Experimental Analysis on Character Recognition using Singular Value Decomposition and Random Projection

Manjusha K.[1], Anand Kumar M.[2], Soman K. P.[3]
Centre for Excellence in Computational Engineering and Networking, Amrita Vishwa Vidyapeetham
Amrita School of Engineering, Coimbatore, India
[1] manjushagecpkd@gmail.com
[2] m_anandkumar@cb.amrita.edu
[3]kp_soman@amrita.edu

***Abstract*—Character recognition, a specific problem in the area of pattern recognition is a sub-process in most of the Optical Character Recognition (OCR) systems. Singular Value Decomposition (SVD) is one of the promising and efficient dimensionality reduction methods, which is already applied and proved in the area of character recognition. Random Projection (RP) is a recently evolved dimension reduction algorithm which can scale with large dataset. In this paper, we have applied SVD and RP as feature extraction technique for character recognition and experimented with k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) classifiers. Our experiments are conducted on MNIST handwritten digit database and Malayalam character image database. On both databases SVD features could achieve lower misclassification rate than RP based features. SVD features with SVM classifier using RBF kernel could achieve a misclassification rate of 2.53% on MNIST database and 2.87% on Malayalam character image database.**

Keyword-Character Recognition, Singular Value Decomposition, Random Projection, Pattern Recognition, Support Vector Machine, Dimensionality Reduction

I. INTRODUCTION

Character recognition research dates back to 1950's. It is the process of identifying the character or symbol represented by the given image. It comprises of feature extraction followed by classification process. Feature extraction in the context of character recognition is mapping the segmented character image to a real valued vector which can better describe the character on that image. Features extracted from character images of different classes should have a good separating boundary to achieve good accuracy in character classification. Thus the feature extraction algorithm used character recognition can greatly influence the accuracy of whole character recognition process. Sometimes the different feature extraction algorithms can be combined together to achieve a higher accuracy in character classification. To deal with high dimensional feature vector, an efficient dimension reduction technique would be desirable.

The best way to reduce the dimension of data is to project the data vector to a reasonably low-dimensional orthogonal subspace that can capture most of the variations present in the high dimensional data. The amount of distortion in data and the computational complexity are the two evaluation parameters for dimension reduction algorithms [1]. SVD is one of the most widely used dimension reduction method which is a matrix factorization technique, which breaks the data matrix into simple and meaningful pieces. It has been applied in several data processing applications successfully [15][16][17]. It can transform the data matrix from higher to lower dimension without much information loss. SVD is applied to the character image data set and the eigenvectors associated with the most significant principal components are retained for classification. The disadvantage of SVD is that it becomes computationally expensive as the data dimension increases.

Random Projection (RP) is recently emerged simple, powerful computationally efficient dimension reduction technique [6]. RP is based on random matrix to project data to lower dimensional subspace whose columns have unit length. Another advantage of RP over SVD is that it is data independent. The random directions we are choosing does not depend upon the dataset we are using. RP have strong theoretical backgrounds [2] and successfully applied in different machine learning applications [4][5][6][8]. Our paper is a comparative study between SVD and RP on character recognition process.

After feature extraction, the feature vector is mapped to a class label or membership scores with the help of machine learning algorithms. Different supervised classification algorithms can be used with character recognition. k-Nearest Neighbor (k-NN) and Support Vector Machine (SVM) algorithms are used for our experimental purpose. k-NN finds k closest neighbors of given feature vector from the training samples and decides the final class label while SVM creates hyper plane which maximally separates different classes, and is used for classifying the new test feature vector [18]. Instead of applying the dimensionality reduction on image itself, it can be applied on features extracted from the image. Wavelet and Radon transform can produce robust feature descriptors [19][20]. With wavelet transform, the image can be decomposed into four sub-images with the high- and low-pass filters. Radon transform represents the directional features in the image [20]. Besides transform domain features, structural features can also produce good feature descriptors.

Our experiments are conducted on MNIST hand written digit database and Malayalam character image database. The challenges in terms of recognition for Malayalam Language script, the official language of Kerala are the similarity among character symbol shapes and large number of characters present in the script. Malayalam language have a large number of character symbols representing, vowels, consonants, half-consonants (known as chillu), vowel modifiers and conjunct characters (formed from consonant-consonant/vowel combination). Script revision and nonstandard font design are other main challenges in Malayalam character recognition [12]. Neeba and Jawahar [21] conducted an empirical study on different pattern classification schemes for Malayalam language script and observed that SVM outperforms other classification approaches on Malayalam character recognition. But the performance of SVM degraded with the increase in number of classes. Neural Network classifiers with wavelet features on Malayalam character recognition [22] could achieve an accuracy of 92%. SVD have applied on 90 classes of Malayalam Language [14] for dimensionality reduction and then classified the features in lower dimension using k-Nearest Neighbor (k-NN) algorithm with an accuracy of 94%.

In this paper, we are applying the technique described in [14] for creating SVD based features. RP features are created using the algorithm mentioned in Section 3. The performance of both the features are evaluated using k-NN and SVM classifiers on Malayalam character image database and MNIST hand written digit image database.

