KANNADA HANDWRITTEN CHARACTER RECOGNITION USING MULTI FEATURE EXTRACTION TECNHIQUES

Article *in* International Journal of Science and Research (IJSR) · October 2014

Aravinda C V
NMAM Institute of Technology

KANNADA HANDWRITTEN CHARACTER RECOGNITION USING MULTI FEATURE EXTRACTION TECNHIQUES

Aravinda.C.V*[1], Dr.H.N.Prakash†[2][2]

[1] Asst. Professor Department of Information Science & Engineering, S.J.B.I.T. B.G.S Health & Education City, Kengeri, Bengaluru ,Karnataka.
[2]Head Department of Computer Science & Engineering, Rajeev Institute of Technology, Hassan, Karnataka.

Abstract

In this paper, we have presented a method of feature extraction for handwritten character recognition. Handwritten character recognition is a complex task because of various writing styles of different individuals.Our Method yeilds good classification accuracy on handwritten characters, apart from complexity. Normalization and binarization are the pre-processing techniques used for getting accurate results of classification process in handwritten character recognition. To select a set of features is an important step for implementing a handwriting recognition system. In this work, we have extracted various features, namely-Hu's Invariant moments, Zernike moments, Zonal features, Fouier-Wavelet cofficients. The recognition process is carried out using Back Propagation Neural Network.

Keywords :- Character Segmentation, Optical Character Recognition, Hand Written Character Recognition, Optical Hand Character Recognition, Neural Network Classifier.

1 Introduction

Handwritten Character Recognition is an Optical Character Recognition problem for handwritten characters. It is very valuable in terms of the variety of applications and also as an academically challenging problem. When HCR is used as a solution for inputting regional language data and also as a solution for converting paper information to soft form, HCR solutions become powerful component in addressing the digital divide. It also provides a solution to processing large volumes of data automatically. Hence extensive work is happening in this field on different scripts. But on Indian language scripts, very little work is reported. Indian scripts share a large number of structural features due to common Brahmi origin. The written form has more curves than straight or slant lines and has lots of similarities between different alphabets of the same scripts and also between the scripts of different languages. Here Kannada script is used for experiments. Handling Kagunita, the Kannada language script is the hardest part in Indian language HCR, since most vowels when following a consonant modifies the consonants shape. The modifications are normally not drastic, making it hard to spot the vowel. On the other side, the modification reduces consonant recognition due to the resulting distortion in the shape.

1.1 Kannada Language Script

Kannada, the official language of the South Indian state of Karnataka, is spoken by about 48 million people. The Kannada alphabets were developed from the Kadamba and Chalaukya scripts, descendants of Brahmi which were used between

*aravinda.cv@gmail.com
†prakash_hn@yahoo.com

the 5th and 7th century A.D. The basic structure of Kannada script is distinctly different from Roman script. Unlike many North Indian languages, Kannada characters do not have shirorekha (a line that connects all the characters of any word) and hence all the characters in a word are isolated. This creates difficulty in word segmentation. Kannada script is more complicated than English due to the presence of compound characters. However, the concept of upper/lower case characters is absent in this script. Modern Kannada has 50 base characters, called as Varnamale as shown in Figure.1 comprising 13 vowels, 2 yogawahakas. There are five groups of consonants each with five letters, a total of twenty five letters called as grouped consonants or structured consonants. When a consonant character is used alone it results in a dead consonant (mula vyanjana). Vowel modifiers can appear to the right, on the top or at the bottom of the base consonant. Figure 2 shows a consonant modified by all the 15 vowels. Such consonant-vowel combinations are called live consonants (gunithakshara) as shown in Figure 3.

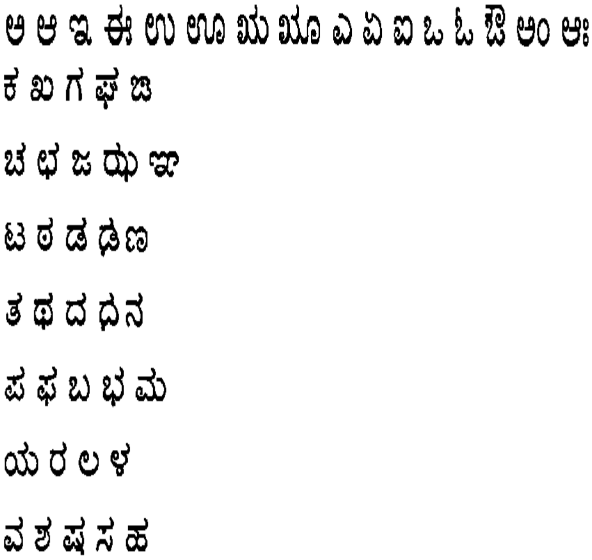

Figure 1: Varnamale

2 RELATED WORK

Various works done on English character recognition was studies. It was found that the results were nearing 100%. Kannada scripts have less research done on the character recognition due to their structures which differs from person to person. The accuracy found out for Kagunita, Kannada script is about 40-50%.

