Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 3159-0040
Vol. 2 Issue 12, December - 2015

Cluster Based Ear Biometric System for Personal Identification:

An Exploitative Approach

M.A.Jayaram
Director, Department of MCA
Siddaganga Institute of Technology
Tumkuru, Karnataka, India
jayaram_mca@sit.ac.in

Prashanth G.K.,
Assistant Professor, Department of MCA
Siddaganga Institute of Technology
Tumkuru, Karnataka, India
prashanthgk@sit.ac.in

***Abstract*—It is becoming increasingly clear among biometric research fraternity that Ear as a biometric articulation in human beings provides exclusive and unique advantages when compared with other kinds. In this paper, we present a person identification system which is based on clustering of ears. For the development of the system, a database of 605 ear images was considered. Shape based biometric features that are related to moment of inertia (MI) of the ear such as MI about major axis, MI about minor axis, radii of gyration with reference to major and minor axis and the area of the ear were elicited presuming that the ear as a planar surface of irregular shape. Extensive computational clustering experiments were done involving k-means, back propagation network (BPN), four similarity metrics (Jaccard, Dice, Cosine, and Overlapping), radial basis function network (RBFN), and support vector machine (SVM). The results showed that Jaccard, Dice similarity measures and RBNF showed excellent performance in terms of minimal overlapping among groups followed by recognition accuracy ranging from 94%-98%. Further, there was a decrease in CPU time for image matching and retrieval of personal details to an extent of 14 % when compared with the time expended for the same task when the image database was presented in an unorganized way.**

Keywords— Ear biometrics, shape related parameters, moment of inertia, radius of gyration, major axis, minor axis, similarity measures.

I. INTRODUCTION

In recent years, recognizing people by their ear has gained momentum and copious literature is also available. The reasons that merit consideration to account this trend are: firstly, ear recognition is typically a noncontact biometrics, such as face recognition; secondly, it is the most widely used in combination with the face in the context of multi-pose face recognition; and third, the ear can be used for human recognition in surveillance videos where the face may be occluded completely or in part. Further, the ear appears to degrade little with age. Above all, the ear can easily be captured from a distance, even if the subject is not fully cooperative. This makes ear recognition especially interesting for smart surveillance tasks and for forensic image analysis [1].

Even though current ear detection and recognition systems have reached a certain level of maturity, their success is limited to controlled indoor conditions. In addition to variation in illumination, other open research problems include hair occlusion, ear print forensics, ear symmetry, ear classification, and ear individuality [2]. Most of the researchers in general and criminal investigators in particular have found the shape of the outer ear as a valuable means for personal identification. The French criminologist Alphonse Bertillon was the first to herald the potential use for human identification through ears, more than a century ago [3]. It is well documented in the literature on ear biometrics that the American police Officer Alfred Iannarelli used a huge database of 10000 ear images and conclusively proved the discriminative potential of outer edge of the ear. Further, his experimentation had an outcome of 12 characteristic features needed to unambiguously identify a person [4]. Iannarelli also performed studies on twins and triplets, discovering that ears are even unique among genetically identical persons. Though there is no complex theoretical basis for Iannarelli's work, it is commonly believed that the shape of the outer ear is unique. The studies as reported in [5] and [6] show that all ears of the investigated databases posses individual characteristics, which can be used for distinguishing. Because of the lack of a sufficiently large ear database, these studies can be regarded as hints, not evidence, for the outer ear's uniqueness.

The work presented in this paper is thoroughly different and novel for the reason that the biometric features are extracted under the premise that the ear can be treated as a planar surface of some peculiar shape. Thus, shape related features of planar surface like Moment of Inertia (MI) about two mutually orthogonal axes, the respective radii of gyration and the planar area are extracted from 605 ear images. The sole motive of this work is to identify the persons by performing classification of ear images in the database by using a variety of clustering methods and designing a person identification system by choosing the best method that shows minimal overlapping among groups. The rest of the paper is organized as follows. Section II presents the related research work , the data for the work and the feature extraction process is briefed in section III, section IV illustrates briefly the various classification techniques used , a comparative analysis is of various classification techniques is made in section V , section VI presents the efficiency of the recognition system ,and the paper concludes in section VII.

II Related Research Work

Considerable research on person identification through classification of biometric data has been reported. However, all these works pertain application of individual classification methods. There is a paucity for an explorative research involving a galaxy of classification methods. This is exactly the motivation for this paper. A new method based on a combination of supervised and unsupervised learning for clustering data without any preliminary assumption on the cluster shape is implemented for Iris dataset. This is obtained by extracting the dissimilarity relations directly from the available data [7]. A novel approach directed towards the automatic clustering of x-ray images has been attempted. The clustering was carried out based on multi-level feature of given x-ray images such as global level, local level and pixel level. The approach involves a combination of k-means and hierarchical clustering techniques this work has reported for having shown high level of accuracy [8]. Xi Cheng et al [9] have used similarity measures in multi-sample biometric systems. Both Pearson's correlation and Cosine similarity are used. Computational experiments have shown a better performance than using raw matching scores. Roman V. et al [10] have compared performance of similarity measure functions to that obtained from customized field-specific approach in the domain of strategy-based behavioral biometrics. While all similarity measure functions showed a relatively high accuracy levels during user verification, weighted Euclidian similarity measures has slightly outperformed than general approaches such as Manhattan distance or Mahalanobis distance as claimed. Satya Chaitanya Sripada et al [11] have compared the for K-means and Fuzzy C means clustering using the Purity and Entropy. The paper reported that, The K-means has lower value of purity and high value of entropy compared to Fuzzy C Means. The Fuzzy C means clustering is more accommodating for medical data sets when compared to K means. Vikas Thada et al [12] have focused on comparative analysis for finding out the most relevant document for the given set of keywords by using three similarity measures viz Jaccard, Dice and Cosine similarity measures by using genetic algorithm approach. Due to the randomized nature of genetic algorithm the best fitness value is the average of 10 runs of the same code for a fixed number of iterations. The result states that the best fitness values were obtained using the Cosine similarity coefficients followed by Dice and Jaccard.