Section 2 and 3 describes about the technical details about SVD and RP respectively. Section 4 is about the classification algorithms used in our experiments. Section 5 represents our detailed experiments and results. Section 6 is the conclusion and summary of our work with SVD and RP on character recognition.

II. Singular Value Decomposition (SVD)

SVD represents or transforms the data in the matrix to those axes in which variation among the data is maximum while keeping the total variation among data same. It is basically a matrix factorization method, which factorizes any rectangular matrix A of size $m \times n$ to three matrices U, V and S. U is an $m \times m$, orthogonal matrix and its columns are eigen vectors of AA^T. V is an $n \times n$ orthogonal matrix with its columns, eigen vectors of A^TA. S is an $m \times n$ rectangular diagonal matrix with positive real values on its diagonal and the diagonal values are decreasingly ordered [25]. The basic SVD process can be alternatively described as:

If the rank of A is n, then $U = (u_1\ u_2\ u_3\ \ldots.\ u_n)$, $V = (v_1\ v_2\ v_3\ \ldots\ v_n)$ and $S = \text{diag}\ (s_1, s_2, s_3, \ldots s_n)$. A can be represented as in (1)

$$A = \sum_{i=1}^{n} s_i u_i v_i^T \tag{1}$$

In the summation given in (1), even if we discard the last few terms it can approximate A. In last few terms actually noise is captured and those values don't contribute to the information contained in A. This property is actually helping for dimensionality reduction and noise removal using SVD.

In our experiments, we have created the SVD features of training and testing character images using the technique described in [14]. The algorithm finds the U, S and V factors of data matrix D using SVD. The data matrix is created by appending vectorized form of all the character images in training dataset. Then subset of V and S are multiplied together to represent the reduced representation of training images. These are training features. The testing features are created by projecting each testing character image vector to the subset of U. The subset of V, S and U are selected based on the number of singular values we are considering for the experiment. The number of singular values is the varying parameter for our SVD based feature creation process.

III. Random Projections (RP)

Random Projection is achieved by projecting the d-dimensional data on to k dimensional space using random matrix where $k << d$ (k is very less than d). Let our data matrix is denoted by D of size $d \times N$ where, D denotes the feature data dimension and N is the number of data observations. Let R be the random matrix of size $k \times d$. The columns of R are normalized to get an approximate orthonormal matrix. Then the random projection P can be calculated as shown in equation (2).

$$P = RD \tag{2}$$

The dimension of P will be of $k*N$. Here we can see that the data dimension is reduced from d to k. The RP can also be considered as a linear transformation in which the basis is changed to random vectors. The whole process is of order O(dkN). Random projection relies on the result that; in high dimensional space, there is large number of almost orthogonal than orthogonal directions [2]. As the dimension increases, the random matrix vectors will be almost orthogonal and the matrix $R^T R$ will be approximate identity matrix.

Random projection preserves the pair-wise distances between the data vectors even in the reduced low dimension while SVD does not give guarantee about this. The inner product between the two data vectors are distorted by random mapping is almost zero on the average and its variance is at most the inverse of the reduced dimension if the random matrix is chosen from normal distribution with zero mean and variance equals to one [1]. Apart from inner product, Euclidean measure can be utilized to find the similarity between data vectors. Johnson-Lindenstrauss lemma [3] states that the data dimension can be reduced to $k > \log(d)/\varepsilon$ such that distances between the points are approximately preserved (i.e., not distorted more than a factor of $1\pm \varepsilon$, for any $0 < \varepsilon < 1$). But some previous studies [1] suggests that the lower bound for k is not tight and even lower k values gives good results.

Random projection algorithm for feature extraction can be summarized as follows

1. Generate random matrix, R of size $k\times d$
2. Normalize the columns of R
3. Find the random projection result $P = RD$

The 3rd step can be parallelized if the number of data samples we are taking is very large. Each column of D can be taken and find the projection vector. In RP, the feature size is dependent the parameter k, the row size of random matrix R. Feature size increases with increase in k.

IV. Feature Classification

After mapping the character images from image space to feature space, the feature vector has to be mapped or classified to one of the predefined character classes. Supervised and unsupervised machine learning algorithms can be applied for feature classification purpose. In this paper, we have used two supervised machine learning algorithms for classifying the feature vector, k-Nearest Neighbor (k-NN) and Support Vector Machine.

A. K-Nearest Neighbor Classifier

k-NN is one of the simple machine learning technique, which can be used for both classification and regression. In k-NN, no explicit training phase is needed, it usually store all the features extracted from the training images along with their class labels. During testing phase, it labels the new instances (features extracted from the testing images) with that class which is most common among the k training feature neighbors calculated with the help of similarity measuring techniques. In k-NN, two parameters can be varied; k value and the similarity measuring technique for calculating distance. Cross validation approach can be used for estimating k value for which k-NN gives good accuracy on the dataset. Different distance functions can be used for calculating similarity between testing feature and training features. In our experiments, we have used Euclidean distance as the distance function.