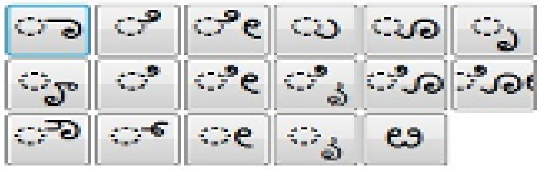

Figure 2: Varnamale

Figure 3: Varnamale

3 PROBLEM STATEMENT

Handling Kagunita is the hardest part in Indian language HCR, since most vowels when following a consonant modifies the consonant's shape. The modifications are normally not drastic, making it hard to spot the vowel. On the other side, the modification reduces consonant recognition due to the resulting distortion in the shape. The distortions are not drastic and hence results in many similar shaped characters. The Table 1.1 shows few of the characters in the Kannada script with similar shape. The characters in the same row have similar shape with minute differences between the characters which make them difficult to recognize.

Table 1.2

ಅ ಆ	ಇ ಈ
ಐ ಣ ಟ ಚ	ಖ ಬ ಭ ಒ
ದ ಡ ಪ ಹ ವ ಏ	ಥ ಧ ಢ ಫ ಷ ಘ ಛ
ಠ ರ ಗ	ನ ಸ
ಳ ಶ	ಯ ಋ ಮ
ಕ ತ	ಙ ಜ ಓ
ಅಂ ಅಃ	

4 STAGES INVOLVED

The system is intended to isolate each character from an input image containing handwritten Kannada text, recognize the character and also print its corresponding Unicode onto a text file which can be edited and saved for future use. This system mainly breaks down the recognition process into four fundamental sequential stages:

4.1 PREPROCESSING TECHNIQUES

Preprocessing generally consist of series of image-to-image transformations. It is preliminary step which transforms the data into a form that is more easily and effectively processed. The main task of preprocessing is to process the scanned image and increase the noise that causes a reduction in the recognition rate and increases the complexities. Hence preprocessing is an essential stage prior to the segmentation, as it controls the suitability of the results for the success of the recognition. Preprocessing is divided into the following sub-modules:

1. Conversion to grayscale
2. Noise removal
3. Conversion to binary

4.2 SEGMENTATION

Segmentation is one of the most important phases of HCR system. By applying good segmentation techniques the performance of HCR can be increased.It is the process of extracting objects of interest from an image. It subdivides an image into its constituent regions or objects, which are certainly characters. This is need because the classifier recognizes only the isolated characters. Segmentation phase is also crucial in contributing to the error due to touching characters, which the classifier cannot properly tackle. Even in good quality documents, some adjacent characters touch each other due to inappropriate scanning resolution. Segmentation technique is divided into the following submodules:

1. Line Segmentation
2. Letter Segmentation
3. Boundary Detection

4.2.1 PROBLEMS IN SEGMENTATION

Problem 1
A problem that occurs with the above segmentation is that in handwritten text, there high chances of slant lines which can lead to an incorrect line cropping. Figure 4 shows the problem of slant line in an image.

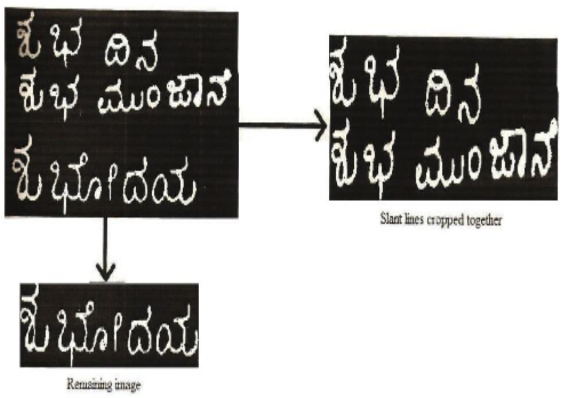

Figure 4: Slant Line Problem

Solution
The slant line problem can be overcome by sending such images into the line crop function once again after the letter crop. By using this flow, the slant lines in the incorrectly segmented lines are divided vertically by the letter crop making the

lines shorter and straight, as a result of which the characters get segmented separately in the second line crop.