Mai V et al [13] proposed a new method to identify people using Electrocardiogram (ECG). QRS complex (Q waves, R waves, S waves) which is a stable parameter against heart rate variability is used as a biometric feature. This work has reported for having achieved a classification accuracy of 97% using RBF. Sulong et al [14] have used a combination of maximum pressure exerted on the keyboard and the time latency between the keystrokes to recognize the authenticate users and to reject imposters. In this work, RBFNN is used as a pattern matching method. The system so developed has been evaluated using False Reject Rate (FRR) and False Accept Rate (FAR). The researchers have affirmed the effectiveness of the security system designed by them. Chatterjee et al [15] have proposed a new biometric system which is based on four types of temporal postural signals. The system employs S-transform to determine the characteristic features for each human posture. An RBFNN with these characteristic features as input is developed for specific authentication. The training of the network has augmented extended Kalman filtering (EKF). The overall authentication accuracy of the system is reported to be of the order of 95%.

Multi-modal biometric consisting of fingerprint images and finger vein patterns were used to identify the authorized users after determining the class of users by RBFNN as a classifier. The parameters of the RBFNN were optimized using BAT algorithm. The performance of RBFNN was found to be superior when compared with KNN, Naïve Bayesian and non-optimized RBFNN classifier [16]. Ankit Chadha et al have used signature of persons for verification and authentication purpose. RBFNN was trained with sample images in the database. The network successfully identified the original images with the recognition accuracy of 80% for image sample size of 200 [17]. Handwriting recognition with features such as aspect ratio, end points, junction, loop, and stroke direction were used for recognition of writers [18]. The system used over 500 text lines from 20 writers. RBFNN showed a recognition accuracy of 95.5% when compared to back propagation network.

Justino et al[19] have compared SVM and HMM classifiers under two specific conditions, the first being the number of samples used for training, and the second being the use of different types of forgeries. Under both conditions, the SVM showed better results. However, in terms of random forgery acceptance and small number of samples used to training, the SVM showed promising results, demonstrating SVM_s ability to identify simple and simulated forgeries without previous knowledge.

Ramirez et al[20] have developed Handshake biometric system based on feature extraction methodology which is novel. In this work, Identification experiments were carried out using the feature vectors as inputs to recognition system using SVM technique.

An average recognition system of 98.5 is claimed. The verification included False Acceptance Rate and False Rejection Rate.

Tobias et al [21] have developed a Biometric User Authentication System based on dynamic hand writing classification of persons. In this work Support Vector Machines are employed to classify dynamic hand writing sample. The goal of SVM in this work is to carry out binary classification and to handle multiple class problems using a combination of different Support Vector Machines

Scheirer et al [22] have reported Face Recognition algorithm which is based on similarity surfaces and Support Vector Machines. Their work has shown that prediction of biometric system failure can be done reliabily using SVM approach

III. Data Gathering and Feature Extraction

In this work, a novel idea that makes use of planner surface area properties has been used. For this, a ear is considered to be a planner surface. The moment of inertia (MI) and its related five parameters are elicited from the ear images. The features considered are given in Table 1. The details of the features, their extraction, the evaluation authentication and the development of the system for human identification making use of these features is elaborated in the seminal work of authors [23]. However, for the sake of completeness the features are briefly explained in the following paragraphs.

The surface area of the ear is the projected area of the curved surface on a vertical plane. Moment of Inertia (MI) is the property of a planar surface which originates whenever one has to compute the moment of distributed load that varies linearly from the moment axis. Moment of Inertia is also viewed as a physical measure that signifies the shape of a planar surface and it is proved that by configuring the shape of planar surface and hence by altering the moment of inertia, the resistance of the planar surface against rotation with respect to a particular axis could be modulated or altered [24]. Therefore in this work, moment of inertia of ear surface with respect to two axes i.e. the major axis and the minor axis are considered to be the best biometric attributes that could capture the shape of irregular surface of the ear in a more scientific way.

Further, major axis is the one which has the longest distance between the two points on the edge of the ear, the distance here is the maximum among point to point Euclidean distance. The minor axis is drawn in such way that it passes through tragus and is orthogonal to the major axis. Therefore, with different orientation of ears the orientation of major axis also changes. Being perpendicular to major axis, the orientation of minor axis is fixed.

The projected area is assumed to be formed out of segments. The area of an ear to the right side of the major axis is considered to be made out of six segments. Each of the segments thus subtends 300 with respect to the point of the intersection of the major axis and minor axis. The extreme edge of a sector is assumed to be a circular arc. Typical ear edge with measurements is shown in Figure 1.

The measurements are

θ → Inclination of the central radial axis of the segment with respect to minor axis (in degrees).

r → The length of the radial axis (in mm).

