ABSTRACT

INTRODUCTION:

Handwritten digit recognition (HDR) is the ability of a computer to elucidate an obtained comprehensible handwritten input from sources such as paper documents, images and some other devices. Off-line HDR includes transformation of the input into machine recognisable ASCII code which can be used by computer and other applications necessary for the recognition procedure. The data obtained in this process is regarded as a static representation of handwritten digit. Though machine printed digit and character recognition problems are almost sorted, handwritten digit and character recognition still need much effort to accomplish. So, HDR is such a field of research which needs more cultivation. In addition to that, in India, different languages are used in different states. For that we need a general and relatively simple approach that can handle the digit images written in multiple languages with ease.

HDR is considered as an important module of an OCR (Optical Character Recognition) system to be developed for recognizing handwritten textual document images. This involves photo-scanning of the text taking one character at a time, analysis of the scanned-in image, and then conversion of the scanned-in image into character codes, such as ASCII, generally used in data processing through computer system. HDR is also used in: (a) National ID number recognition system (b) Postal office automation with code number recognition on Envelope and (c) Handwritten cheque processing [10]. Recently HDR has been implemented in tablet-based learning and e-content development [11] .

Methods to obtain higher accuracy and less computational requirements are prime requirements in order to develop a comprehensive HDR system. But along with that, in a multilingual environment, we also need such methods that will be able to recognise numerals on a general basis, i.e. which will work on numerals written in different languages.

Neural network classifiers have been immensely used in this field [19, 20] and some are based on Multi-layer Perceptron (MLP) too [19, 20, 21]. The latter has been used to recognize Arabic numerals by a method proposed by Das et al. [3] and again another approach based on MLP has been made by Basu et al. [1] for Bangla numerals. The combination of supervised and unsupervised learning methods on a single algorithm has been implemented on Devanagari numerals by Patil et al. [14] to ensure its use for pure classification and clustering and hybrid classification/clustering. Y. Le Cun et al. [15] have used back-propagation network for HDR. Das et al. [4] have used a novel convex hull based feature set calculated over various bayescharacteristics of the convex hull of a pattern to recognize handwritten Bangla numerals.

The challenges in handwritten digit recognition arise not only from the different ways in which a single digit can be written, but also from the different requirements imposed by some specific applications. Writing style varies from person to person and even it differs for a single person . Thus, it is a more difficult issue to recognise unconstrained handwritten numerals. Any prior hint is also not provided from which one can get any idea about the writing style. So this is a very challenging and practical problem to manage. This implies that the main demand of the system is to develop such a method which will be able to detect the handwritten digit images irrespective of its writing style. Apart from these, there are also such numerals in almost all languages which belong to different class but has nearly similar kind of shapes (e.g. in two pairs of Gurmukhi numerals seem almost similar, one is 1 and 7 and another one is 8 and 9, for Bangla and Assamese numerals 1 and 2 may seem similar sometimes and another pair is 3 and 6).

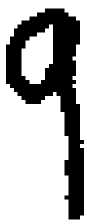
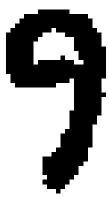
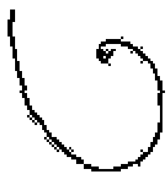
		

fig1: 1 in Gurumukhi	fig2: 7 in Gurumukhi	fig3: 3 in Assamese	fig4: 6 in Assamese

Related Work:

