ISSN: 0976-9102(ONLINE)
DOI: 10.21917/ijivp.2014.0111

ICTACT JOURNAL ON IMAGE AND VIDEO PROCESSING, FEBRUARY 2014, VOLUME: 04, ISSUE: 03

WATERSHED ALGORITHM BASED SEGMENTATION FOR HANDWRITTEN TEXT IDENTIFICATION

P. Mathivanan[1], B. Ganesamoorthy[2] and P. Maran[3]

[1]*Department of Electronics and Communication Engineering, Velammal Engineering College, India*
E-mail: mathiece05@gmail.com
[2]*Department of Electronics and Communication Engineering, Adhiparasakthi Engineering College, India*
E-mail: bganesamoorthy@gmail.com
[3]*Department of Electronics and Communication Engineering, SSN College of Engineering, India*
E-mail: maranece07@gmail.com

Abstract

In this paper we develop a system for writer identification which involves four processing steps like preprocessing, segmentation, feature extraction and writer identification using neural network. In the preprocessing phase the handwritten text is subjected to slant removal process for segmentation and feature extraction. After this step the text image enters into the process of noise removal and gray level conversion. The preprocessed image is further segmented by using morphological watershed algorithm, where the text lines are segmented into single words and then into single letters. The segmented image is feature extracted by Daubechies'5/3 integer wavelet transform to reduce training complexity [1, 6]. This process is lossless and reversible [10], [14]. These extracted features are given as input to our neural network for writer identification process and a target image is selected for each training process in the 2-layer neural network. With the several trained output data obtained from different target help in text identification. It is a multilingual text analysis which provides simple and efficient text segmentation.

Keywords:

Slant Correction, Morphological Watershed Algorithm, Daubechies'5/3 Integer-to-Integer Wavelet Transform, Neural Network

1. INTRODUCTION

Handwriting text image analysis is usually involved in security purposes of expensive and ancient historical documents. In our practice, handwriting text image analysis involves preprocessing, segmentation, feature extraction and writer identification. This handwritten image analysis is used to analyze a document of various size and texture. It is a multilingual analysis, text analysis are performed based on two important concepts they are text independent and text dependent methods. In text independent methods more than one writer has been considered here we consider individual character as words. But in text dependent the image is statistically computed, a writer has to write the identical text to perform identification.

In our handwritten text analysis we use both text independent and text dependent method due to its flexibility. Here the key process is to extract the characteristic from the handwriting image quickly and efficiently by using slants removal and watershed segmentation [1, 2].

In this process of handwritten text analysis previously several other techniques are employed like OCR (optical characteristic recognition) [2], GLCM (Gray Level Co-occurrence Matrix) and GLRL (Gray Level Run Length matrix) and so on. OCR is a multilingual analysis in this process usually we convert our text image into electronics form. In some case we do texture classification of handwritten document by using GLCM and GLRL matrix [5, 7]. A lot of segmentation is done in the past on handwritten text images. The various existing methods for segmentation are categorized as projection based, Hough transform based, smearing, grouping, graph based, CTM (Cut text Minimum) approach, block covering and linear programming, but in our paper we used watershed for segmentation, ITI WT for feature extraction and neural tool for training the given text image [1, 18].

Our paper is organized as follows. Section 2 discusses the preprocessing stage and in the section 3 segmentation process is proposed. In section 4 feature extraction using Daubechies biorthogonal 5/3 wavelet transform is proposed. In the last section neural classifier is discussed.

2. PROPOSED METHOD

In the first stage input image is taken from a stored file or photographed one. In the next Stage preprocessing for noise removal, slant correction, binarization and normalization has been done. In the third stage segmentation process has been done by means of watershed algorithm. In the fourth stage feature extraction of segmented image has been done by reversible integer to integer wavelet transform. In the last stage classification is done by neural classifier. Fig.1 shows the algorithm for handwritten text analysis.

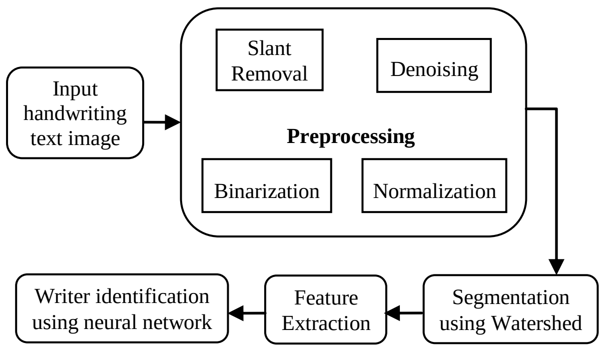

Fig.1. Block diagram for handwritten text analysis

2.1 PREPROCESSING

Preprocessing is the process of converting input image into more suitable form for the future process of any image analysis. In the preprocessing stage we perform some important process like noise removal, slant removal, grayscale conversion, binarization and normalization. Fig.1 shows the various steps involved in the preprocessing [1].

2.1.1 Noise Removal:

In all photographed or scanned document most common noise is the impulse noise. It is created due to the lens vibration and other disturbances during scanning or photographing. In our input image we have a common noise called salt and pepper or impulse noise which affects our handwritten analysis at a great level for further process of analysis. So in order to remove such unwanted data from our image we use several filters to remove such noise [1, 10]. Here we use median filter to remove impulse noise shown in the Fig.2 and Fig.3.

