supplied)) * 100

2) Search time

It is the time required in predicting values like wrinkle and geometric features of a image.

3) Memory used

Memory utilized for this framework is characterized as required main memory for complete working of the system to detect the age of a person [3].

B. Results

1) Accuracy

Accuracy of k-NN is kept high in most of the cases. But as size of dataset increases we can see accuracy of both system decreases [2]. But here shows some

result and we must say in overall accuracy KNN has more effective work. Some accuracy results are shown in table 1

Age Group	SVM(in%)	k- NN(in%)
senior	61.32	58.1
adult	35.60	36.15
child	29.51	55

Table 1:Table showing the accuracy of both algorithm

2) Search time

Time taken to calculate all required values for result is defined as search time [16]. Search time results are shown in table 2.

Data set size	SVM	k- NN
1000	.0651	.259
500	.0663	.525
200	.0640	.524
100	.0640	.228
50	.0640	.102

Table 2. Table showing the Search time

k-NN is keep high in most of the cases. But as size of dataset increases we can see time consumed for predict values of KNN system increases. But in case of SVM remain constant [5].

3) Memory use:

Memory used by algorithms is remaining constant in all cases.

VI.CONCLUSION

In this work, we used two different classifiers for image classification [3]. With a proper observation we have results as nearest neighbor classifiers typically have good predictive accuracy in low dimensions, yet may not in high dimension data sets. They have high memory usage, and are not easy to interpret. Overall after execution we found that K-NN is a quit good classifier. Furthermore, support vector machine has been picked as it represents a framework both interesting from a machine learning perspective [5]. But MATLAB's SVM classifier does not support multiclass classification and only supports binary classification. Its accuracy is less as compared to KNN in many executions.

However the performance of different classification methods still depends extraordinarily on the general characteristics of the data to be classified [11]. The definite relationship between the information to be arranged and the execution of different characterization techniques still stays to be found. To determine the best classification method for a certain dataset we need to use trial and error in different field to find the best performance [13].

Future Work

In future, we can utilize more various types of classifications that would be troublesome for the computer to classify look at more complex classifiers. Another area for research would be to find certain characteristics in various image categories that make one classification method better than another. We can implement best classifier in real life applications and can apply rules and regulation according to the age classification of a person like driving licence, movie theatre etc.

REFERENCES

[1] A.Lanitis and C.J.Taylor, "Towards Automatic Face Identification Robust to Ageing Variation", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol.24, no.24, p.442-455, 2002.

[2] Facial age estimation, http://www.scholarpedia.org/article/Facial_Age_Estimation, accessed 13.02.2013F

[3] Lloyd-Williams,M. ―Case studies in the data mining approach to health information analysis‖, Knowledge Discovery and Data Mining (1998/434), IEEEColloquium on,8May1998, 1996 Page(s): 1/1 -1/4

[4] D.A. Forsyth and J. Ponce, Computer Vision, A Modern Approach, Prentice Hall, 2003.

[5] N.Ramanathan and R. Chellappa, "Face verification across age progression", in Proc. IEEE Conf. Computer Vision and Pattern Recognition, San Diego, CA, 2005, vol.2, pp.462-469.

[6] N.Ramanathan and R. Chellappa, "Modeling Age Progression in young faces", in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), vol.1, pp.387-394, 2006.

[7] G. Guo, A. Jain, W. Ma, and H. Zhang, "Learning similarity measure for natural image retrieval with relevance feedback," IEEE Trans. Neural Netw., vol. 13, no. 4, pp. 811–820, Apr. 2002.

[8] G. Guo and S. Li, "Content-based audio classification and retrieval by support vector machines," IEEE Trans. Neural Netw., vol. 14, no. 1,pp. 209–215, Jan. 2003.

[9] G. Guo, S. Li, and K. Chan, "Support vector machines for face recognition," Image Vis. Comput., vol. 19, no. 9-10, pp. 631–638, 2001.

[10] G. Guo, S. Li, and K. Chan, "Face recognition by support vector machines," in Proc. IEEE Conf. FG, 2000, pp. 196–201.

[11] X. F. He and P. Niyogi, "Locality preserving projections," presented at the NIPS, 2003.

[12] T. Joachims, "Making large-scale SVM learning practical," in Advances in Kernel Methods— Support Vector Learning, B. Schölkopf, C. Burges, and A. Smola, Eds. Cambridge, MA: MIT Press, 1999.

[13] Y. Kwon and N. Lobo, "Age classification from facial images," Comput. Vis. Image Understand., vol. 74, no. 1, pp. 1–21, 1999.

[14] [14] A. Lanitis, C. Taylor, and T. Cootes, "Toward automatic simulation of aging effects on face images," IEEE Trans. Pattern Anal. Mach. Intell.,vol. 24, no. 4, pp. 442–455, Apr. 2002.

[15] A.Lanitis, C.Draganova, and C.Christodoulou, "Comparing different classifiers for automatic age estimation", IEEE Trans.Syst.Man, Cybern.B, Cybern, vol34, no.1, pp.621-628, Feb.2004.

[16] Chin Heng Wan, Lam Hong Lee, rajprasad Rajkumar, Dino Isa(2012). A Hybrid text classification approach with low dependency on parameter by integrating K-nearest neighbour and support vector machine, Expert Systems with Applications, elsevier journal,39 ,11880-11888.