To deal with the above mentioned challenges, a number of effective and comprehensive techniques have been put forward by many research groups around the world. Based on the properties of the features used, the proposed methods are predominantly divided into two groups – structure based features and texture based features. Structure based features mainly focus on structural and topological properties of a numeral image either taken from the entire image shape or after subdividing the image into different sized zones or sub images. Basu et al. in their paper [1] have proposed a 76-element feature vector containing 16 centroid feature, 36 longest run feature and 24 shadow feature for recognition of handwritten Bangla numerals. Thereafter Multi-Layer Perceptron (MLP) is used for classification purpose. In another version of their work [2], Basu et al. have come up with an application of Dempster-Shafer (DS) method for combination of classification decisions obtained from two MLP classifiers using two feature vectors providing complement information. In [3] Das et al. have used a feature vector of length of 88 comprising 16 octant and 72 shadow features followed by MLP classifier to recognize Arabic numerals. In the paper [4] Das et al. have designed a novel convex hull based feature set calculated over various bays characteristics of the convex hull of a pattern, for effective recognition of isolated handwritten Bangla characters and numerals. Dongre et al. in their paper [5] have used geometric and structural features to recognize handwritten Devanagari numerals and characters. In this method, every image is divided into nine partitions. To combine both local and global effects, eight structural features are computed from each partitions and from entire image. The classification is carried out using Multi-Layer Perceptron Neural Network (MLP-NN). In [6] Lehal et al. have proposed a system to recognize Devanagari and English numerals using a set of global and local features, obtained from the right and left projection profiles of the numeral images. The primary flaw of structure based features is that it falters in case of similar structure of two different images. Besides, predominant structure based methods use local information from subdivided images which make it computationally inefficient especially in case of repetitive subdivision. On the other hand, texture based features intent to calculate the data pixel density or statistical measures from a group of pixels. The process of calculating features is divided into two types – spatial and spectral based feature extraction approaches. For the former approach, the texture features are extracted by computing the pixel statistics or finding the local pixel structures in the original image, whereas the latter transforms an image into frequency domain and then calculates feature from the transformed image. In the paper [7], Hassan et al. have introduced an approach for handwritten Bangla numeral recognition using three different variations of Local Binary Pattern (LBP) - the basic LBP, the uniform LBP and the simplified LBP followed by a K-NN classifier. In the paper [8] Karthik et al. have put forward a technique based on HOG (Histogram of Oriented Gradients) for the recognition of handwritten Kannada numerals. HOG descriptors are invariant to geometric transformation and hence they are regarded as one among the best descriptors for numeral recognition. Multi-class Support Vector Machine (SVM) is used for the classification purpose. In the paper [9], Singh et al. have suggested a novel Mojette transform (also called projection histograms features) based feature vector to recognize handwritten numerals of four major Indic scripts namely, Bangla, Devanagari, Arabic and Telugu. After that principal component analysis (PCA) is accomplished to reduce the feature dimension, and then this reduced feature vector is fed to MLP for classification of the handwritten numeral images. The main drawback of texture based feature is that it is very sensitive towards the orientation of the numeral image as the texture feature extraction is highly influenced the spatial position of a pixel. It becomes inefficient in the cases of poor handwriting, rotation of the images while scanning etc. Singh et al. in their paper [10] have proposed a 130-element feature set for efficient handwritten numeral recognition. The proposed feature descriptor is essentially a combination of six different types of moments which are geometric moment, moment invariant, affine moment invariant, Legendre moment, Zernike moment and complex moment. Ashiquzzaman et al. in their paper [11] proposed a deep learning based novel approach for recognition of Arabic numeral recognition. The key idea behind the method is to use a suitable activation function and a regularization layer in the neural network. In another work [12], Ahmed et al. have used a LSTM (Long short-term memory) architecture for Bangla handwritten numeral recognition. The suggested LSTM methodology in [12] first normalizes the images and then two-layer LSTM is employed to classify the numeral. Many alternative approaches have been performed by other researchers to deal with handwritten numeral recognition; some of those can be found in the recent survey paper [13] by Singh et al.

In the proposed method, refraction of light has been used and numerals have been recgnised by tracing different paths of light traversed due to refraction. The depths of mediums has been varied and the type of light sources too.