Probem 2
Another issue is in case the characters are written too close such that there is no enough empty space between the characters to be detected by the letter crop function. In this case the characters are not segmented correctly which leads to incorrect recognition.
Solution
The boundary determination is used to overcome this error. Here assumption is made that all slant line problem is overcome and only an image consisting of images in a single line is passed into the boundary detection. **Boundary Detection** In this stage the connected objects are given a label and a rectangular box is plotted around each connected object. The value of each label is extracted and each rectangular box is cropped to get the isolated character. This isolated character is then sent for feature extraction.Figure 5 shows the problem of close images being solved by the boundary detection.

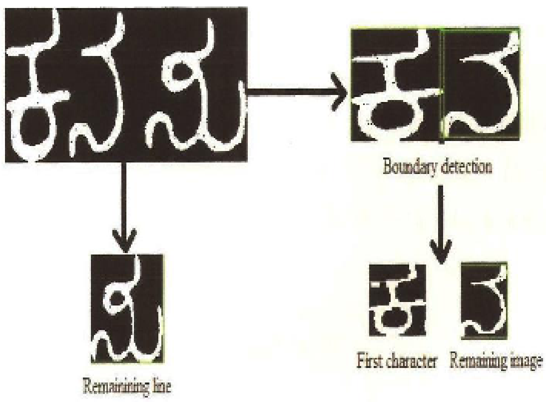

Figure 5: Boundary Detection

4.2.2 FEATURE EXTRACTION

Feature extraction is a crucial step for character recognition, and most research has been devoted to finding measures that concisely represent a pattern and at the same time contain enough information to ensure reliable recognition. Feature extraction is the problem of extracting from raw data, the information which is most relevant for classification purposes, in the sense of minimizing the within-class pattern variability while enhancing the between-class pattern variability. In general, good features must satisfy the following requirements: First, intra-class variance must be small. Secondly, the interclass separation should be large. In order to recognize many variations of the same character, features that are invariant top certain transformations on the character need to be used. Various feature extraction methods employed for recognizing the normalized segmented characters are discussed.

Feature Set I: Hu's Invariant Moments
Feature Set II: Zernike Moments
Feature Set III: Zonal Features
Feature Set IV: Fourier-Wavelet Co-efficients
The feature extraction module is as shown in Figure 6

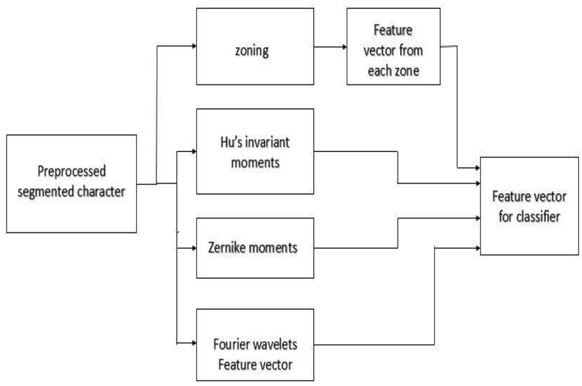

Figure 6: Flow of Feature Extraction Module

Feature Set I: Hu's Invariant Moments
Hu introduced the use of moment invariants as features for pattern recognition. The general form of a regular moment function mpq of order(p+q)of an image intensity function f(x,y) is defined in equation below. For a digital image, the central moments, which are invariant to translation is defined in equation of mu_{pq} , where $x_c = m_{01}/m_{00}$ are the co-ordinates of centroid.

