A NEW ACTIVE LEARNING TECHNIQUE USING FURTHEST NEAREST NEIGHBOUR CRITERION FOR K-NN AND SVM CLASSIFIERS

Indu. M. T[1], Prof. S. Bhuvana[2], Prof. P. Tamije Selvy[3] and Dr. R. Radhakrishnan[4]

[1]PG Student, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore
indumt.anjali@gmail.com

[2]Assistant Professor, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore
bhuvana_anju@rediffmail.com

[3]Assistant Professor, Department of Computer Science and Engineering, Sri Krishna College of Technology, Coimbatore
tamijeselvy@gmail.com

[4]Principal, Sri Ramakrishna Engineering College, Coimbatore
rlgs14466@gmail.com

ABSTRACT

Active learning is a supervised learning method that is based on the idea that a machine learning algorithm can achieve greater accuracy with fewer labelled training images if it is allowed to choose the image from which it learns. Facial age classification is a technique to classify face images into one of the several predefined age groups. The proposed study applies an active learning approach to facial age classification which allows a classifier to select the data from which it learns. The classifier is initially trained using a small pool of labeled training images. This is achieved by using the bilateral two dimension linear discriminant analysis. Then the most informative unlabeled image is found out from the unlabeled pool using the furthest nearest neighbor criterion, labeled by the user and added to the appropriate class in the training set. The incremental learning is performed using an incremental version of bilateral two dimension linear discriminant analysis. This active learning paradigm is proposed to be applied to the k nearest neighbor classifier and the support vector machine classifier and to compare the performance of these two classifiers.

KEYWORDS
Active learning, age classification, machine learning, supervised learning, support vector machine

1. INTRODUCTION

1.1. Active Learning

Active learning is a technique that is well-motivated in many modern machine learning problems [2]. It is used in situations where unlabeled data may be abundant or easily obtained, but labels are difficult, time-consuming, or expensive to obtain. Active learning may also be called as “query learning” or “optimal experimental design”. The main idea in active learning is

that a machine learning algorithm can achieve greater accuracy with fewer training labels if it is allowed to choose the data from which it learns.
For sophisticated supervised learning problems, the labelled instances are very difficult to obtain. Active learning systems attempt to overcome this difficulty by asking queries in the form of unlabeled instances to be labelled by an oracle (e.g., a human trainer). The active learner aims to achieve high accuracy for a given problem using as few labelled instances as possible, thereby minimizing the cost of obtaining labelled data [2].

An active learner begins with a small number of instances in the labelled training set, requests labels for one or more carefully selected instances, learns from the query results, and then uses its new knowledge to choose which instances are to be queried next. Once a query has been made, there are usually no additional assumptions on the part of the learning algorithm. The new labelled instance is simply added to the labelled set and the learner proceeds from there in a standard supervised way.

Pool-based active learning for classification was introduced by Lewis and Gale (1994) [4]. A large quantity of unlabeled data will be readily available in most of the classification problems. The active learner has access to this pool of unlabeled data and can request the class label for a certain number of instances in the pool from the human annotator. The main issue with active learning is finding a way to choose good requests or queries from the pool.

1.2. Classification

In machine learning and pattern recognition, classification refers to an algorithmic procedure for assigning a given piece of input data into one of a given number of categories. An example would be assigning a given email into "spam" or "non-spam" classes or assigning a diagnosis to a given patient as described by the observed characteristics of the patient like gender, blood pressure, presence or absence of certain symptoms, etc. An algorithm that implements classification is known as a classifier. The piece of input data is formally termed as an instance, and the categories are termed as classes.

Classification normally refers to a supervised procedure, i.e., a procedure that learns to classify new instances based on learning from a training set of instances that have been properly labelled by hand with the correct classes. The corresponding unsupervised procedure is known as clustering, and involves grouping data into classes based on some measure of inherent similarity (e.g., the distance between instances, considered as vectors in a multi-dimensional vector space).

In supervised learning, we have training examples and test examples. A training example is an ordered pair <x, y> where x is an instance and y is a label. A test example is an instance x with unknown label. The goal is to predict labels for test examples. The name “supervised” comes from the fact that some supervisor or teacher has provided explicit labels for training examples.
The different types of classifiers include k- Nearest Neighbour classifier, Bayesian Classifier, Neural Network classifier, Support Vector Machine classifier, etc. Here we concentrate on k-NN and SVM classifiers. K-NN depends upon distance measures while SVM depends upon probability estimates for classification.