Present work:

Refraction of light depends on the refractive indices of two mediums, one from where the light ray is coming and another from where it is getting reflected. We are using the method of refraction to distinguish different numerals because each numeral has different slopes at different sections, so, with the help of refraction of light rays, passing through different mediums, deviation of the light rays will also differ and thus, we can recognise different numerals. This method can be applied on numerals of different languages in more or less same manner and a good result can be expected. Thus this method generalises the way of handwritten digit (numeral) recognition. Two types of light sources have been used in this method, one is a point source and second one is a parallel beam of light rays. Both are explained below.

Feature using refraction of light rays coming from a point source:

Light rays are beamed on the numeral from four vertices of the image making angle with the adjacent boundaries with an interval of 5°, making the number of light rays from each corner = 19.

Thus the $number\ of\ light\ rays = 19 \times 4 = 76$ (as light rays are beamed from all the four corners). The image is divided into two halves, the upper part is considered to be filled with air(**µ1 = 1**) and the lower half is filled with a denser medium. The refractive index of the denser medium is **µ2 = 1.5** and is of depth **d** (≤*image size*).. The numeral is also considered to be as a denser medium with refractive index **µ3 = 2**. The refraction in medium 3, i.e the medium of the numeral is considered as glass slab refraction. Thus the light rays will suffer at least one refraction and at most two depending on the values of **d** and the slope of the numeral at that slope.

There can be two types of refractions, one due to the medium 2, i.e the normal refraction and another is due to medium 3, i.e the glass slab refraction. We are considering the second refraction as glass slab due to the fact that the section of the numeral is considered to have parallel surfaces. The light rays can go through three types of processes:

1) Some light rays will undergo both the refractions. But the light rays with angle of incidence close to 85° will suffer only glass slab refraction (i.e the refraction due to medium 3). Those light rays will not suffer the first refraction, i.e. the normal one. So, for them we will calculate the horizontal deviation only and that will be the final deviation for them.

$$horizontal\ deviation = \frac{lateral\ shift}{\cos r}$$

$$final\ deviation = horizontal\ deviation$$

2) Those light rays which will encounter both of the denser mediums, after entering the second medium the light ray will suffer normal refraction and then that refracted ray will strike thenumeral (medium 3) and will again undergo through glass slab refraction. The final horizontal deviation (displacement is taken, i.e. it can be both negative and positive) is considered as the feature. For glass slab refraction in third medium,

$$lateral\ shift = width \times \frac{\sin(i-r)}{\cos r}$$

$$final\ deviation = (R \tan(i-d) \tan(r) + horizontal\ deviation)$$

$$\mu = \frac{\sin i}{\sin r}$$

where $i = angle\ of\ incidence$, $r = angle\ of\ refraction$

and $\mu = refractive\ index\ of\ the\ medium\ wrt\ that\ of\ air$.

$r' = second\ angle\ of\ refraction = slope - first\ angle\ of\ refraction$

(3)Only a few light rays, i.e one or two light rays close to 90° will suffer no refraction,i.e they will traverse just a straight line with no deviation and so,

$$final\ deviation = 0$$

To get the horizontal deviations, help of the above mentioned formulae has been taken. For each kind of refraction, first the data pixel, on which the light ray is striking has been found out and then deviation at that section is calculated taking the help of local slope and local width. Then if further refractions are possible, the method is carried out in the same manner.

Feature using refraction of a beam of parallel light rays:

A parallel beam of light rays is thrown on the numeral from the top of the image. $Number\ of\ light\ rays\ per\ beam = number\ of\ distinct\ columns\ of\ the\ image.$ The numeral here acts as denser medium with refractive index **μ2 = 1.5** and that of the background is same as that of air, i.e **μ1 = 1**. At first the data pixels where the light rays strike in first place are found out. Here also each small section of the numeral is considered as a glass slab,(i.e it is considered to have same slope or parallel surfaces for each small section) and so the horizontal deviation which we get by glass slab refraction for each ray is taken as our feature. For calculating the horizontal deviation, the local width and local slope are found out and accordingly the lateral shift and horizontal deviation (displacement is considered, i.e we can get both positive and negative values) are calculated using the following equations:

$$lateral\ shift = width \times \frac{\sin(i-r)}{\cos r}$$

$$horizontal\ deviation = \frac{lateral\ shift}{\cos r}$$

where $i = angle\ of\ incidence$, $r = angle\ of\ refraction$ and $refractive\ index\ (\mu) = \frac{\sin i}{\sin r}$.