B. Support Vector Machine

SVM classifier is basically a binary (two-class) classifier based on the use of discriminant functions, which represents a surface that separates the observations so that the observations from the two classes lie on the opposite sides of the surface. Consider labeled training dataset (x_i, y_i) ; $i = 1, ... n$; y_i € {-1, 1}, x_i € R^d where n is the total number of samples in training set, d is the dimensionality of each sample. Suppose we have a hyperplane that separates the positive from the negative samples. The points x which lie on the hyperplane satisfy $wx + b = 0$, where w is normal to the hyperplane and b is the offset [26]. For the linearly separable case, SVM finds a separating hyperplane with largest margin; the point where the training data satisfy the following constraints.

$$w.x_i + b \geq 1, \text{ for } y_i = +1$$
$$w.x_i + b \leq 1, \text{ for } y_i = -1$$

Identification of the optimal hyperplane for separation involves maximization of an appropriate objective function, i.e, solving the following quadratic optimization problem in (3) during training an SVM.

Maximize

$$\sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i, x_j)$$
$$subject\ to\ \ \alpha_i \geq 0 \tag{3}$$
$$\sum_{i=1}^{n}\alpha_i y_i = 0;\ i = 1, 2, ...n$$

where α_i are the Lagrangian multipliers corresponding to each of the training data points x_i. The function K is the kernel function. It is defined as $K(x, y) = \Phi(x). \Phi(y)$, where ø maps the data points in d dimensions to a higher dimensional space [26] . Kernels used in our experiments are linear, polynomial and radial basis function (RBF).

1. Linear Kernel :

$$K(x, y) = x^{\mathrm{T}}y \qquad (\Phi \text{ is identity – no mapping})$$

2. Polynomial Kernel :

$$K(x, y) = \left(x^{\mathrm{T}}y\right)^d \text{ or} \left(1 + x^{\mathrm{T}}y\right)^d \quad (d \text{ represents degree})$$

3. Radial Basis Function (RBF) Kernel :

$$K(x, y) = \exp\left[-\gamma \left\|x - y\right\|^2\right] \ (\gamma \text{ represents kernel bandwidth})$$

The result of the training phase in SVM is the identification of a set of labeled support vectors and a set of coefficients α_i. Support vectors are the samples near the decision boundary and most difficult to classify. The decision of test sample x is made from (4).

$$y = sign(\sum_{i=1}^{n}\alpha_i y_i K(x_i, x)) \tag{4}$$

For extending binary SVM to multi class classification problem, there are two approaches: - Direct and Indirect [26]. In indirect approach several binary SVM are constructed and are combined together to find final target class (one-against-one). In direct approach all the classes are considered in a single optimization formulation (one-against-rest). In our experiments we have used one-against-one approach in multiclass SVM based classification.

V. EXPERIMENTAL RESULTS

Experiments are conducted on our Malayalam character's image database and MNIST hand written digit database [24]. Malayalam character image database is built from scanned document images of Malayalam language story books and magazines. The image database includes the consonants, vowels, vowel modifiers (diatrics), conjunct characters of Malayalam language script. In total 135 character classes are considered. 75% character images of each character classes is taken as training images and the rest 25% as testing images. MNIST hand written digit database contains totally 60,000 training samples and 10,000 testing samples. MATLAB tool is used as experimental platform for conducting our trials. Misclassification rate is considered as the evaluation measure to decide the performance of techniques used in our experiments. Misclassification rate is found by calculating the percentage of testing samples classified incorrectly among total number of test samples.

For the first experiment, SVD features extracted from character images in Malayalam character image database is first tested with k-NN classifier. SVD features for different number of singular values are extracted and are classified with different k values (k=1,2,…9) in k-NN classifier. Misclassification rate obtained for features of different numbers of singular values are analyzed and the range of singular values for which lowest misclassification rate obtained is found out. The result obtained is listed in TABLE. I. The lowest misclassification rate is obtained for k=1 when the number of singular values is in the range 61-72. For all other k values till 9, the lowest misclassifications occurred on the range of 81-90. A graph is drawn in between the lowest misclassifications and different k values in k-NN classifier, and is shown in Fig.1.

TABLE I. LOWEST MISCLASSIFICATION RATE OBTAINED FOR K-NN CLASSIFIER FOR SVD FEATURES ON MALAYALAM CHARACTER IMAGE DATABASE

k	Lowest Misclassification Variation (%)	Range of Number of Singular Values
1	5.94-6.27	61-72
2	6.86-7.36	79-92
3	6.27-6.86	82-92
4	6.44-6.88	81-92
5	6.81-7.18	82-92
6	6.94-7.36	80-90
7	6.78-7.21	82-92
8	7.39-7.86	83-93
9	7.23-7.52	80-91