2. RELATED WORK

2.1. Approaches for Active Learning

A list of techniques that can be used for selecting the query instance in active learning was presented by Burr Settles [2]. Uncertainty sampling is a commonly used query strategy in the case of probabilistic models. The active learner queries the instance about which it is least certain how to label.
The Query-By-Committee method involves maintaining a committee of models which are all trained on the current labeled set, but representing different classes. When an instance is queried, each committee member is allowed to vote on the labeling of the query instance. The most informative query is considered to be the instance about which the models most disagree.
A decision-theoretic approach selects the instance that would impart the greatest change to the current model if we knew its label. Another decision-theoretic approach aims to measure how much the generalization error, caused due to selecting an incorrect instance, is likely to be reduced. The idea is to query the instance with minimal expected future error.

Most of the above active learning techniques work well only with binary classification tasks. For multi-class problems, the main problem has to be split down into many independent binary sub problems.

A multi-label support vector machine active learning method was proposed by Xuchun Li, Lei Wang, and Eric Sung [5] that includes two selection strategies namely Max Loss strategy and Mean Max Loss strategy. SVM can obtain the best classification performance by minimizing the expected loss over the data distribution. Hence, they proposed the selection strategies based on the idea that the decrease of the expected loss between two contiguous learning cycles can be regarded as an indicator of the improvement of classification performance.

A probabilistic variant of the k-nearest neighbour method was proposed for active learning in large multi-class scenarios by Prateek Jain and Ashish Kapoor [6]. This method learns an accurate metric/kernel function over the input space that can be used for classification and similarity search. It defines a probability measure based on the pair wise distances between the data points.

2.2. Facial Age Classification Methods

The first work on facial age classification is based on cranio-facial development theory and skin wrinkle analysis [7]. This approach was applied to the classification of facial images into three age-groups: babies, young adults and senior adults. The primary features of the face, which include eyes, nose, mouth, chin, virtual-top of the head, and the sides of the face, are found first, followed by the analysis of the secondary features of the face. The ratios that distinguish the different age-groups are computed from these features. In secondary feature analysis, wrinkles are detected and analyzed. This step permits the distinguishing of seniors from those in the two younger categories.

The Aging Pattern Subspace method was proposed in [8]. The basic idea is to model the aging pattern by learning a representative subspace. The aging pattern is defined as a sequence of personal aging face images. The proper aging pattern for an unseen face image is determined by the projection in the subspace that can best reconstruct the face image with minimum reconstruction error. Finally the age associated with that position is returned as the estimated age for that image.

T. F. Cootes, G. J. Edwards, and C. J. Taylor developed the Active Appearance Model (AAM) approach. They devised a combined shape and intensity model to represent face images. Age is modelled as a function of the face model parameters. The aging function includes linear, quadratic, and cubic functions [3]. A relationship between model parameter displacements and the residual errors induced between a training image and a synthesised model example is learnt in the training phase. In the testing phase, the current residuals are measured and the model is used to predict the changes to the current parameters.

A comparative study of the different classifiers that can be used for age estimation was carried out by Andreas Lanitis, Chrisina Draganova, and Chris Christodoulou [9]. They compared a classifier based on the use of quadratic functions for modeling the relationship between face model parameters and age, a shortest distance classifier, and an artificial neural network based classifier. Quadratic functions that explain the relationship between the face parameters and the age of a person in the corresponding face image were used for transforming a given set of face parameters to an estimate of the age. In a Shortest Distance Classifier, based on the training data, the distributions of face parameters corresponding to a certain age were defined. A new set of face parameters can be assigned to the closest distribution in order to estimate the age. Supervised neural networks have been trained with a set of face parameters and their corresponding ages so that given an unknown set of parameters they produce an estimate of the age of the person in the corresponding face image.

A craniofacial growth model that characterizes growth related shape variations observed in human faces during formative years was proposed by Narayanan Ramanathan and Rama Chellappa [10]. They characterized facial growth by means of growth parameters defined over facial landmarks often used in anthropometric studies. Constraints based on proportion indices such as the intercanthal index, nasal index, etc. result in linear constraints on the growth parameters while constraints based on proportion indices such as eye fissure index, orbital width index, etc. result in non-linear constraints on the growth parameters.

A combination of the method of error-correcting output codes with boosting using a decision tree based classifier or binary support vector machine classifier was proposed to solve the age categorization problem by J. G. Wang, W. Y. Yau, and H. L. Wang [11]. Gabor and Local Binary Pattern aging features are extracted and combined at the feature level to represent the face images.