The seven non linear functions defined on regular moments which are invariant to rotation, scaling and translation are given in below equation

$$\phi 1 = \eta_{20} + \eta_{02}$$

$$\phi 2 = \left(\eta_{20} + \eta_{02}\right)^2 + 4\eta_{11}^2$$

$$\phi 3 = \left(\eta_{30} + \eta_{12}\right)^2 + \left(3\eta_{21} + \eta_{03}\right)^2$$

$$\phi 4 = \left(\eta_{30} + \eta_{12}\right)^2 + \left(\eta_{21} + \eta_{03}\right)^2$$

$$\phi 5 = \left(\eta_{30} - 3\eta_{12}\right)\left(\eta_{30} - \eta_{12}\right)\left[\left(\eta_{30} + 3\eta_{p12}\right)^2 - 3\left(\eta_{21} + \eta_{03}\right)^2\right] + \left(3\eta_{21} - \eta_{03}\right)\left(\eta_{21} + \eta_{03}\right)\left[3\left(\eta_{30} + \eta_{12}\right)^2 - \left(\eta_{21} + \eta_{03}\right)^2\right]$$

$$\phi 7 = \left(3\eta_{12} - \eta_{30}\right)\left(\eta_{30} + \eta_{12}\right)\left[\left(\eta_{30} + \eta_{12}\right)^2 - 3\left(\eta_{21} + \eta_{03}\right)^2\right] + \left(3\eta_{21} - \eta_{03}\right)\left(\eta_{21} + \eta_{03}\right)\left[\left(3\eta_{30} + \eta_{12}\right)^2 - \left(\eta_{21} + \eta_{03}\right)^2\right]$$

$$\phi 6 = \left(\eta_{20} - \eta_{02}\right)\left[\left(\eta_{30} + \eta_{12}\right)^2 - \left(\eta_{21} + \eta_{03}\right)^2\right] + \left(\eta_{30} + \eta_{12}\right)\left(\eta_{21} + \eta_{03}\right)$$

where $\eta_{pq}$is the central moment given by $\eta_{pq} = \sum_{X=0}^{M-1} \sum_{Y=0}^{N-1} x^p y^q f\left(x, y\right)$ for $p, q = 0, 1.....\infty$. The definition of regular moments has the form of projection of $f\left(x, y\right)$ function onto the basis set $x^p y^q$ The seven moments $\phi 1$ to $\phi 7$are calculated for the pre-processed character and considered as feature vector for recognition. Since the basis set $x^p y^q$ is not orthogonal, the recovery of image from these moments is difficult and has a drawback of information redundancy.

Feature Set II: Zernike Moments

Zernike moments are due to Zernike polynomials introduced by Zernike and have minimum information redundancy. Zernike polynomials are a set of complex polynomials $V_{nm}\left(x, y\right)$ defined in equation below where $n \geq 0, \mid m \mid \leq n$ and $n- \mid m \mid$ is even , r is length of the vector from origin to (x, y) and ϕ is the angle between r vector and the x-axis in the counter clockwise direction. Equation of $R_{nm}\left(r\right)$defines orthogonal radial polynomial R_{nm} which form a complete orthogonal set over interior of the unit circle. The Zernike moments are projections of the input image onto the space spanned by orthogonal V functions. For a digital image, the Zernike moment of order n and repetition m is defined in equation of $Z_{nm.}$ The Zernike moments of a given image is computed by considering the centre of the image as origin and the given image is mapped to the range of the unit circle $x^2 + y^2 \leq 11$ and only those pixels which fall inside the unit circle are considered for moment computation. If the image is rotated by an angle α, the transformed Zernike moment function $Z'_{nm} = Z_{nm} e^{jm\theta}$. The magnitude of the moments$\mid Z_{nm} \mid$ remains the same after rotation. Hence the amplitude of Zernike moments are used as features. Although Zernike moments are invariant to rotation, roust to noise and have minimum information redundancy, they have the drawback of computational complexity. $V_{nm}\left(x, y\right) = V_{nm}\left(r, \theta\right) = R_{nm}\left(r\right).e^{jm\theta}$

$$R_{nm}\left(r\right) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s\left(x^2+y^2\right)\left(\frac{n}{2}-s\right)(n-s)}{s\left(\frac{n-|m|}{2}-S\right)\left(\frac{n-|m|}{2}-S\right)}$$

Feature Set III: Zoning

In this method the image is split into different zones and simple features are extracted from each of the zones.In this method, the segmented character is first area normalized so that the number of ON pixels in all the normalized characters are equal. With this normalized technique the classifier is immune to size changes in the characters. The normalized character is divided into smaller zones. Various regional features such as minor axis length, major axis length, centroid, eccentricity, convex area are calculated. Along with the regional features, structural features such as geometric moments, variance are calculated and used as feature vector.