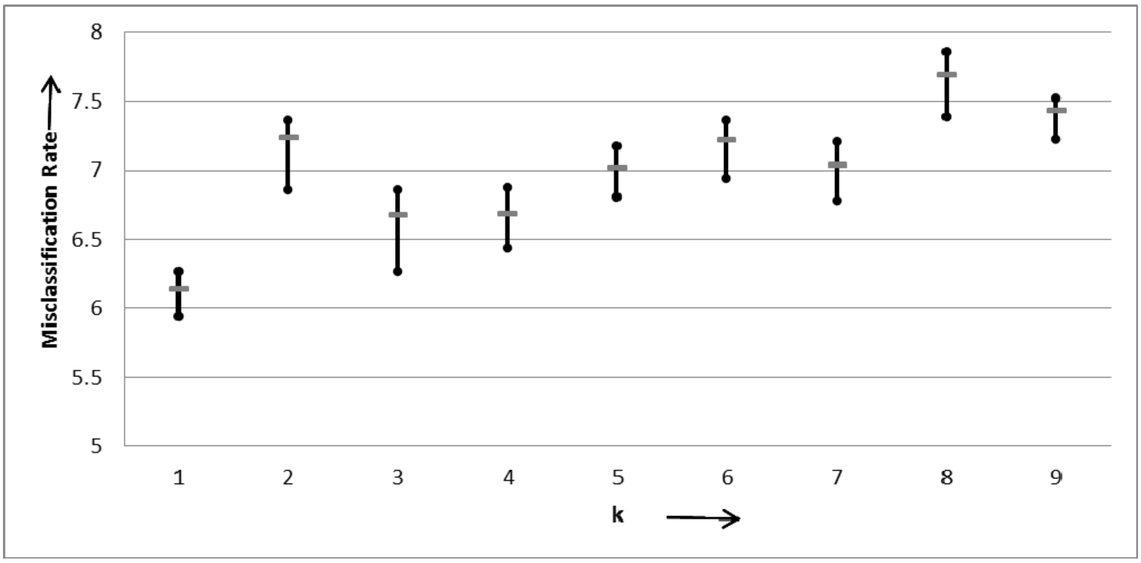

Fig. 1. Visual representation of the misclassification rate variation obtained for the range of singular values listed on TABLE I in k-NN classifier for different k values on Malayalam character image database. Among different k values in k-NN, 1-NN is getting lowest misclassification rate.

Then, the same SVD features are tested with SVM classifier. For testing with SVM multiclass classifier we have used LibSVM tool [23]. During training process, Linear, Polynomial and Radial Basis Function (RBF) kernels in SVM classifier are applied and tested with the features. The misclassification rate obtained for SVD features created by considering different singular values using different kernels on SVM classifier are shown in Fig 2. RBF kernel is performing better than linear and polynomial kernels. Average misclassification rate of 2.87% is obtained with RBF kernel on SVM multiclass classifier for number of singular values in the range 145-155. From Fig. 2 it is obvious that for small number of singular values, the misclassification rate is high; up on increasing the number of singular values the misclassification rate tends to decrease. On Malayalam character image database the SVD - SVM combination performs better than SVD - k-NN combination.

In experiments mentioned above, we have used vectorized form of character images for applying SVD. Instead of using image vector directly, a pre-processing can be applied to character images to extract features. After pre-processing, SVD can be applied on those extracted features. We have used Wavelet Transform, Radon Transform and Projection Profile features on Malayalam character image database for pre-processing. Wavelet transform features are extracted by finding the approximation co-efficient after applying wavelet decomposition using haar wavelets on the character images. Radon transform computes the line integrals from multiple sources along parallel paths, in a certain direction; in our experiments angle is varied in between 0 and 180 degree. Projection profile features are computed by summing the pixel values in character image horizontally and vertically. The same SVD approach is applied for the extracted features with SVM classifier and the misclassification rate obtained is tabulated in TABLE II.

TABLE II. SVD APPROACH WITH DIFFERENT PRE-PROCESSING METHODS ON MALAYALAM CHARACTER IMAGE DATABASE

Feature Name	Least Misclassification Rate Obtained (%)
Raw Image (without pre-processing)	2.81
Wavelet Transform	2.51
Radon Transform	4.5
Projection Profile	17.18

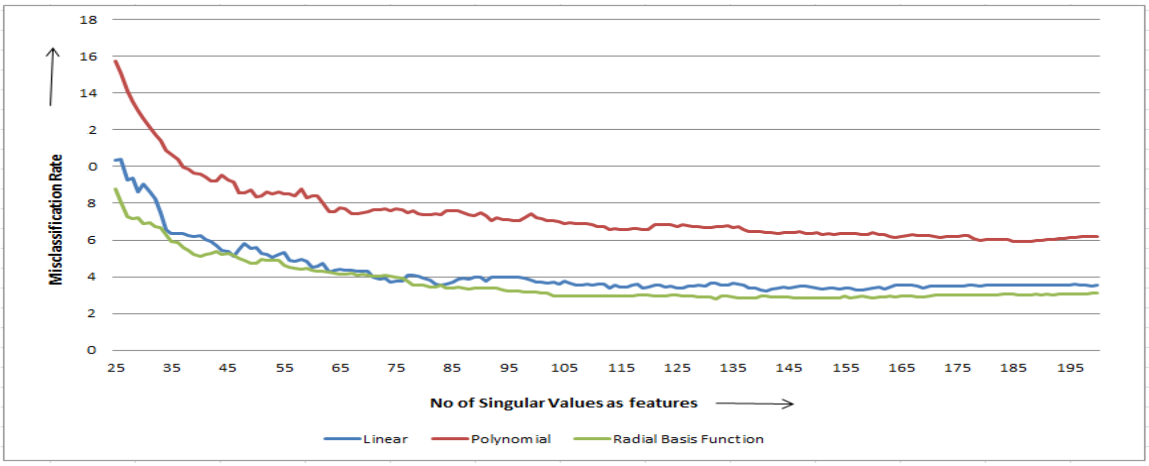

Fig. 2. Graph shows the misclassification rate on Malayalam character image database, when SVD features from different singular values are classified with SVM classifier. Linear, polynomial and radial basis function based kernels inside SVM are experimented.