Y. Fu and T. S. Huang constructed a low-dimensional manifold from a set of age-separated face images and used linear and quadratic regression functions on the low-dimensional feature vectors from the respective manifolds to estimate the age of a face [12]. They proposed that the aging patterns can be effectively extracted from a discriminant subspace learning algorithm and visualized as distinct manifold structures.

However most of the existing systems for facial age classification works using only labelled training images. The proposed approach allows the use of both labelled and unlabeled facial images in the training set.

3. PROPOSED SYSTEM

A study of the system proposed by Jian Gang Wang, Eric Sung, and Wei Yun Yau [1] is carried out by applying it to k-NN and SVM classifiers. The system applies an active learning approach to facial age classification which allows a classifier to select the data from which it learns. The age classification problem considered here is the classification of facial images into four predefined age groups: children, teenager, adult and senior adult. The classifier is initially

trained using a small pool of labeled training images. This is achieved by using the bilateral two dimension linear discriminant analysis (B2DLDA). The B2DLDA reduces the given training images to lower dimensional representation. Next the most informative unlabeled image is found out from the unlabeled pool using the furthest-nearest neighbor criterion (FNN). The selected image is labeled by the trainer and added to the training set. The calculations done by the B2DLDA are updated incrementally using an incremental B2DLDA (IB2DLDA) to incorporate the information from the newly added images. This iterative procedure continues till the unlabeled pool becomes empty. Finally the classification accuracy is tested with both k-NN classifier and SVM classifier. An overall process of the system is shown in Figure 1.