The conversion of number of pixel into linear dimension (in mm) was based on the resolution of the camera expressed in PPI (Pixel Per Inch). In this work 16Mega pixel camera, at 300 PPI was used. The computation of linear distance is straight farward mm=(number of pixel*25.4)/PPI [1 inch=25.4 mm]. With these measurements, the following parameters are computed.

Moment of inertia with respect to minor axis Imin

$$I_{min} = \sum_{i=1}^{6} a_i y_i^2 \qquad (1)$$

Where ai is the area of a the ith segment and yi is the perpendicular distance of the centroid of the ith segment with respect to minor axis.

$$a_i = \theta r^2 \qquad (2)$$

$$y_i = C \sin\theta \qquad (3)$$

Here, C is the centroidal distance of the segment with respect to the intersection point of the axes, which is given by[25];

$$C = \frac{2}{3}\frac{r \sin\theta}{\theta} \qquad (4)$$

Simlarly, moment of inertia with respect to major axis Imax ,xi is the perpendicular distance of the centroid of the ith segment with respect to major axis.

$$I_{max} = \sum_{i=1}^{6} a_i x_i^2 \qquad (5)$$

$$\text{Where } x_i = C\cos\theta \qquad (6)$$

From the computed values of moment of inertia and area of the ear surface, the radii of gyration with respect to minor axis(RGx)and major axis(RGy) were computed. The formulae for radii of gyration are given by[25].

$$RGx = \sqrt{\frac{I_{min}}{A}} \qquad (7)$$

$$RGy = \sqrt{\frac{I_{max}}{A}} \qquad (8)$$

Where , A is the sum of areas of six segments.

$$A = \sum_{i=1}^{6} a_i \qquad (9)$$

Radius of gyration is the distance from an axis at which the mass of a body may be assumed to be concentrated and at which the moment of inertia will be equal to the moment of inertia of the actual mass about the axis. It is also equal to the square root of the quotient of the moment of inertia and the mass. Mathematically, it is equal to the square root of the quotient of the moment of inertia and the area. The axis of inertia is unique to the shape. It serves as a unique reference line to preserve the orientation of the shape. The axis of least inertia (ALI) of a shape is defined as the line for which the integral of the square of the distances to points on the shape boundary is a minimum [26].

Table 1: Ear Shape Based Features in Classification

Sl. No	Attributes
1	Area (mm^2)
2	Moment of Inertia Y (Imax) (mm^4)
3	Radius of gyration Y (RGy) (mm)
4	Moment of Inertia X (Imin) (mm^4)
5	Radius of gyration X (RGx) (mm)

Ear images for this work were acquired from the pupils of Siddaganga group of institutes. The subjects involved were mostly students and faculty numbering 605. In each acquisition session, the subject sat approximately one meter away with the side of the face in front of the camera in outside environment without flash.

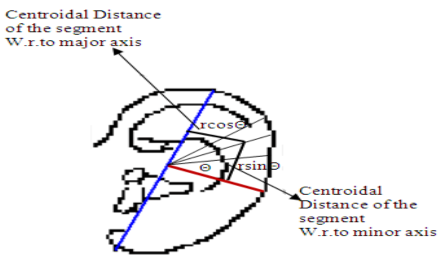

Fig. 1. Typical ear edge with M.I. parameters

The images so obtained were resized in such a way that only ear portion covers the entire frame having pixel matrix. A segment of the database showing right ear images of persons is shown in Figure 2.

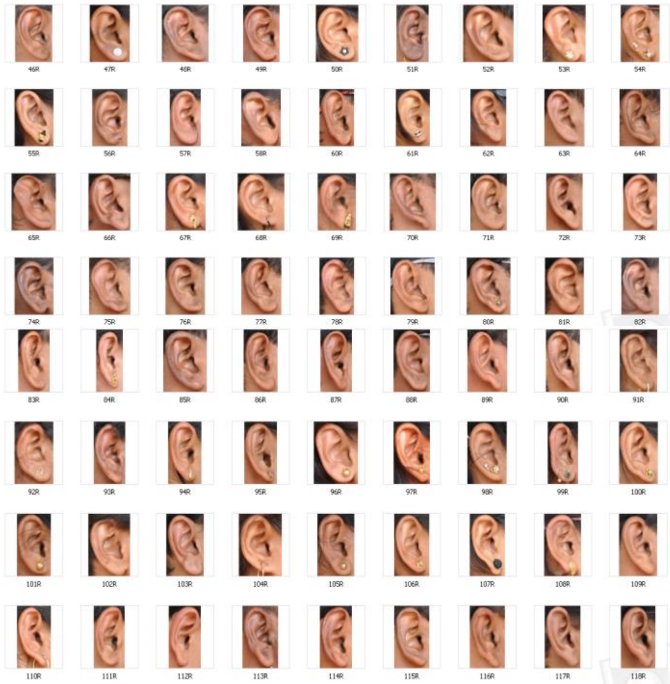

Fig.2. A sample gallery of right ear images.

The color images were converted into gray scale images followed by uniform distribution of brightness through histogram equalization technique. The delineation of outer edge of each ear was obtained using canny edge detection algorithm. The resulting edge was inverted to get a clear boundary shape of the ear. The conceptual presentation of the process involved is shown in Figure 3.