Glass slab refraction:

We know that light traverses in straight line in a medium or through two mediums with same density. Now we need to see what happens when it travels through mediums of different densities. So, when this happens light ray bends at the boundary between two mediums. This phenomenon of bending of light ray is known as **Refraction of light**. Now when a light ray traverses through a glass slab it suffers a parallel shift or lateral shift after exiting the slab. The first angle of refraction and second angle of refraction are equal as the slab is parallel and so does the angle of incidence and angle of emergence for the same reason.

Lateral displacement is the perpendicular distance between the incident and the emergent rays when the light ray is incident on the slab obliquely.

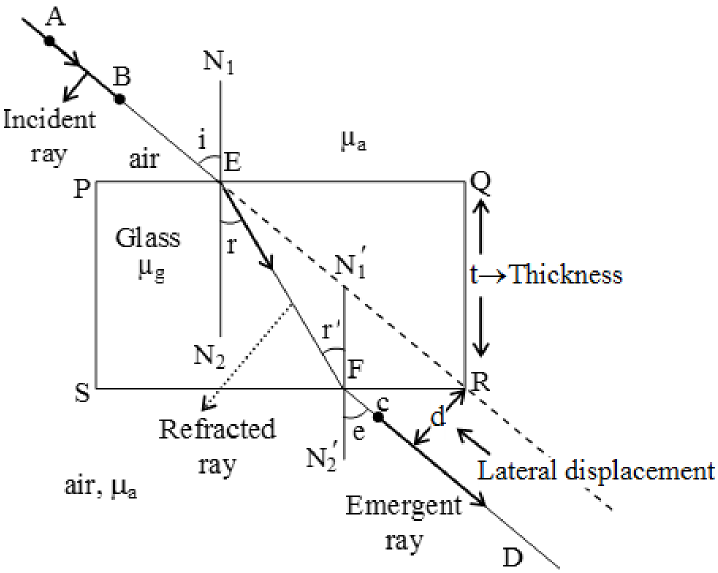
Factors on which the lateral displacement depends:

(1)Lateral displacement is directly proportional to the thickness of the slab.

(2)Lateral displacement is directly proportional to the angle of incidence.

(3)Lateral displacement is directly proportional to the refractive index of the slab.

(4)Lateral displacement is inversely proportional to the wavelength of the incident light ray.



$\mu a = refractive\ index\ of\ background\ pixel = 1$

$\mu g = refractive\ index\ of\ glass = 1.5$

The above image clearly depicts the phenomenon of **glass slab refraction of light**. And the lateral displacement as shown in the figure can be expressed as,

$$lateral\ dispalcement = width \times \frac{\sin(i-r)}{\cos r}$$

And the horizontal deviation or displacement is cosine of the lateral displacement.

Local slope and width calculation:

For both of the above mentioned methods, local slope and local width have been calculated in the following way.

For slope calculation, firstly, the data pixel, on which the light ray first strikes is found out. Then taking the horizontal axis (i.e x-axis) as main axis a straight line is rotated from 0°to 360° with keeping an interval of 10° and along with that the corresponding stroke widths are also noted. As we know that perpendicular distance is always minimum, we take the minimum stroke width as the local width and the complementary angle of corresponding angle of rotation is taken as the local slope. Angle of incidence is thus calculated from the slope with the help of the formula given below:

$angle\ of\ incidence = 90° - slope$

Result Section:

The method of recognition of digits involves refraction of light rays from two different kinds of sources, one is point source of light and another one is parallel beam of light. All parameters have been kept fixed for the second method , including the length of feature vector and input image size. But, for the first one, two lengths of feature vector have been taken and input image size is varied for the both to get optimal length of feature vector corresponding to its input image size. For classification in both cases, each dataset is divided into two parts, viz. train set and test set, the ratios being 7:3, 4:1 and 3:1.