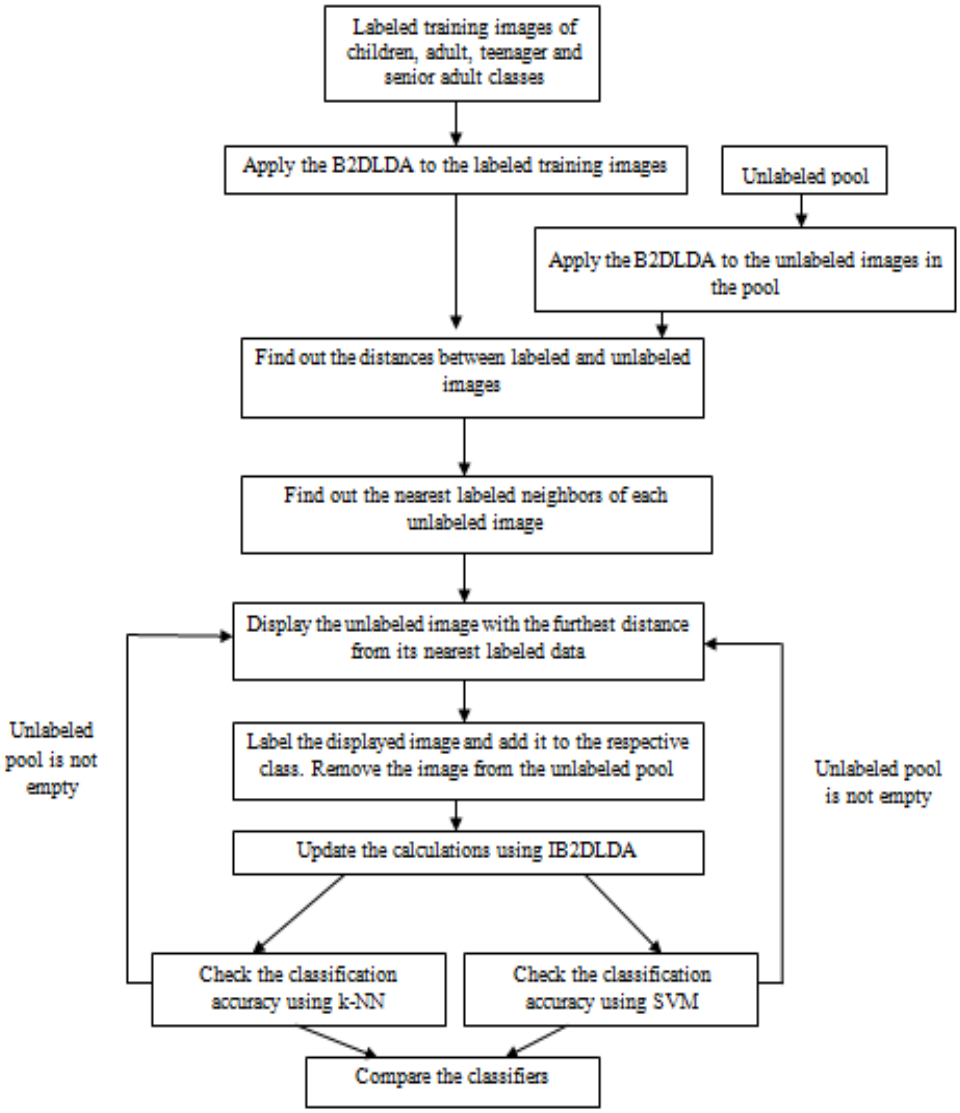

Figure 1. Overall process of the system

The input to the system is a pool of labeled and unlabeled facial images. These images are first reduced using B2DLDA. The distances between labeled and unlabeled images are then calculated. Then the FNN criterion selects the unlabeled image which has the greatest distance to its nearest labeled neighbor. This image is displayed to the trainer who labels the image. The newly labeled image is added to the training pool and all the calculations done in the first

iteration are updated using IB2DLDA. The classification accuracy is tested using a kNN classifier and the iterative learning is continued till the unlabeled pool becomes empty. The entire active learning process using FNN can be applied to an SVM classifier also to check its accuracy.

3.1. Dimension Reduction Using B2DLDA

High-dimensional data that require more than two or three dimensions to represent the data can be difficult to interpret. Hence reduced dimensional representation of the training images has to be prepared. This is done using the bilateral two dimension linear discriminant analysis (B2DLDA) as described in [1]. The input to this step is a collection of labeled images belonging to four predefined classes: children, teenager, adult, and senior adult. The outputs are the reduced dimensional representations of input images. Left and right transformation matrices are created by finding out the eigen vectors of the product of between class scatter matrix and the inverse of within class scatter matrix. The reduced dimensional representations are obtained by multiplying original image matrices with these transformation matrices.

Let *N* be the number of classes, *T* be the total number of training samples and n_k be the number of training images in class *k*. Let X_i^k denote the ith training image belonging to class *k*. B2DLDA computes the set of reduced dimensional representations of input images as follows:

1. Calculate the mean matrix $\overline{X_k}$ of each class *k*, *k=1,2,..,N*.

$$\overline{X_k} = \left({}^{1}/_{n_k}\right)\sum_{i=1}^{n_k} X_i^k \tag{1}$$

2. Compute the global mean *M* of all the training samples.

$$M = \sum_{k=1}^{N}\left({}^{n_k}/_{T}\right)\overline{X_k} \tag{2}$$

3. Find left between class scatter matrix S_{bl} and left within class scatter matrix S_{wl}.

$$S_{bl} = \sum_{k=1}^{N} n_k\,(\overline{X_k} - M)^T(\overline{X_k} - M) \tag{3}$$

$$S_{wl} = \sum_{k=1}^{N}\sum_{i=1}^{n_k}\left(X_i^k - \overline{X_k}\right)^T\left(X_i^k - \overline{X_k}\right) \tag{4}$$

4. Compute the first m_l eigenvectors of $S_{wl}^{-1}S_{bl}$ to get the left transformation matrix W_l.
5. Find right between class scatter matrix S_{br} and right within class scatter matrix S_{wr}.

$$S_{br} = \sum_{k=1}^{N} n_k\,(\overline{X_k} - M)(\overline{X_k} - M)^T \tag{5}$$

$$S_{wr} = \sum_{k=1}^{N}\sum_{i=1}^{n_k}\left(X_i^k - \overline{X_k}\right)\left(X_i^k - \overline{X_k}\right)^T \tag{6}$$

6. Compute the first m_r eigenvectors of $S_{wr}^{-1}S_{br}$ to get the right transformation matrix W_r.
7. Compute the left reduced dimensional representations LB_i^k and right reduced dimensional representations RB_i^k.

$$LB_i^k = X_i^k \times W_l \tag{7}$$

$$RB_i^k = (X_i^k)^T \times W_r \tag{8}$$

3.2. Data Selection Using FNN Criterion

The data selection is the most important step in the application of active learning principle to age classification. It is used to select an unlabeled sample that will be the most informative sample to improve the classifier. The input to this step is a pool of labelled and unlabeled images. The output is the selected image. The furthest nearest neighbour technique proposed in [1] performs the following steps:

1. Find out the distances between labelled and unlabeled samples.
2. Find out the nearest labelled neighbours of each unlabeled sample.
3. Find out the furthest unlabeled sample among those selected in step 2.
4. Give the selected sample to the trainer for labelling.
5. Add the labelled sample to the set of training images.