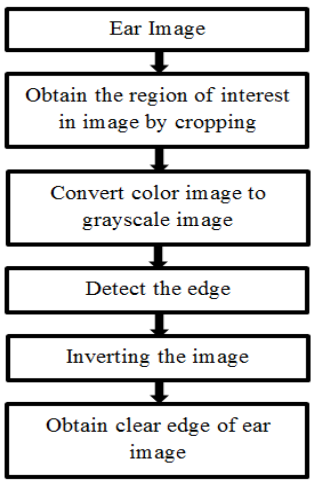

Fig 3. The Steps involved in ear edge extraction

A sample segment of the feature database is shown in Table 2. The descriptive statistics of the entire data of the right ear features is presented in Table 3, this table oestensibly gives summary of feature data in a clear and understandble way. The centeral tendeancy of the feture values is depicted by mean and median. The range is indicative of the dispersion of the feature values. Skewness showcases the devation of the distribuation of the feature values from symmetric, and it is clearly seen that distribution is asymetrical. Kurtosis, which is an indicator of peakedness of the distribution is clearly different from zero conveying that the distribution is peaked than normal. Standard devation and variance indicates the scatterness of attribute values with respect to their average. There is a significant variability of feature values among the ears as the value of standard devation and variance is appeciably high.

Table 2: A sample segment of the database(Right Ear)

A mm^2	**I_{min} mm^4**	**I_{max} mm^4**	**RGx mm**	**RGy mm**	**Major Axis mm**	**Minor Axis mm**
1054.09	13548.79	33529.17	2.46	5.68	50.62	29.47
1477.39	18653.79	35448.46	1.85	4.76	55.45	19.90
1815.42	20458.75	44501.02	4.37	5.70	65.74	29.27
1641.17	21354.79	38746.51	2.15	4.86	53.06	26.00
1759.08	18546.79	40263.80	3.24	5.61	64.47	30.54
1947.47	14568.75	45021.31	1.26	4.16	66.59	26.05
1856.08	22903.51	44550.11	1.08	2.15	65.25	25.52
1630.32	61208.56	38520.81	2.07	5.27	61.17	22.96
1909.18	28157.12	44980.40	1.81	5.80	65.86	24.96
1212.87	16271.48	35089.56	1.61	8.73	49.63	16.22
1054.09	13637.79	33619.17	2.66	5.86	52.22	29.52
1477.38	18766.75	35529.46	1.41	4.75	55.84	19.93
1815.42	20314.31	44056.13	4.24	5.51	65.17	29.33
1641.17	21210.35	38301.63	2.02	4.67	54.65	25.03
1759.08	18402.35	39818.91	3.11	5.42	64.30	29.50
1947.47	14424.31	44576.42	1.14	3.97	66.80	26.12
1856.08	22759.07	44105.22	0.95	1.96	65.77	24.90
1630.32	61208.56	38520.81	2.07	5.27	61.10	23.05
1909.18	28157.12	44980.40	1.81	5.80	65.05	24.96
1212.87	16271.48	35089.56	1.61	8.73	18.93	17.21

Table 3. Descriptive statistics of ear features for the entire database.

	A(m^2)	I_{max} mm^4	I_{min} mm^4	RGx mm	RGy mm
Mean	1920.519	63249.8548	37828.8147	7.189329	3.53472
Median	1737.422	42775.6391	23230.3368	6.618619	3.205432
Standard Deviation	694.3276	91346.1497	57085.9967	2.520325	1.693496
Variance	482090.8	8344119073	3258811020	6.35204	2.867927
Kurtosis	0.10165	42.0363847	10.8897931	-0.26037	2.277937
Skewness	0.830819	6.19088895	3.47763469	0.588988	1.151528
Min	553.5184	2884.27825	1593.23797	1.862836	0.8134
Max	4097.718	802410.363	310819.142	14.84563	11.9742
Range	3544.2	799526.085	309225.904	12.9828	11.1608

IV. Classification Techniques Used

In all, five clustering methods were used for exploring a way to group the ears with minimum overlapping. The methods include KNN, Back Propagation Neural Network (BPN), Radial Basis Function Network (RBFN), Support Vector Machine(SVM), similarity functions that consisted of Cosine, Jaccard, Dice and overlapping measure. For the sake of completeness all the methods considered are briefed.

A. K-Means Algorithm

K-means clustering is an iterative, data-partitioning algorithm that assigns n observations to exactly one of k clusters defined by centroids, where k is chosen before the algorithm starts. The purpose of applying the k-means clustering algorithm is to find a set of clustered centres and a partition of training data into subclasses. Normally, the centre of each cluster is initialized to a randomly chosen input datum. Then each training datum is assigned to the cluster that is nearest to itself. After training data have been assigned to a new cluster unit, the new centre of a cluster represents the average of the training data associated with that cluster unit. After all the new clusters have been calculated, the process is repeated until it converges [27].

Computational Experiments were carried out to find the appropriate number of clusters using K-Means. The algorithm settled for three distinct clusters with a minimum overlapping. In fact, K-means algorithm was also used for the purpose of ascertaining possible number of groups with minimum overlapping. During runs of the algorithm, number of groups less than three and more than three proved to be untenable with huge overlapping.

B. Back Propagation Neural Network

In the present study we implement back propagation Neural Network algorithm for classifying the image dataset. Figure 3 represent architecture of a simple BPN used in this work. It has 3 layers i.e., input layer consisting of 5 neurons, central layer consisting of just 3 neurons and a lone neuron in output layer. Inputs data is fed forward through the network to optimize the weights between neurons. Adjustment of the weights is done by backward propagation of the error during training phase. The network takes the input and target values in the training data set and changes the value of the weighted links to reduce the difference between the output and target values. The error is minimized across many training cycles called epoch. During each cycle network reaches to specified level of accuracy. The number of processing elements per layer, as well as the number of layers, greatly affects the abilities of the network. Sigmoid activation function is used. The training algorithm used in this work is Levenberg–Marquardt back propagation (trainlm). This algorithm locates the minimum of a multivariate function that can be expressed as the sum of squares of non-linear real-valued functions. It is an iterative technique that works in such a way that performance function will always be reduced in each iteration of the algorithm. This feature makes trainlm the fastest training algorithm for networks of moderate size. Similar to

Broyden–Fletcher–Goldfarb–Shanno(trainbfg), trainlm function has drawback of memory and computation overhead caused due to the calculation of the gradient and approximated Hessian matrix [28].