Square images of two sizes are taken as input, one with length = 32 and another with 64. For both cases, the another parameters which are involved in the method involving refraction of light rays from point source, are depth of the air medium, d = 2/3 times of the length of the input image, the refractive index of medium upto d = 1, the refractive index of medium placed below that = 2 and the refractive index of the medium that forms the digit = 3. This test has been conducted on online dataset of Assamese and Devnagari and offline dataset of CMATER of Hindi, Bengali, Arabic and Telegu. The accuracies, obtained by conducting this test, has been recorded in table and is shown in fig. - 1 and fig. - 2 given below.

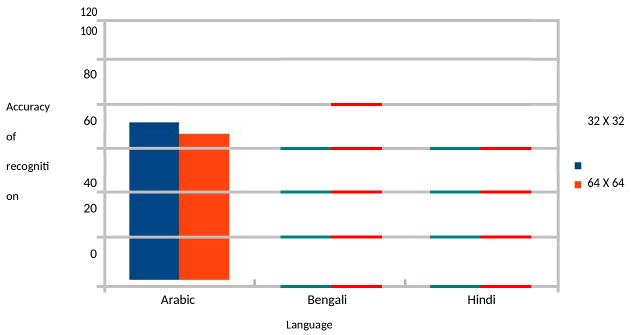


Fig.1 : Accuracy of recognition v/s Language bar chart for input images of two sizes with feature vector length = 76

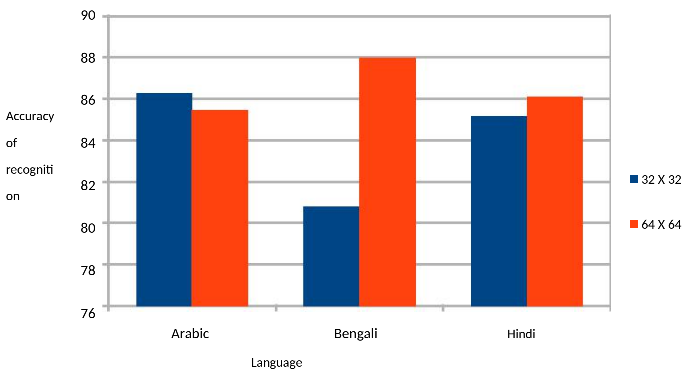


Fig.2 : Accuracy of recognition v/s Language bar chart for input images of two sizes with feature vector length = 323

The length of feature vectors are taken as 76 and 323 for one method and 768 for the second one. The length of feature vector has been varied based on number of sorces of light placed in the image. The chart given below shows the variation in accuracy of recognition(%) due to change in length of feature vector.

In fig.1, the input image size of 32 X 32 is giving better result, but in fig.2 64 X 64 input image size is giving better result and that too surpasses the highest result for each language by 10% – 15% acuuracy of recognition. So, based on both the figures, the optimal feature vector length is 323 with optimal input image size being 64 X 64.

Now, the optimal feature vector has been merged with the second feature vector to make a unique feature vector, whose length is 1172 and then, Random Forests are applied on the feature vectors.Applying this, the obtained accuracy is charted below in fig.3.