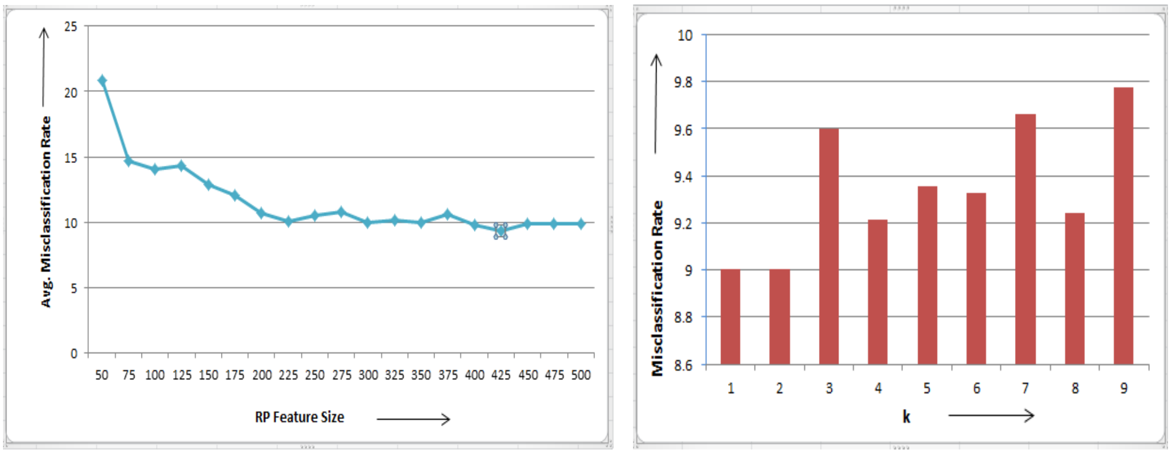

Fig. 3. a)Graph shows the average misclassification rate obtained for the RP features classified with k-NN on Malayalam character image database. The lowest misclassification is for RP feature size = 425 b) Misclassification for different k in k-NN classifier for RP feature size 425.

In our approach, SVD works on whole of the training dataset to generate its factor matrices. As the number of images in training set increases, applying SVD becomes computationally expensive. RP is another dimension reduction algorithm which scales with growing data set. We have applied RP on Malayalam character image database and tested with k-NN and SVM classifiers. The random matrix size determines the number of features in RP. In our experiments with RP on Malayalam character image database, we have varied the RP feature size from 50, 75,100 to 500. These RP features are classified with k-NN and SVM classifier. The misclassification rate obtained for different RP feature size on k-NN classifier is shown in Fig 3. From Fig 3, it is obvious that the misclassification rate decreased with the increase in RP feature size initially. When the RP feature size crossed 425, the misclassification rate is below 10% for all k in k-NN classifier. We have obtained lowest misclassification rate of 9% for RP feature size 425 in 1-NN and 2-NN classifier. The RP features from Malayalam character image database are classified with SVM classifier for different kernels: linear, polynomial and radial basis function. The result obtained is graphically shown in Fig 4. RBF kernel performs better than the other two kernels on RP features. As the RP feature size increased and it reaches 350, linear kernel could achieve almost the same result as that of RBF kernel. RP features of size 350 could achieve a misclassification rate of 4.21% with SVM classifier using RBF kernel.

We have performed the same set of experiments with SVD and RP on MNIST handwritten digit database. In MNIST database each character image is of size 28×28. All the images inside MNIST database are converted to binary images before applying SVD and RP. On vectorized form each binarized character image represents a feature vector of size 784. All the binarized training image vectors are appended together to create the training data matrix. SVD features are extracted for different numbers of singular values from the factorized training data matrix. The obtained features are first experimented with k-NN classifier. The misclassification rate obtained for different values of k in k-NN classifier is listed in TABLE III. From TABLE III it is clear that with for feature size of range 32-38, the 5-NN classifier could achieve the lowest misclassification rate of 3.24%. The results imply that the reduced dimension in thirties from the original dimension of 784 could capture most of the information contained in the training data matrix.