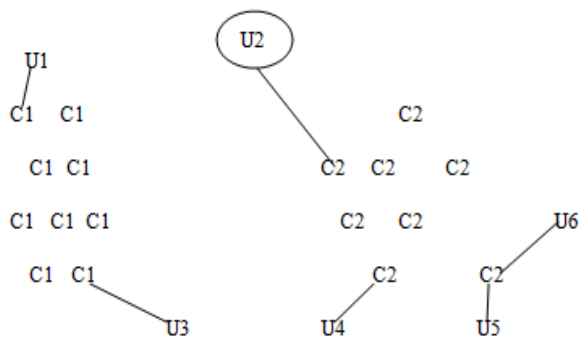

Figure 2. Data Selection using FNN

The data selection process using furthest nearest neighbour criterion is explained in Figure 2. The labelled training samples belonging to two classes C1 and C2 are shown in the figure. U1 to U6 are the unlabeled instances. Lines are drawn from each unlabeled instance to its corresponding nearest labelled instance. The instance U2 will be selected by the FNN criterion in this case because it is the unlabeled instance which is the furthest from its nearest labelled neighbour.

3.3. Incremental Learning Using IB2DLDA

The application of active learning technique to age classification makes the learning process an iterative one. The data selection criterion FNN selects the most informative unlabeled image, which is labelled by the oracle and added to the training set. Instead of a single image, a batch of images can be selected from the unlabeled pool, labelled and added to the training set. The scatter matrices calculated in the previous iteration have to be updated now with the new training set. If B2DLDA is used for this, a large amount of update calculations will be needed. To avoid calculation overhead, the updates can be performed incrementally using an incremental version of B2DLDA, called incremental bilateral two dimension linear discriminant analysis (IB2DLDA) as explained in [1].

Let $Y_1, Y_2, .., Y_t$ be *t* new labeled images that were added to the training set. Let $l_1, l_2, .., l_t$ be their respective labels. Let q_m images of these *t* images $\{Y_p\}_{Y_p \epsilon Y \cap l_p = m}$ belong to *m*th class. Let $\bar{y}_k$ be the mean of the new samples that belong to the *k*th class and let $\Sigma_{kl}$and Σ_{kr} be the left and right scatter matrices of the original training samples in the *k*th class. The calculations are performed as follows:

1. The number of samples in *m*th class is updated as:

$$n'_m = n_m + q_m \tag{9}$$

2. The mean of the *m*th class is updated as:

$$\bar{X}'_m = \frac{\left(n_m \bar{X}_m + \sum_{Y_p \epsilon Y \cap l_p = m} Y_p\right)}{n'_m} \tag{10}$$

3. The global mean is updated as:

$$M' = \frac{nM + \sum_{i=1}^{t} Y_i}{n+t} \tag{11}$$

4. The left between class scatter matrix is updated as:

$$S'_{bl} = \sum_{k=1}^{N} n'_k (\bar{X}'_k - M')^T (\bar{X}'_k - M') \tag{12}$$

5. The right between class scatter matrix is updated as:

$$S'_{br} = \sum_{k=1}^{N} n'_k (\bar{X}'_k - M')(\bar{X}'_k - M')^T \tag{13}$$

6. The left within class scatter matrix is updated as:

$$S'_{wl} = \sum_{k=1}^{N} \left(\Sigma_{kl} + \frac{n_k q_k^2}{(n_k + q_k)^2}\right)(\bar{y}_k - \bar{X}_k)^T (\bar{y}_k - \bar{X}_k) + \frac{n_k^2}{(n_k + q_k)^2} \times \sum_{Y_p \epsilon Y \cap l_p = k} (Y_p - \bar{X}_k)^T (Y_p - \bar{X}_k) + \frac{q_k(q_k + 2n_k)}{(n_k + q_k)^2} \times \sum_{Y_p \epsilon Y \cap l_p = k} (Y_p - \bar{y}_k)^T (Y_p - \bar{y}_k) \tag{14}$$

7. The right within class scatter matrix is updated as:

$$S'_{wr} = \sum_{k=1}^{N} \left(\Sigma_{kr} + \frac{n_k q_k^2}{(n_k + q_k)^2}(\bar{y}_k - \bar{X}_k)(\bar{y}_k - \bar{X}_k)^T + \frac{n_k^2}{(n_k + q_k)^2} \times \sum_{Y_p \epsilon Y \cap l_p = k} (Y_p - \bar{X}_k)(Y_p - \bar{X}_k)^T + \frac{q_k(q_k + 2n_k)}{(n_k + q_k)^2} \times \sum_{Y_p \epsilon Y \cap l_p = k} (Y_p - \bar{y}_k)(Y_p - \bar{y}_k)^T\right) \tag{15}$$