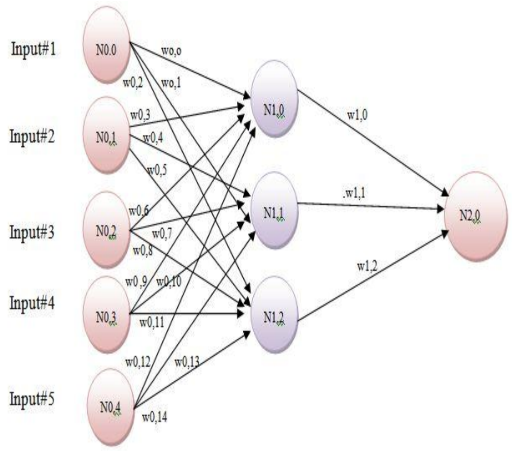

Fig. 3: Architecture of BPN

C. Similarity Measures

Before clustering, a similarity/distance measure must be determined. The measure reflects the degree of closeness or separation of the target objects and should correspond to the characteristics that are believed to distinguish the clusters embedded in the data. In many cases, these characteristics are dependent on the data or the problem context at hand, and there is no measure that is universally best for all kinds of clustering problems [29]. Moreover, choosing an appropriate similarity measure is also crucial for cluster analysis, especially for a particular type of clustering algorithms. In general, similarity/distance measures map the distance or similarity between the symbolic descriptions of two objects into a single numeric value, which depends on two factors— the properties of the two objects and the measure itself. Deferent measure not only results in deferent partitions, but also imposes deferent requirements for the same clustering algorithm [30]. The four similarity measures used in this work are briefed.

Cosine similarity: Cosine similarity is a measure of similarity between two vectors of an inner product space that measures the Cosine of the angle between them [31]. The Cosine of 0° is 1, and it is less than 1 for any other angle. It is thus a judgment of orientation and not magnitude: two vectors with the same orientation have a Cosine similarity of 1, two vectors at 90° have a similarity of 0, and two vectors diametrically opposed have a similarity of -1, independent of their magnitude. Cosine similarity is particularly used in positive space, where the outcome is neatly bounded in [0,1].

$$f(x,y)=\frac{\sum_{i=1}^{p} x_i y_i}{\sqrt{\sum_{i=1}^{p} x_i^2 \sum_{i=1}^{p} y_i^2}} \quad (10)$$

One of the reasons for the popularity of Cosine similarity is that it is very efficient to evaluate, especially for sparse vectors, as only the non-zero dimensions need to be considered.

Jaccard similarity: The Jaccard similarity coefficient is a statistic used for comparing the similarity and diversity of sample sets [32]. The Jaccard coefficient measures similarity between finite sample sets, and is defined as the size of the intersection divided by the size of the union of the sample sets.

$$f(x,y)=\frac{\sum_{i=1}^{p} x_i y_i}{(\sum_{i=1}^{p} x_i^2+\sum_{i=1}^{p} y_i^2-\sum_{i=1}^{p} x_i\, y_i)} \quad (11)$$

Overlapping similarity: The overlapping coefficient (or, Szymkiewicz-Simpson coefficient) is a similarity measure related to the Jaccard index that measures the overlap between two sets, and is defined as the size of the intersection divided by the smaller of the size of the two sets [33].

$$f(x,y)=\frac{\sum_{i=1}^{p} x_i y_i}{\mathrm{Min}(\sum_{i=1}^{p} x_i^2,\sum_{i=1}^{p} y_i^2)} \quad (12)$$

If set X is a subset of Y or the converse, then the overlap coefficient is equal to one.

Dice similarity: The Dice Similarity coefficient of two vectors is twice the sum of dot product of the vector divided by the sum of the second degrees of the vectors [33]. It is given by,

$$f(x,y)=2\,\frac{\sum_{i=1}^{p} x_i y_i}{(\sum_{i=1}^{p} x_i^2+\sum_{i=1}^{p} y_i^2)} \quad (13)$$

D. Radial Basis Function Neural Network (RBFNN)

A Radial Basis Function (RBF) network is a special type of neural network that uses a radial basis function as its activation function [34]. RBF networks are very popular for function approximation, curve fitting, time series prediction, and classification problems. The radial basis function network is different from other neural networks, possessing several distinctive features. Because of their universal approximation, more compact topology and faster learning speed, RBF networks have been widely applied in many fields, in science and engineering. The learning of RBFNN happens in three steps

- Finding the cluster centres of the radial basis function using the K-means clustering algorithm.
- Determining the width of the radial basis function.
- Computing the weights.

A block diagram of an RBF network used in this work is presented in Figure 4. The input layer is the 5-dimensional vector which has to be classified. The entire input vector is passed to each of the RBF neurons.