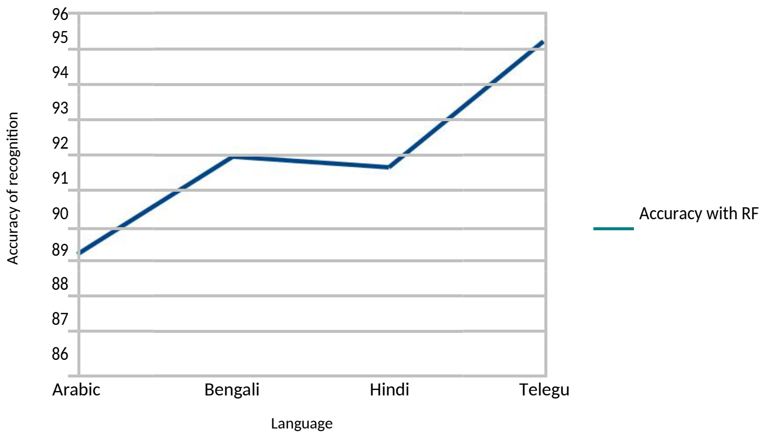


Fig.3 – Line Chart to show accuracy of recognition v/s language after applying RF

Random Forest Algorithm:

Random Forests are an ensemble learning method for classification and regression and other tasks. Random forest algorithm is a supervised classification algorithm. As the name suggest, this algorithm creates the forest with a number of trees.In general, the more trees in the forest the more robust the forest looks like. In the same way in the random forest classifier, the higher the number of trees in the forest gives the high accuracy results. Random Forest classifire can handle the missing values and it will not overfit when more trees are present in the forest. This algorithm is widely used in field of Medicine, Banking, Stock Market, E – commerce and others.

Reference:

(1) Basu, Subhadip, Nibaran Das, Ram Sarkar, Mahantapas Kundu, Mita Nasipuri, and Dipak Kumar Basu. "An MLP based Approach for Recognition of Handwritten Bangla 'Numerals." arXiv preprint arXiv:1203.0876 (2012).

(2) Basu S., Sarkar R., Das N., Kundu M., Nasipuri M., Basu D.K. (2005) Handwritten Bangla Digit Recognition Using Classifier Combination Through DS Technique. In: Pal S.K., Bandyopadhyay S., Biswas S. (eds) Pattern Recognition and Machine Intelligence. PReMI 2005. Lecture Notes in Computer Science, vol 3776. Springer, Berlin, Heidelberg.

(3) Das, Nibaran, Ayatullah Faruk Mollah, Sudip Saha, and Syed Sahidul Haque. "Handwritten arabic numeral recognition using a multi layer perceptron." arXiv preprint arXiv:1003.1891 (2010).

(4) Das, Nibaran, Sandip Pramanik, Subhadip Basu, Punam Kumar Saha, Ram Sarkar, Mahantapas Kundu, and Mita Nasipuri. "Recognition of handwritten Bangla basic characters and digits using convex hull based feature set." arXiv preprint arXiv:1410.0478 (2014).

(5) Dongre, Vikas J., and Vijay H. Mankar. "Devanagari offline handwritten numeral and character recognition using multiple features and neural network classifier." In Computing for Sustainable Global Development (INDIACom), 2015 2nd International Conference on, pp. 425-431. IEEE, 2015.

(6) Lehal, G. S., and Nivedan Bhatt. "A recognition system for Devnagri and English handwritten numerals." In Advances in Multimodal Interfaces—ICMI 2000, pp. 442-449. Springer, Berlin, Heidelberg, 2000.

(7) Hassan, Tasnuva, and Haider Adnan Khan. "Handwritten bangla numeral recognition using local binary pattern." In Electrical Engineering and Information Communication Technology (ICEEICT), 2015 International Conference on, pp. 1-4. IEEE, 2015.

(8) Karthik, S., and K. Srikanta Murthy. "Handwritten Kannada Numerals Recognition Using Histogram of Oriented Gradient Descriptors and Support Vector Machines." In Emerging ICT for Bridging the Future-Proceedings of the 49th Annual Convention of the Computer Society of India CSI Volume 2, pp. 51-57. Springer, Cham, 2015.

(9) Singh, Pawan Kumar, Supratim Das, Ram Sarkar, and Mita Nasipuri. "Recognition of handwritten indic script numerals using mojette transform." In Proceedings of the First International Conference on Intelligent Computing and Communication, pp. 459-466. Springer, Singapore, 2017.