Feature Set IV: Fourier- Wavelet Coefficients

Fourier transform is a powerful tool for pattern recognition. Fourier transform is translation and rotation invariant, but the frequency information of Fourier transform is global and so a local variation of the shape will affect the Fourier coefficients. Wavelet transform has multi-resolution ability but is translation variant. A small shift of the original signal will lend totally different wavelet coefficients. Therefore Fourier and Wavelet transforms are combined to obtain a feature vector which is not only invariant to translation and rotation, but also has multi- resolution ability.

Given an NxN pattern $f\left(x, y\right)$, translation invariance is achieved by translating the origin of the coordinate system to the center of the mass of the pattern, denoted by $\left(x_0, y_0\right)$. The scale invariance is obtained by transforming the pattern in Cartesian coordinate to polar coordinate system. The longest distance from the center$\left(x_0, y_0\right)$.to a point(x, y)on the pattern is given by Equation below. $d = max\sqrt{\left(x - x_0\right)^2 + \left(y - y_0\right)^2}$N concentric circles are drawn centered at(x_0, y_0) with radius r as given in Equation $r = d * \frac{t}{N}, i = 1, 2....N$ N angularly equi-spaced radial vectors$\theta_i$are formed departing from(x_0, y_0) with angular step $2\pi/N$. For any small region The average value of $f(x, y)$ is calculated over this region, and the average is assigned to $g(r, \theta)$ in the polar co-ordinate system.1-D Fourier transform then applied along the axis of the polar angle$\theta$of $g(r, \theta)$ to obtain its spectrum. Since the Fourier transform of circularly shifted signals do not change, a feature is obtained which is rotation invariant. As wavelet coefficients represent pattern features at different resolution levels, Wavelet transform is applied along the radius of the resulting $G(r, \phi)$ which yields$WT_r(G(r, \phi))$ so that features can be obtained at different resolution levels.

5 Classification Using NN Classifier

The feature vector extracted from the segmented character is assigned a label using a classifier. Recognition of segmented characters is performed using NN Classifier.The recognition performance of Back Propagation network will highly depend on the structure of the network and training algorithm. Feed forward back propagation neural network has been selected to train the network. The number of nodes in input, hidden and output layers will determine the network structure. All the neurons of one layer are fully interconnected with all neurons of its just preceding and just succeeding layers (if any)

5.1 Back Propagation Neural Network Algorithm

1. Initialize the weights to small random values.
2. Randomly choose an input pattern $x(^{\mu})$
3. Propagate the signal forward through the network.
4. Compute δ_i^L in the output layer $\left(o_l = y_l^L\right)$ $\delta_l^L = g'\left(h_l^L\right)\left[d_l^L - y_l^L\right]$,whereh_l^Lrepresents the net input to the ith unit in lth layer and $g'$is the derivative of activation function g.
5. Compute the deltas for the preceding by propagating the error backwards.$\delta_i^L = g'\left(h_i^L\right)\sum_j w_{ij}^{l+1}\delta_j^{l+1}, for I = L-1......1$
6. Update weights using $w_{ij}^l = \eta\delta_i^L y_j^{i-1}$
7. Go to step 2 and repeat for the next pattern until the error in the output layer is below a pre-specified threshold or maximum number of iterations is reached.

5.2 Conversion to Editable Format

Based on index value (for template based matching), the Unicode corresponding to the character stored at the obtained index value is stored into a variable letter. The letter value is stored into a word array. The recognized characters are printed on to a notepad which can be further edited and saved.

6 Conclusion and Future Enhancement

HCR is the process of identifying the handwritten characters. The text in an image is converted into other letter codes which are usable within computer and text processing applications.Here recognition is done using NN classifier.It attempts to increase overall efficiency and accuracy of the HCR. Various feature extraction techniques are incorporated to improve the efficiency. Also the image is converted into an editable format. The editable text can be saved and opened for further editing.The current system can be combined with other features to improve the efficiency. An overall architecture for HCR incorporating all these features can be developed to improve the accuracy. Such a structure will help to exploit further domain information in the recognition process. The current system can be extended to recognize votaksharas.