The SVD features from MNIST database are applied on SVM classifier. The misclassification rate obtained for SVD features on different kernels with SVM classifier is shown in Fig 5. The RBF kernel outperforms the linear and polynomial kernel on the SVD features from MNIST. An average misclassification rate of 2.53% is obtained with RBF kernel in SVM classifier on feature size on the range 26-45. From Fig 5, it can be seen that the misclassification rate is high for very small number of singular values and is decreased with the increase in number of singular values and reached a minimum range. Misclassification Rate is slightly increased when the feature size is increased beyond 50 in both RBF and linear kernel. It can be seen that the SVD - SVM combination could achieve a misclassification rate lower than that of SVD - k-NN combination on MNIST database.

The RP features extracted from hand written digit images MNIST database are classified with k-NN and SVM classifier. The RP feature size we have considered from 50 to 500 and tested with k-NN for different k values. The result obtained is listed in TABLE IV. The size of RP features increased, accuracy also increased in k-NN. 4-NN is getting lowest misclassification rate in majority cases of different RP feature sizes. It can be seen that in higher feature dimensions the difference in misclassification rate for different k in k-NN classifier is very less. The lowest misclassification rate is obtained for 4-NN for RP feature size of 450 and 475. The same RP features are tested with SVM classifier with linear, polynomial and radial basis function kernels and the results are listed in TABLE IV. It can be concluded that RBF kernel performs better than linear and polynomial kernels. The lowest misclassification rate of 5.56% is obtained in RP feature size of 150 for SVM classifier with RBF kernel. The results of RP features on k-NN and SVM classifiers shows that there is not much difference in the lowest misclassification rate obtained for both of the classifiers. k-NN could achieve the lowest misclassification rate of 5.98% for RP feature size of 450 while SVM could achieve 5.56% for RP feature size of 150. k-NN could achieve almost comparable misclassification rate as that of SVM classifier but in higher dimension of RP feature on MNIST database.

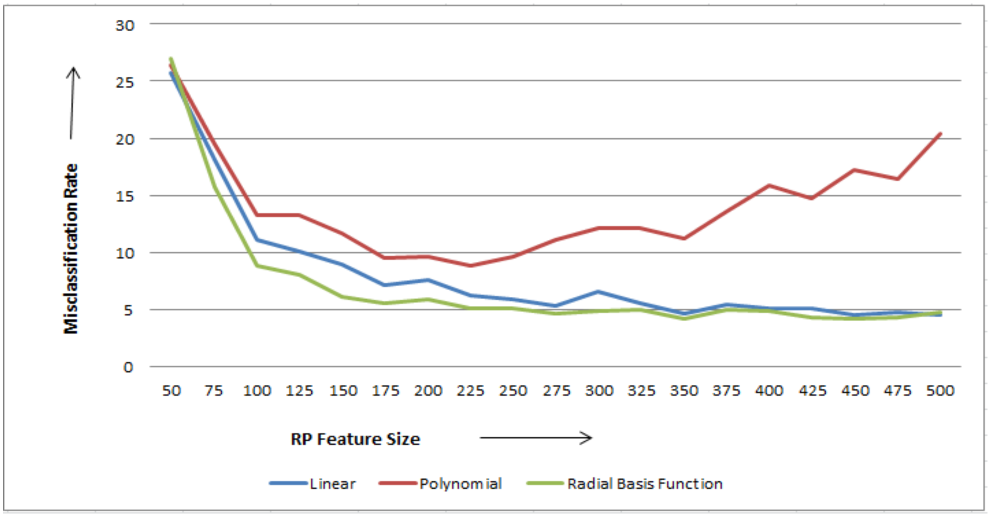

Fig. 4. Graph of misclassification rate Vs RP features, when classified with SVM on Malayalam character image database

TABLE III. LOWEST MISCLASSIFICATION RATE OBTAINED FOR K-NN CLASSIFIER FOR SVD FEATURES ON MNIST DATABASE

k	Range of Number of Singular Values for which Lowest Misclassification Rate obtained	Average Misclassification Rate on the range (%)
1	33-38	3.840
2	33-38	3.840
3	32-36	3.410
4	34-38	3.286
5	32-38	3.248
6	31-38	3.253
7	29-36	3.369
8	25-34	3.392
9	25-34	3.517