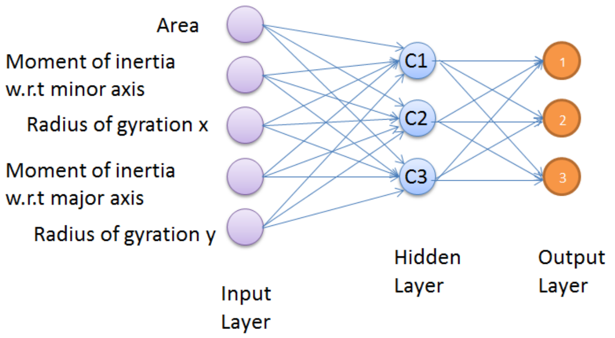

Fig.4. Architecture of RBFNN

A prototype vector is stored by each RBF neuron which is just one of the vectors from the training set. Each RBF neuron compares the input vector to its prototype, and outputs a value between 0 and 1 which is a measure of similarity [35]. Each RBF neuron computes a measure of the similarity between the input and its prototype vector taken from the training set. Input vectors which are more similar to the prototype return a result closer to 1. There are different possible kernel functions, but the most popular is based on the Gaussian which is given by

$$\Phi = \exp\left[-\frac{\sum_{i=1}^{3}||x-\mu_i||^2}{2\sigma^2}\right] \qquad (14)$$

Where x is the input pattern, μ is the centroid of RBF unit for input variables and σ is the width of the RBF unit. Each RBF neuron will produce its largest response when the input is equal to the prototype vector. The linear sum of multiples of output of the central layer neurons and weights of connections will lead to outputs of the three neurons present in output layer. The gradient descent method is applied for finding centres, spread and weights by minimizing the (instantaneous) squared error. Activation functions become the input to this method. The RBFNNs trained by the gradient-descent method is capable of providing the equivalent or better performance compared to that of the multi-layer feed forward neural network trained with the back propagation [36].

E. Support Vector Machine (SVM)

Support Vector machines can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory that implements a learning bias derived from statistical learning theory. Support vector machine was initially popular with the Neural Information Processing Systems (NIPS) community and now is an active part of the machine learning research around the world. SVM becomes famous when, using pixel maps as input; it gives accuracy comparable to sophisticated neural networks with elaborated features. SVM is a supervised learning model with associated learning algorithm that analyses data and recognize patterns, used for classification and regression analysis. It takes a set of input data and predicts, for each given input which of two possible classes forms output, making it a non-probabilistic binary linear classifier. SVM performs classification tasks by constructing hyper planes in a multidimensional space.

Sequential Minimal Optimization (SMO) is a simple algorithm that can quickly solve the SVM without any extra matrix storage and without using numerical optimization steps at all. SMO decomposes the overall quadratic problem (QP) into QP sub-problems[37]. Unlike the previous methods discussed, SMO chooses to solve the smallest possible optimization problem at every step. For the standard SVM QP problem, the smallest possible optimization problem involves two Lagrange multipliers, because the Lagrange multipliers must obey a linear equality constraint. At every step, SMO chooses two Lagrange multipliers to jointly optimize, finds the optimal values for these multipliers, and updates the SVM to reflect the new optimal values.There are four steps to implement in SMO which is shown in Figure 5.

Step1: Find α^1 as the initial feasible solution. Set $k=1$

Step2: If α^k is a stationary point of (2), stop. Otherwise, find a two-element working set $B=\{i,j\}\subset\{1,\ldots,l\}$.

Define $N\equiv\{1,\ldots,l\}\setminus B$ and α_B^k and α_B^k as sub vector of α^k corresponding to B and N respectively.

Step3: If $a_{i,j}=K_{ii}+K_{jj}-2K_{ij}>0$

Solve the following sub-problem with the variable α_B:

$$\min \frac{1}{2}\begin{bmatrix}\alpha_i & \alpha_j\end{bmatrix}\begin{bmatrix}Q_{ii} & Q_{ij}\\ Q_{ij} & Q_{jj}\end{bmatrix}\begin{bmatrix}\alpha_i\\ \alpha_j\end{bmatrix}+(p_B+Q_{BN}\alpha_N^k)^T\begin{bmatrix}\alpha_i\\ \alpha_j\end{bmatrix} \quad (4)$$

Subject to, $0\le\alpha_i,\alpha_j\le C,$

$y_i\alpha_i+y_j\alpha_j=\Delta-y_N^T\alpha_N^k,$

Else solve

$$\min \frac{1}{2}\begin{bmatrix}\alpha_i & \alpha_j\end{bmatrix}\begin{bmatrix}Q_{ii} & Q_{ij}\\ Q_{ij} & Q_{jj}\end{bmatrix}+(p_B+Q_{BN}\alpha_N^k)^T\begin{bmatrix}\alpha_i\\ \alpha_j\end{bmatrix}$$

$$\frac{\tau-\alpha_{ij}}{4}((\alpha_i-\alpha_i^k)^2+(\alpha_j-\alpha_j^k)^2)$$

(5)

subject to constraints of (4)

Step4: Set α_B^{k+1} to be the optimal solution of (4) and $\alpha_N^{k+1}=\alpha_N^k$ set $k\leftarrow k+1$ and go to **step 2.**

Fig. 5. The SMO algorithm

V. Comparative Analysis

It is very difficult to conduct a systematic study comparing the impact of supervised classification algorithms on the quality, of the groups they generate. This is because objective evaluation of cluster quality is difficult in itself. In practice, manually assigned category labels are usually used as baseline criteria for evaluating clusters. As a result, the clusters, which are generated in an unsupervised way, are compared to the priory defined category structure, which is normally created by human experts. This kind of evaluation assumes that the objective of clustering is to replicate human thinking, so a clustering solution is good if the clusters are consistent with the manually created categories. However, in practice datasets often come without any manually created categories and it is exactly here that the significance of clustering lies [38].