7 Results Obtained

The developed technique is tested with multiple images of Kannada Handwritten Text where the image undergoes the testing of each individual module. Once the testing is complete, the recognized characters are printed onto a text pad and the editable text file is checked to find the recognition rate

Table1:-Results Obtained

Feature Set	Obtained Recogn-tion Accuracy for Vowels in %	Obtained Recogn-tion Accuracy for Consonants in %
Hu's Invariant Moments	35-40	30-40
Zernike Moments	40	30-35
Zonal Features	45-55	40
Fourier Wavelet Features	55	45-50
Hu's Invariant & Zernike Moments	40-45	30-35
Hu's Invariant & Zonal Features	50	40
Hu's Invariant & Fourier Wavelet	50	40-45
Zernike & Zonal Features	60-65	40-45
Zernike & Fourier Wavelet	60	30-40
Zonal & Fourier Wavelet	75	55-60
Hu's Invariant , Zonal, Zernike & Fourier Wavelet	85-90	80-85

```
Standard Approach - Features extraction for character recognition... please wait.

feature_vector =

  Columns 1 through 10

    0.1903    0.0000    0.0000    0.0000   -0.0000   -0.0000   -0.0000    0.0380   22.8750    0.9215

  Columns 11 through 20

    0.0889  -45.0000    0.3822    0.0889    0.0896    0.0348   34.2500    0.8183    0.2067   25.3102

  Columns 21 through 30

    0.3511    0.1556    0.0643    0.0369   26.1250    0.6728    0.1022   53.2717    0.4933    0.1022

  Columns 31 through 40

    0.0759    0.0562   15.7500    0.9954    0.2500   74.9261    0.1156    0.0667    0.0797    0.0077

  Columns 41 through 50

   30.1250    0.8481    0.2500  -26.1068    0.2489    0.0889    0.0529    0.0280   20.1250    0.9667

  Columns 51 through 60

    0.1407  -63.0811    0.2711    0.0844    0.0902    0.0231   24.8750    0.5121    0.0978   44.2370
```