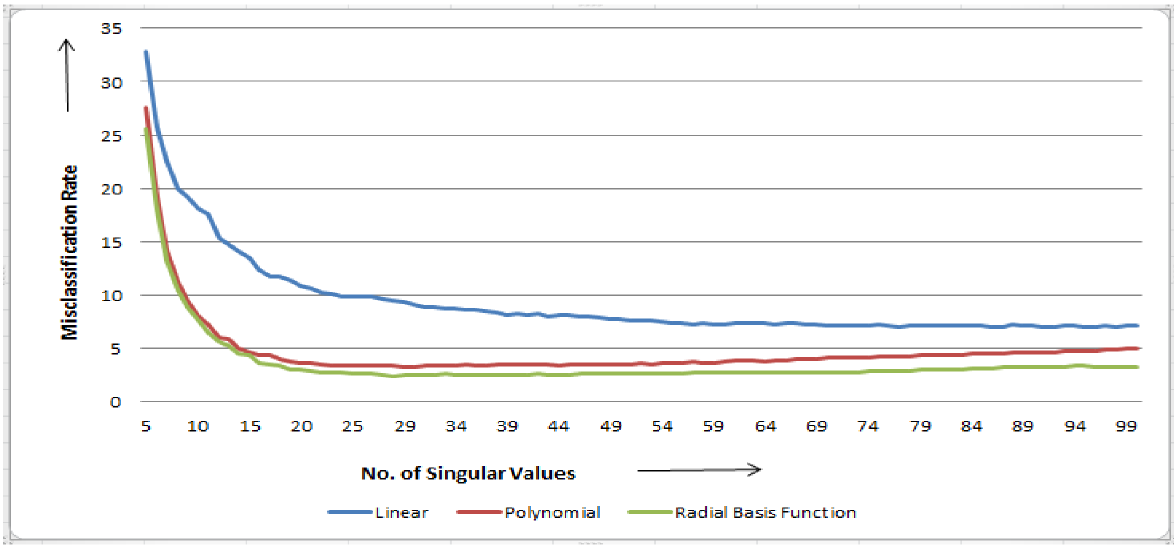

Fig. 5. Graph shows the misclassification rate obtained for SVD features using SVM classifier with linear, polynomial and radial basis function kernels on MNIST database

TABLE IV. MISCLASSIFICATION RATE OF RP FEATURES ON MNIST DATABASE

RP Feature Dimension	**RP+k-NN Misclassification Rate (%)**									**RP+SVM Misclassification Rate (%)**		
	k=1	**k=2**	**k=3**	**k=4**	**k=5**	**k=6**	**k=7**	**k=8**	**k=9**	**Linear**	**Polynomial**	**RBF**
50	14.17	14.17	12.8	12.24	12.22	12	11.85	11.74	12.04	19.49	8.05	7.57
75	11.46	11.46	10.43	10.06	9.86	9.92	9.94	9.98	10.3	15.25	6.48	6.08
100	9.19	9.19	8.52	8.07	8.25	8.03	8.25	8.34	8.42	12.99	7.12	6.01
125	9.26	9.26	8.39	8.16	8.32	8.2	8.39	8.23	8.44	11.08	7.36	5.59
150	8.4	8.4	7.69	7.5	7.8	7.62	7.72	7.63	7.86	10	7.84	$\boxed{5.56}$
175	8.24	8.24	7.61	7.35	7.38	7.46	7.76	7.66	7.83	9.32	9.33	5.66
200	7.55	7.55	7.26	7.12	7.31	7.27	7.45	7.27	7.45	9.06	10.18	5.69
225	7.7	7.7	7.03	6.82	6.93	6.83	6.96	7	7.09	9.11	12.37	5.87
250	7.38	7.38	6.8	6.63	6.7	6.71	7.07	6.9	7.04	8.56	13.08	5.81
275	7.4	7.4	6.9	6.62	6.84	6.87	7.01	7.07	7.27	8.16	13.83	5.93
300	7.46	7.46	6.55	6.42	6.4	6.5	6.73	6.7	7.05	8.24	15.72	6.22
325	6.61	6.61	6.46	6.05	6.57	6.34	6.8	6.73	6.9	7.96	17.05	6.3
350	7.03	7.03	6.39	6.24	6.13	6.05	6.36	6.38	6.58	7.98	18.33	6.13
375	6.84	6.84	6.42	6.23	6.13	6.15	6.42	6.44	6.66	8	20.96	6.32
400	6.9	6.9	6.2	6.29	6.43	6.17	6.46	6.43	6.78	7.54	23.76	6.46
425	6.81	6.81	6.33	6.16	6.26	6.16	6.21	6.26	6.63	7.69	26.38	6.24
450	6.95	6.95	6.22	$\boxed{5.98}$	6.17	6.18	6.44	6.31	6.42	7.61	27.32	6.54
475	6.68	6.68	6.18	$\boxed{5.98}$	6.09	6.13	6.42	6.29	6.5	7.75	31.63	6.68
500	6.93	6.93	6.29	6.13	6.33	6.06	6.31	6.35	6.57	7.68	33.82	6.75

VI. CONCLUSION

Dimension reduction algorithms play a great role in almost all machine learning applications. We have tried to compare the performance of SVD and RP, two dimension reduction techniques with the help of two classification algorithms, k-NN and SVM on the context of character recognition. The experiments are conducted on Malayalam character image database (135 different classes) and MNIST handwritten digit database. SVD features are giving better accuracy with SVM classifier than k-NN on both the databases. In case of RP features, on Malayalam character image database RP – SVM combination is performing better than RP – k-NN combination. On MNIST database the performance of RP – k-NN and RP – SVM is almost same, except that k-NN performs well on high dimensions of RP feature. On comparing the performance of SVD and RP features, SVD dominates over RP on both the image databases we have used in our experiments. But for applications, handling massive data, the SVD approach we have used becomes computationally expensive while RP scales well. Our future work includes improving the performance of RP to make its performance comparable to that of SVD and on making SVD approach practical for applications involving growing datasets.

REFERENCES

[1] Ella Bingham and Heikki Mannila, "Random projection in dimensionality reduction: Applications to image and text data," in Proc. KDD'01, pp. 245–250, 2001.