3.4. Age Classification Using k-Nearest Neighbor Approach

After the training phase, the classifier has to be evaluated by providing test images and verifying the correctness of the class label assigned by the classifier [1]. Age classification is performed using *k*-nearest neighbor method. The *k*-nearest neighbor algorithm (*k*-NN) is a method for classifying objects based on closest training examples in the feature space. A test image is classified by a majority vote of its neighbors, with the image being assigned to the class most common amongst its *k* nearest neighbors. If *k* = 1, then the image is simply assigned to the class of its nearest neighbor.

The training phase of the algorithm consists only of storing the feature vectors and class labels of the training samples. In the classification phase, *k* is a user-defined constant, and an unlabeled vector (a query or test point) is classified by assigning the label which is most frequent among the *k* training samples nearest to that query point.

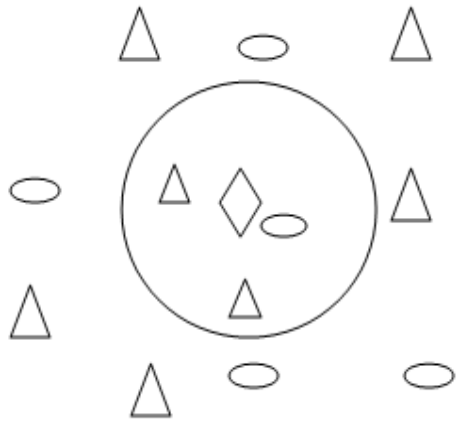

Figure 3. *k*-NN classification

The classification method using *k*-NN is shown in Figure 3. The test instance to be classified is represented using the diamond symbol. Labeled training images belonging to two classes are shown using triangles and ovals. The task is to classify the test instance represented using diamond to one of the two classes shown by ovals and triangles. If k=1, the test instance will be classified to the class represented by oval because it is the nearest neighbor. If k=3, the oval and the two triangles shown within the circle in the figure are the nearest neighbors. The test instance will be classified to the class represented by triangle because two out of the three nearest neighbors belong to the triangle class.

3.5. Active learning in Support Vector Machine

It has been stated in [1] that for very large databases with high feature dimensions and with large number of classes, the active SVM approach will be intractable. In order to make SVM handle large feature dimensions, a technique is proposed to combine B2DLDA with SVM. B2DLDA can be used to reduce images to a lower dimensional space, the most informative unlabeled image can be found out from the unlabeled pool using FNN, incremental learning can be done using IB2DLDA and finally classification can be done by SVM. A one versus all SVM can be used for multiclass classification since ordinary SVM is a binary classifier.

Support Vector Machine is a binary classification method that finds the optimal linear decision surface between two classes. The decision surface is nothing but a weighted combination of the support vectors. The standard SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the input, making the SVM a non-probabilistic binary linear classifier. Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other. An SVM model is a representation of the examples as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible. New examples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on.

The SVM takes in labelled training examples (u_i , v_i), where u_i represents the features of the training samples and v_i represents the corresponding class labels, that could be either 1 or -1. By the training phase, a set of *m* support vectors s_i multipliers α_i, y_i and the term *b* are obtained.

The hyperplane is represented as:

$$w * x + b = 0 \tag{16}$$

where $w = \sum_{i=1}^{m} \alpha_i \, y_i \, s_i$, the terms *w* and *x* determine the orientation of the hyperplane and *b* represents the actual position of the hyperplane.

The classification by SVM is shown in Figure 4. The hyperplane has a margin that allows incorporating some tolerance to the exact partitioning of feature space into two classes. Two classes are labelled 1 and -1. Images fall in class 1 if $w * x + b > 1$ and images fall in class 2 if $w * x + b < -1$.