Figure 7: Feature Vector

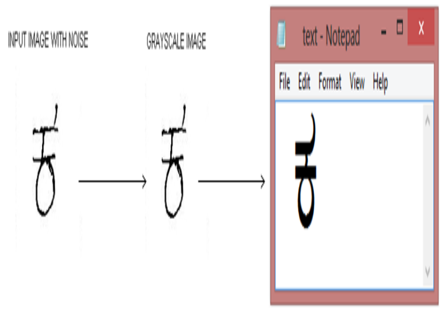

Figure 8: Output

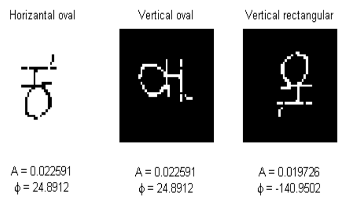

Figure 9: Zernike Moment

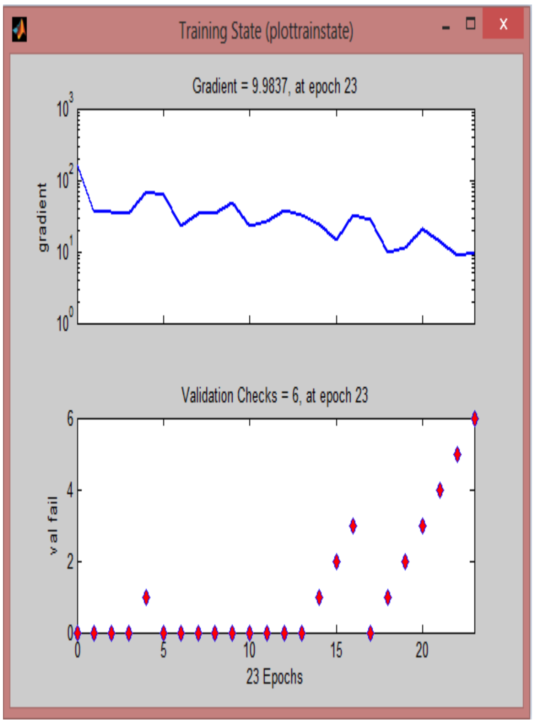

Figure 11: Trainingstate

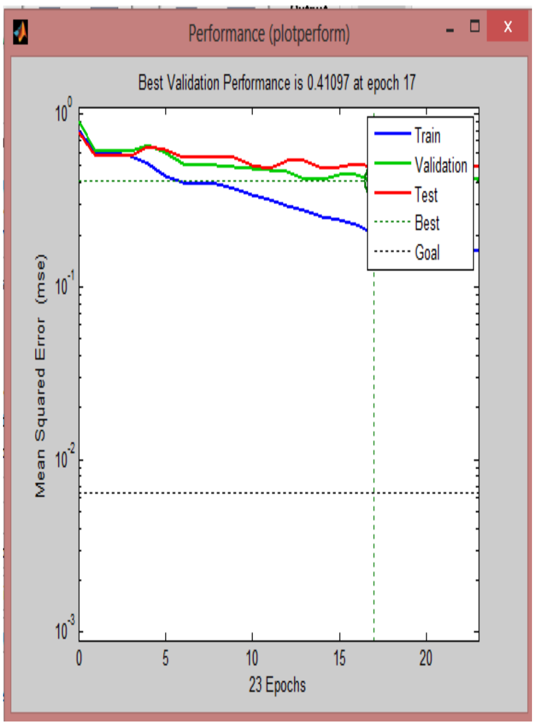

Figure 10: Performance Plot

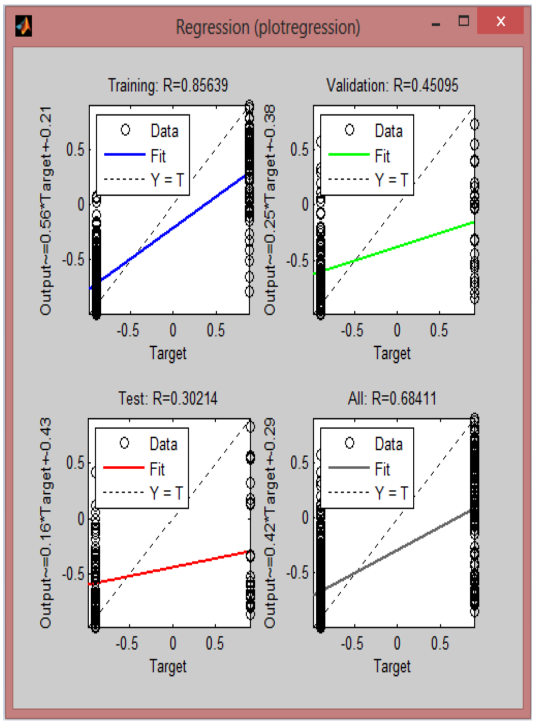

Figure 12: Regression

References

[1] Bhardwaj Anurag, Jose Damien and Govindaraju Venu. 2008. ?Script Independent Word Spotting in Multilingual Documents.? In the Second International workshop on ? Cross Lingual Information Access-2008?.

[2] Chung Yuk Ying, Wong Man To and Bennamoun Mohammed, 1998,Handwritten Character Recognition by contour sequence moments and Neural Network, In International Conference on Systems, Man, and Cybernetics, Volume 5, Issue , 11-14 Oct 1998 page(s): 4184 - 4188

[3] Liao Simon X and Lu Qin, 1997, ?A study of moment functions and its use in Chinese Character Recognition?, In Proceedings of the 4th International conference on Document Analysis and Recognition (ICDAR-97) page(s): 572-575

[4] Kunte R Sanjeev and Samuel Sudhaker R D. December 2007. ?A Bilingual Machine-Interface OCR for Printed Kannada and English Text Employing Wavelet Features.? In Tenth IEEE International conference on information technology. page(s): 202-207.

[5] Pal U; Chaudhuri B B, ?Indian Script Character Recognition: A Survey?, In Elsevier Ltd. Pattern Recognition Society- 2004, page(s): 1887 ? 1899,

[6] Ragha Leena and M Sasikumar. October 07. Fusion of Geometric, Positional and Moments features for Handwritten Character Recognition using Neural Network, In National Conference on Computer Vision, AI and Robotics ? (NCCVAIR - 07) at Chennai. India. page(s): 74-80.

[7] Ragha Leena and M Sasikumar. ?Adapting Moments for Handwritten Kannada Kagunita Recognition? In International Conference on Machine Learning and Computing (ICMLC-2010) at Bangalore. India. page(s): 125-129.

[8] Mallat S G, 1998, ?Multifrequency Channel Decompositions of Images and Wavelet Models,? In IEEE Transactions on Image Processing, 37(12), page(s): 2091?2110.

[9] V. N. Manjunath Aradhya, G. Hemanth Kumar and S.Noushath, Robust Unconstrained Handwritten Digit Recognition Using Radon Transform, Proceedings of IEEE- ICSCN 2007, pp-626-629, (2007).