[2] R. Hecht-Nielsen, "Context vectors: general purpose approximate meaning representations self-organized from raw data," in Computational Intelligence: Imitating Life, IEEE Press, pp. 43–56, 1994.

[3] W. Johnson and J. Lindenstrauss, "Extensions of lipschitz mapping into a hilbert space," in Conf. on Modern Analysis and Probability, vol. 26, pp. 189–206, 1984.

[4] S. Kaski, "Dimensionality reduction by random mapping," in Proc. Int. Joint Conf. on Neural Networks, vol. 1, pp. 413–418, 1998.

[5] S. Dasgupta, "Experiments with random projection," in Proc. UAI '00, pp. 143–151, 2000.

[6] Tetsuya Takiguchi, Jeff Bilmes, Mariko Yoshii and Yasuo Ariki, "Evaluation of Random Projection Based Feature Combination on Speech Recognition," in IEEE International Conf. on Acoustics Speech and Signal Processing (ICASSP), pp. 2150–2153, 2010.

[7] Xiaoli Zhang Fern and Carla E. Brodley, "Random Projection for High Dimensional Data Clustering: A Cluster Ensemble Approach," in Proc. of the 20th International Conference on Machine Learning, pp. 186–193, 2003.

[8] Navin Goel, George Bebis, and Ara Nefian, "Face Recognition Experiments with Random Projection," in Proc. of the SPIE, vol. 5779, p. 426–437, 2005.

[9] Armin Eftekhari, Massoud Babaie-Zadeh and Hamid Abrishami Moghaddam, "Two-dimensional random projection," Signal Processing, vol. 91, pp. 1589–1603, 2011.

[10] Michael B. Wakin and Richard G. Baraniuk, "Random Projections Of Signal Manifolds," in Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing, vol. 5, 2006.

[11] Magnus Sahlgren, "An Introduction to Random Indexing," in Methods and Applications of Semantic Indexing Workshop at the 7th International Conf. on Terminology and Knowledge Engineering, 2005.

[12] Setlur, Srirangaraj. Guide to OCR for Indic Scripts. Edited by Venu Govindaraju. Springer London, 2009.

[13] U. Pal and BB Chaudhuri, "Indian script character recognition: A survey," in Pattern Recognition, Elsevier vol. 37, no. 9, pp. 1887-1899, 2004.

[14] S. Sachin Kumar, K. Manjusha and K. P. Soman, “Novel SVD Based Character Recognition Approach for Malayalam Language Script,” Recent Advances in Intelligent Informatics Advances in Intelligent Systems and Computing, vol. 235, pp. 435-442, 2014.
[15] Soren Holdt Jensen, Per Christian Hansen, Steffen Duus Hansen and John Aasted Sorensen, “Reduction of Broad-Band Noise in Speech by Truncated QSVD,” in IEEE Transactions on Speech and Audio Processing, vol.3, no.6, 1995.
[16] Vladimir I. Gorodetski,Leonard J. Popyack,Vladimir Samoilov and Victor A. Skormin, “SVD-Based Approach to Transparent Embedding Data into Digital Images,” in Proc. MMM-ACNS '01, pp. 435-442, 2001.
[17] W. Zhao, R. Chellappa and A. Krishnaswamy, “Discriminant Analysis of Principal Components for Face Recognition,” in Proc. FG '98, p. 336, 1998.
[18] Mohamed Cheriet, Nawwaf Kharma, Cheng-Lin Liu and Ching Y.Suen,“Character Recognition Systems, A Guide for Students and Practitioners” WILEY-INTERSCIENCE: John Wiley & Sons, Inc., Hoboken, New Jersey, 2007.
[19] Mallat, S., "A theory for multiresolution signal decomposition: the wavelet representation," in IEEE Pattern Anal. and Machine Intell., vol. 11, no. 7, pp. 674-693, 1989.
[20] Dattatray V. Jadhav and Raghunath S. Holambe, “Feature extraction using Radon and wavelet transforms with application to face recognition,” in Neurocomputing, vol.72, pp.1951–1959, 2009.
[21] N. V. Neeba and C. V. Jawahar,“Empirical Evaluation of Character Classication Schemes,” in ICAPR, 7th International Conference on Advances in Pattern Recognition, 2009
[22] M. Abdul Rahiman, M. S. Rajasree, “OCR for Malayalam Script Using Neural Networks,” ICUMT '09, International Conference on Ultra Modern Telecommunications & Workshops, 2009.
[23] Chih-Chung Chang and Chih-Jen Lin, “LIBSVM: a library for support vector machines,” in ACM Transactions on Intelligent Systems and Technology, 2011.
[24] LeCun Y. and Cortes C, 'MNIST handwritten digit database', http://yann.lecun.com/exdb/mnist/ (Accessed on: 10 Nov. 2014).
[25] Kalman, D., “A Singularly Valuable Decomposition: The SVD of a Matrix” in The College Mathematics Journal, 1996.
[26] Soman K. P., Loganathan R., and Ajay V. Machine learning with SVM and other kernel methods. PHI Learning Pvt. Ltd. 2009.