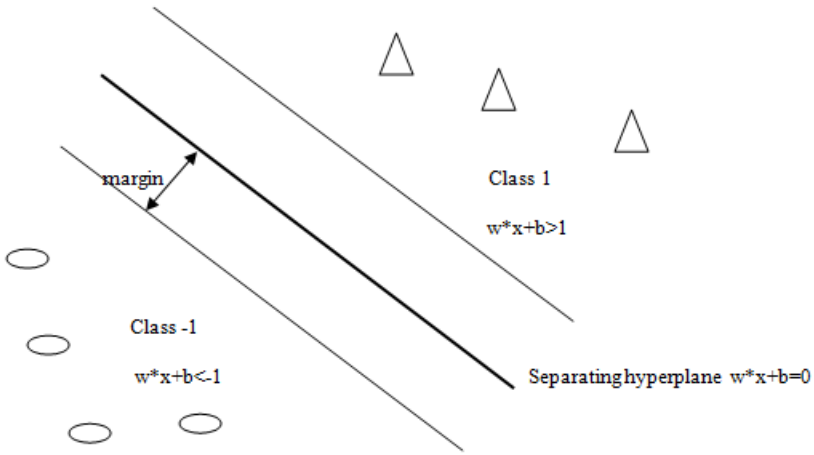

Figure 4. SVM classification

4. EXPERIMENTATION

The study experiment is currently ongoing using the FG-NET Aging Database and it has been completed till the data selection step. It is an image database containing face images showing a number of subjects at different ages. The images have different illumination, pose, expression, and includes faces with beards, moustaches, spectacles, and hats.

The input is a pool of labeled facial images and a pool of unlabeled facial images as shown in Figure 5 to Figure 9. The labeled images belong to four predefined classes- children, teenager, adult, and senior adult.

Figure 5. Training samples in children class.

Figure 6. Training samples in teenager class.

Figure 7. Training samples in adult class.

Figure 8. Training samples in senior adult class.

Figure 9. Unlabeled samples

The classifier is initially trained with the labeled training images using the bilateral two dimension linear discriminant analysis (B2DLDA). The B2DLDA converts the training images to reduced dimensional representations. The main goal of B2DLDA is to maximize the ratio of the between class distance to the within class distance so that a good separation between the given classes can be achieved. Then the unlabeled images are also reduced using B2DLDA.

The FNN criterion calculates the distances between the labeled and the unlabeled images and finds out the unlabeled image which is the furthest from its nearest labeled neighbor. The image selected in such a way is considered to be the most informative sample to improve the classifier. The image selected from the unlabeled pool using the FNN technique is shown in Figure 10. A menu is displayed to the trainer and the selected image is labeled by the trainer by providing the proper menu option as shown in Figure 11. The labeled image is then added to the respective class in the training set.

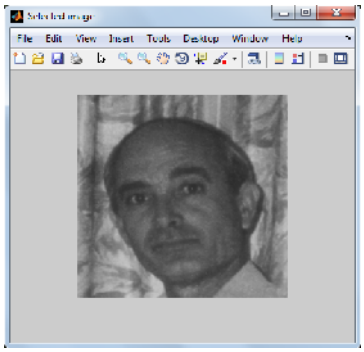

Figure 10. The image selected by the FNN criterion from the unlabeled pool.

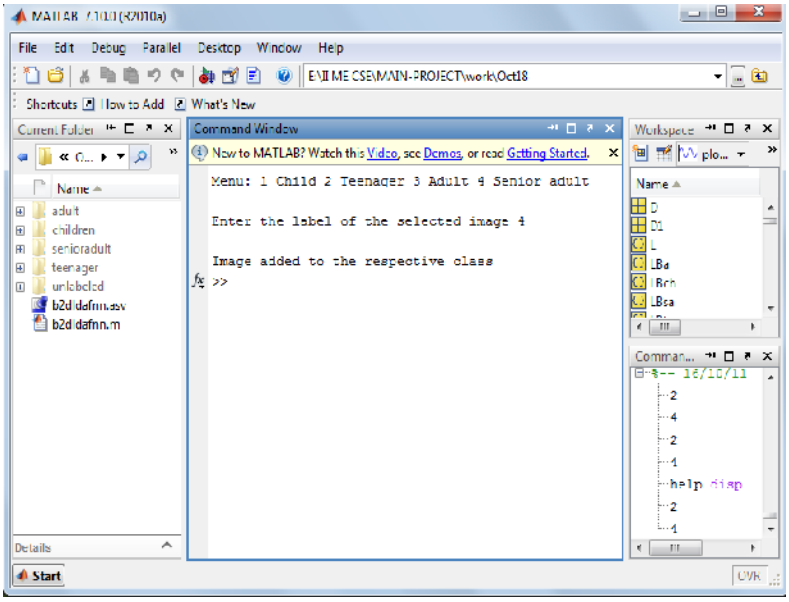

Figure 11. The labeling of the selected image by the trainer.

The calculation performed by B2DLDA has to be updated and classification accuracy has to be checked in both k-NN and SVM classifiers.