(10) Singh, Pawan Kumar, Ram Sarkar, and Mita Nasipuri. "A study of moment based features on handwritten digit recognition." Applied Computational Intelligence and Soft Computing 2016 (2016): 4.

(11) Ashiquzzaman, Akm, and Abdul Kawsar Tushar. "Handwritten Arabic numeral recognition using deep learning neural networks." arXiv preprint arXiv:1702.04663 (2017).

(12) Ahmed, Mahtab, M. A. H. Akhand, and MM Hafizur Rahman. "Handwritten Bangla Numeral Recognition Using Deep Long Short Term Memory." In Information and Communication Technology for The Muslim World (ICT4M), 2016 6th International Conference on, pp. 310-315. IEEE, 2016.

(13) Singh, Pawan Kumar, Ram Sarkar, and Mita Nasipuri. "A comprehensive survey on Bangla handwritten numeral recognition." International Journal of Applied Pattern Recognition 5, no. 1 (2018): 55-71.

(14) P.M. Patil∗, T.R. Sontakke. “Rotation, scale and translation invariant handwritten Devanagari numeral character recognition using general fuzzy neural network”

(15) Y. Le Cun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel. “Handwritten Digit Recognition with a Back-Propagation Network “ AT&T Bell Laboratories, Holmdel, N. J. 07733

(16) Basu S., Sarkar R., Das N., Kundu M., Nasipuri M., Basu D.K. (2005) Handwritten Bangla Digit Recognition Using Classifier Combination Through DS Technique. In: Pal S.K., Bandyopadhyay S., Biswas S. (eds) Pattern Recognition and Machine Intelligence. PReMI 2005. Lecture Notes in Computer Science, vol 3776. Springer, Berlin, Heidelberg.

(17) Blumenstein and B. Verma. Neural–based solutions for the segmentation and recognition of difficult handwritten words from a benchmark database. In Proc. 5th International Conference on Document Analysis and Recognition, pages 281–284, Bangalore, India, 1999.

(18) J. Dong, A. Krzyzak, and C. Y. Suen. Local learning framework for recognition of lowercase handwritten characters. In Proc. International Workshop on Machine Learning and Data Mining in Pattern Recognition, pag to appear, Leipizig, Germany, 2001.

(19) P. D. Gader, M. A. Mohamed, and J. H. Chiang. Handwritten word recognition with character and intercharacter neural networks. IEEE Transactions on Systems, Man and Cybernetics – Part B, 27:158–164, 1994.

(20) Siddhartha Banerjee*, Bibek Ranjan Ghosh, Arka Kundu. “Handwritten Character Recognition from Bank Cheque”. International Journal of Computer Sciences and Engineering.

(21) Berrin A. Yanikoglu, Aytac Gogus, Emre Inal. “Use of handwriting recognition technologies in tablet-based learning modules for first grade education”. Article in “Educational Technology Research and Development”- July, 2017.

Result:

Language	Accuracy	No. of features	Image size
Light rays diveriging from a point source			
Hindi	67.03	76	32*32
Bengali	71.02	76	32*32
Arabic	71.57	76	32*32
Hindi	86.1	323	64*64
Bengali	87.98	323	64*64
Arabic	85.43	323	64*64
Telegu	91.67	323	64*64
Gurumukhi	87.5	323	64*64
Light rays coming in parallel			
Hindi	91	768	64*64
Arabic	87.23	768	64*64
Bengali	90.86	768	64*64
Telegu	93.1	768	64*64
Gurumukhi	81.2	768	64*64
Assamese	92.7	768	64*64
Merge			
Hindi	94.37	1172	
Bengali	95.45	1172	
Arabic	94.17	1172	
Telegu	97.33	1172	
Gurumukhi	92.5	1172	
Asamese	99.73	1172	