The performance measures like accuracy, precision, sensitivity, specificity and F-measure were used to evaluate the classification methods used. These measures are briefly presented.

Accuracy: The accuracy of the classifier is the percentage of test samples that are correctly classified. It is given by,

$$\text{Accuracy}=\frac{\text{Number of true positives+number of true negatives}}{\text{Number of true positives+false positives+false negatives+true negatives}}$$

Precision: Precision is defined as the proportion of the true positives against all the positive results i.e., both true positives and false positives.

Precision: Precision is defined as the proportion of the true positives against all the positive results i.e., both true positives and false positives.

$$\text{Precision}=\frac{\text{Number of true positives}}{\text{Number of true positives+number of false positives}}$$

Sensitivity: Sensitivity is also referred to as true positive rate, i.e., proportion of tuples that are correctly identified.

$$\text{Sensitivity}=\frac{\text{Number of true positives}}{\text{Number of true positives+false positives}}$$

Specificity: Specificity is the true negative rate that is the proportion of negative tuples that are correctly identified.

$$\text{Specificity}=\frac{\text{Number of true negatives}}{\text{Number of true negatives+number of false positives}}$$

The three groups as assorted by the respective algorithms is presented in Table 4.

Table 4. Assorting of 605 ear images to groups

	Group-1	Group-2	Group-3
K-means	247	278	80
BPN	226	289	90
Jaccard	213	300	92
Cosine	177	230	198
Dice	213	300	92
Overlapping	265	339	1
RBFNN	278	247	80
SVM	331	162	82

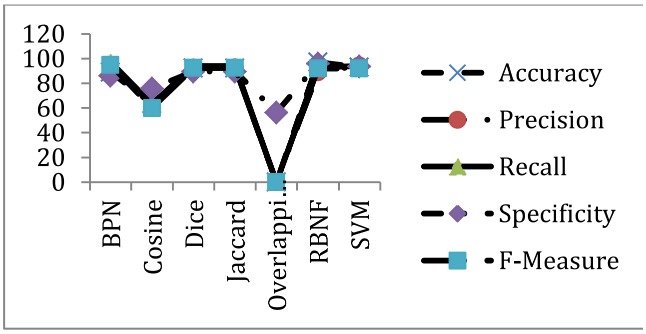

Fig. 6: Performance measures of various classifiers

From Table 4, it is imperative that K-means, BPN, Jaccard and Dice similarity measures, RBFNN and SVM showed marked consistency in terms of number of assorted ear images into three groups. However cosine and overlapping similarity measure proved to be different.

References

[1] Anika Pug, Christoph Busch, Ear Biometrics: A Survey of Detection, Feature Extraction and Recognition Methods, IET Biometrics, Institution of Engineering and Technology, Darmstadt, Germany, 2012, pp 1-35.

[2] Ayman Abaza, Arun Ross, Christina Hebert, Mary Ann F. Harrison, Mark S. Nixon, A Survey on Ear Biometrics, ACM Computing Surveys, Vol. 45, No. 2, 2013.

[3] Bertillon A. La Photographie Judiciaire: Avec Un Appendice Sur La Classification Et L'Identification Anthropometriques'. Gauthier-Villars, Paris; 1890

[4] Iannarelli AV, Ear identification, Paramount Publishing Company, 1989.

[5] Meijerman L, Sholl S, Conti FD, Giacon M, van der Lugt C, Drusini A, Exploratory study on classification and individualization of ear prints, Forensic Science International, 2004;140(1), pp 91-99.

[6] Singh P, Purkait R., Observations of external ear An Indian study, HOMO -Journal of Comparative Human Biology, 2009, 60(5), pp 461-472.

[7] Satish R. Kolhe, Ranjana S. Zinjore, "Clustering Iris Data using Supervised and Unsupervised Learning" , International Journal of Computer Science and Application Issue 2010, ISSN 0974-0767.

[8] Chhanda Ray, Krishnendu Sasmal, "A New Approach for Clustering of X-ray Images" , IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 4, No 8, July 2010 ISSN (Online): 1694-0784 ISSN (Print): 1694-0814

[9] Xi cheng, Sergey Tulyakov and Venu Govindaraju, "Utilization of Matching Score vector Similarity Measures in Biometric Systems" , 978-1-4673-1612-5/12/$31.00 2012 IEEE.

[10] Roman V, Yampolskiy and Venu Govindaraju, "Similarity Measure Function for Strategy-Based Biometrics", World Academy of science,Engineering and Technology 24 2006.

[11] Satya Chaitanya Sripada Dr. M.Sreenivasa Rao , "Comparison of purity and entropy of k-means clustering and fuzzy c means clustering", Indian Journal of Computer Science and Engineering (IJCSE).

[12] Vivek Jaglan, Vikas Thada , "Comparison of Jaccard, Dice, Cosine Similarity Coefficient To Find Best Fitness Value for Web Retrieved Documents Using Genetic Algorithm", published in International Journal of Innovations in Engineering and Technology (IJIET).

[13] Mai V, Khalil I, Meli C, "ECG biometric using multilayer perceptron and radial basis function neural networks", 33rd Annual International Conference of the IEEE EMBS Boston, Massachusetts USA, August 30 - September 3, 2011.