6. CONCLUSION

The study of active learning technique using furthest nearest neighbour criterion for facial age classification is being carried out. The bilateral two dimension linear discriminant analysis has been used to convert images to reduced representations. The unlabeled image to be labeled next was selected using the furthest nearest neighbor criterion. It was labeled by the user and added to the training set. The classifier needs to be updated using the samples added during each iteration of active learning. The accuracy of the classifier needs to be evaluated. The application of active learning to k-NN and SVM is expected to produce better results than the ordinary k-NN and SVM classifiers.

REFERENCES

[1] Jian Gang Wang, Eric Sung, and Wei Yun Yau, "Active Learning for Solving the Incomplete Data Problem in Facial Age Classification by the Furthest Nearest-Neighbor Criterion," in *IEEE Trans. Image Process.*, vol. 20, no. 7, pp. 2049-2062, Jul. 2011.

[2] B. Settles, "Active learning literature survey," Univ. of Wisconsin-Madison, Madison, WI, Computer Sciences Tech. Rep.1648, 2009.

[3] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance model," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.

[4] S. Tong and D. Koller, "Support vector machine active learning with applications to text classification," *J. Mach. Learning Res.*, pp. 45–66, 2001.

[5] X. Li, L. Wang, and E. Sung, "Multi-label SVM active learning for image classification," in *Proc. IEEE Int. Conf. Image Process.*, 2004, pp. 2207–2210.

[6] P. Jain and A. Kapoor, "Active learning for large multi-class problems," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 1–8.

[7] Y. H. Kwon and N. D. V. Lobo, "Age classification from facial images," *Comput. Vis. Image Understanding*, vol. 74, no. 1, pp. 1–21, 1999.

[8] X. Geng, Z. H. Zhou, Y. Zhang, G. Li, and H. Dai, "Learning from facial aging patterns for automatic age estimation," in *Proc. ACMConf. Multimedia*, 2006, pp. 307–316.

[9] A. Lanitis, C. Draganova, and C. Christodoulou, "Comparing different classifiers for automatic age estimation," *IEEE Trans. Syst., Man, Cybernet.*, vol. 34, no. 1, pp. 621–628, Jan. 2004.

[10] N. Ramanathan and R. Chellappa, "Modeling age progression in young faces," in *Proc. CVPR*, 2006, pp. 387–394.

[11] J. G. Wang, W. Y. Yau, and H. L. Wang, "Age categorization via ECOC with fused gabor and LBP features," in *Proc. IEEE Workshop Applic. Comput. Vis.*, 2009, pp. 1–6

[12] Y. Fu and T. S. Huang, "Human age estimation with regression on discriminative aging manifold," *IEEE Trans. Multimedia*, vol. 10, no. 4, pp. 578–584, Jun. 2008.

Authors

Indu. M. T. was born in Kerala, India, in 1988. She received her Bachelor of Technology degree in Computer Science and Engineering from Calicut University, Kerala, India in 2010. She is currently pursuing M. E in Computer Science and Engineering under Anna University, Coimbatore, India.

Prof. S. Bhuvana received her Bachelor of Engineering degree in Computer Science and Engineering from Sri Ramakrishna Engineering College, India, in 1999 and her Master of Engineering degree in Computer Science and Engineering from Kumaraguru College of Technology, India in 2007. At Present, she is working as Assistant Professor in the department of Computer Science & Engg., Sri Krishna College of Technology, Coimbatore. Currently she is pursuing her research work in Image Retrieval. Her research interests include Image Processing and Machine Learning.

Prof. P.Tamije Selvy received B.Tech (CSE), M.Tech (CSE) in 1996 and 1998 respectively. Since 1999, she has been working as faculty in reputed Engineering Colleges. At Present, she is working as Assistant Professor(SG) in the department of Computer Science & Engg , Sri Krishna College of Technology, Coimbatore. She is currently pursuing Ph.D under Anna University, Chennai. Her Research interests include Image Processing, Data Mining, Pattern Recognition and Artificial Intelligence.

Dr. R. Radhakrishnan is the Principal of Sri Ramakrishna Engineering College. He obtained his Bachelor's degree in Engineering in Electronics and Communication Engineering from Bharathidasan University, received his Masters Degree from P.S.G.College of Technology, Coimbatore, in the year 1997 and the Ph.D from Anna University Chennai in the year 2008. His research interest includes Wireless Communication, Signal Processing, Networking and Mobile Communication. He has published more than 11 Journal papers in the field of CDMA systems, Mobile communication, Wireless Networking and Signal Processing.