[14] A. Sulong , Wahyudi and M.D. Siddiqi, "Intelligent Keystroke Pressure-Based Typing Biometrics Authentication System Using Radial Basis Function Network", 2009 5th International Colloquium on Signal Processing & Its Applications (CSPA).

[15] Chatterjee, A., Fournier, R., Nait-Ali, A., Siarry, P., "A Postural Information-Based Biometric Authentication System Employing S-Transform, Radial Basis Function Network, and Extended Kalman Filtering", Instrumentation and Measurement, IEEE Transactions on (Volume:59 , Issue: 12).

[16] Anand Viswanathan, S. Chitra, "Optimized Radial Basis Function Classifier for Multi Modal Biometrics", Research Journal of Applied Sciences, Engineering and Technology 8(4): 521-529, 2014 ISSN: 2040-7459; e-ISSN: 2040-7467.

[17] Ankit Chadha, Neha Satam, Vibha Wali, "Biometric Signature Processing & Recognition Using Radial Basis Function Network", CiiT International Journal of Digital Image Processing, ISSN 0974 – 9675 (Print) & ISSN 0974 – 956X (Online) September 2013.

[18] Ashok.J, Rajan.E.G, (2010)"Writer Identification and Recognition Using Radial Basis Function", Int. Jour. of Computer Science and Information Technologies, 1(2), 51-57.

[19] Edson J.R. Justino, Flavio Bortolozz, Robert Sabourin, A comparison of SVM and HMM classifier in the offline signature verification, available online at www.sciencedirect.com.

[20] Juan-Manuel Ramirez-Cortes, Pilar Gomez-Gil, Vicente Alarcon-Aquino, David Baez-Lopez, Rogerio Enriquez-Caldera, A Bometric System Based on Neural Networks and SVM Using Morphological Feature Extraction from Hand-Shape Images, INFORMATICA (2011), Vol. 22, No. 2, 225–240.

[21] Tobias Scheidat, Marcus Liech, Mark Alxender, and Claus Vielhauer, Support Vector Machines for Dyanmic Biometric Handwriting Classification, AIAI-(2009) Workshops Proceedings, pp 118-125.

[22] W.J. Scheirer, A. Bendale, and T.E. Boult, Predicting Biometric Facial Recognition Failure with Similarity Surfaces and Support Vector Machines, IEEE-Conference on Biometrics (2008), pp 1-8.

[23] M.A.Jayaram, Prashanth.G.K., Sachin.C.Patil,(2015) ,Inertia based ear biometrics: A novel approach, Journal of Intelligent systems, De_Gruiter publishing(ahead of print version-online).

[24] Egor P Popov, “Engineering mechanics of solids, Easter economy edition, 2nd edition 1998”.

[25] Braja M. Das, Paul C. Hassler, “Statics and Mechanics of materials”, Prentice Hall, 1988.

[26] Yang Mingqiang, Kpalma Kidiyo1 , Ronsin Joseph, A Survey of shape feature extraction techniques, Pattern Recognition Techniques, Technology and Applications, availabl at: www.intechopen.com , pp1-49.

[27] Nitin Bhatia, Vandana, Survey of Nearest Neighbor Techniques, International Journal of Computer Science and Information Security, Vol. 8, No. 2, 2010, pp 302-305.

[28] D.Pham, S. Sagiroglu, “Training multilayered perceptrons for pattern recognition: a comparative study of four training algorithms”, International Journal of Machine Tools and Manufacture, vol.41, 2001, pp. 419–430.

[29] Anna Haugh, “Similarity measures for text Document clustering”, NZCSRSC 2008,April 2008,Christchurch,New Zealand

[30] B. Larsen and C. Aone, “Fast and effective text mining using linear-time document clustering” , In Proceedings of the Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1999.

[31] Meghana Katri, Cosine similarity for the temporal dynamic web data, International Journal of Computer Science and Engineering Technology, Vol 3, No 8, 2012, pp315-318.

[32] Suphatkit Niwattanakul,Jatsada Singthongchai,Ekkaachai Naenudorn and Supachanun Wanapu ,Using of Jaccard coefficient for keywords similarity, Proceedings of the International MultiConference of Engineers and Computer Scientists 2013 Vol I, IMECS 2013, March 13 - 15, 2013, Hong Kong.

[33] Reza Bosagh Zadeh and Ashish Goel , Dimension Independent Similarity Computation, Journal of Machine Learning Research (2012).

[34] GhanaShyam Nath Nayakam, Study of similarity coefficients using map reduce programming model, A Paper Submitted to the Graduate Faculty of the North Dakota State University of Agriculture and Applied Science

[35] Broomhead, D.; Low, D. Multivariable functional interpolation and adaptive networks. Complex Systems 1988, 2, 321–355.

[36] Riyadh A.K. Mehdi, The Effect of Dimensionality Reduction on the Performance of Software Cost Estimation Models, International Journal of Engineering and Innovative Technology (IJEIT) Volume 4, Issue 9, March 2015.

[37] C.A.Cao, S.S.Keerthi, C.. Ong, P, Uvaraj, X.J. Fu, H.P.Lee.,Developing a parallel sequential minimal optimization for fast training support vector machine ,Neurocomputing, Vol 70(3), 2006, pp93-104

[38] Y. Zhao and G. Karypis, Evaluation of hierarchical clustering algorithms for document datasets Proceedings of the International Conference on Information and Knowledge Management, ACMNew York, 2002, pp 515-524.