Fig.2. Image with salt and pepper noise occurred to external or internal disturbance

Fig.3. Image after noise removal using median filter

2.1.2 Slant Correction:

In our text handwritten analysis we can't judge a person's handwriting each time as it varies with individual as well as with a person's mind set which is best explained in the Fig.4.

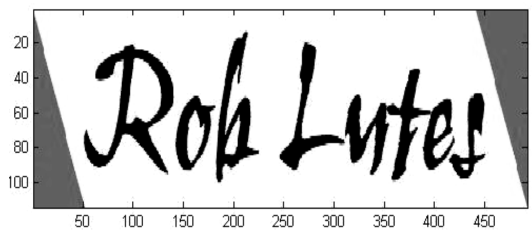

Fig.4. Sheared image

Slant is the angle that the baseline of a word or sentence makes with the horizontal direction. Slope can be removed by rotating the entire text, while slant is corrected by shearing each word which means shifting every line of pixels sideways by an amount depending on its distance from the baseline. In slant removal 2D spatial transform is used to correct the *x* data and *y* data of the given image which automatically shifts the origin of your output image to make as much of the transformed image visible as possible [8].

In character recognition process, removing slant in a document will reduce variation within classes of characters making those classes easier to recognize. It will also make character segmentation easier, since characters that are straight up have more distinct space between them than characters that are at a slant. Another application is writer identification, where slant correction can improve performance by reducing within-writer variation [12].

Slant correction is a projection transformation useful for registering or aligning images. The use of projection transformation as a means to register an image with respect to a different view points. Since this mapping is constrained at four 2-D points, there are eight coordinates and thus eight degrees of freedom in a projective transform [13, 19]. It is possible to show that the general form for the 2-D projective transformation matrix in a homogeneous coordinate system is given by matrix 1& 2 where **T** is the spatial transform matrix, **S** is input image need to be transformed and **S**' is the resultant image after slant removal is shown in the Fig.5.

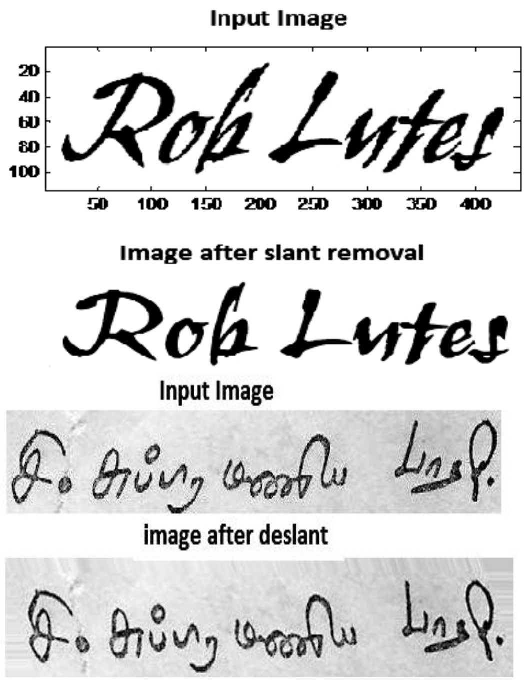

Fig.5. Image after slant removal

$$T = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} S = \begin{bmatrix} x_1 & \dots & x_N \\ y_1 & \dots & y_N \\ 1 & 1 & 1 \end{bmatrix} \tag{1}$$

$$S' = TS = \begin{bmatrix} x_1 & \dots & x_N \\ y_1 & \dots & y_N \\ 1 & 1 & 1 \end{bmatrix}$$

$$S' = TS = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{bmatrix} \begin{bmatrix} x_1 & \dots & x_N \\ y_1 & \dots & y_N \\ 1 & 1 & 1 \end{bmatrix} \tag{2}$$

2.1.3 Binarization and Normalization:

The noise removed RGB image is converted into gray image to make our future process easier. Normalization is done to remove the unwanted backgrounds [1, 3]. In our handwriting analysis it also helps in normalizing the space between horizontal and vertical lines in text image, which helps in segmenting image into individual words and letters in our text image.

3. SEGMENTATION

Segmentation is mainly performed to extract the specified or required region from an image. In handwriting analysis several segmentation process have been used. For example, histogram analysis for segmentation, OCR and GLCM Matrix are used for handwritten image segmentation [2]. In our analysis we use morphological watershed transform to segment a text document. It is more stable segmentation process and it involves in continuous segmentation boundaries. A grey-level image may be seen as a topographic relief, which involves catchment basin watershed and watershed lines. Our proposed Segmentation is performed based on the watershed lines [11].

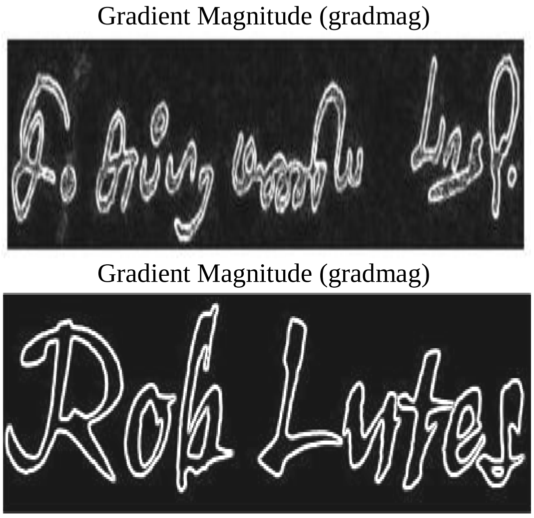

Fig.6. Edge detection image using gradient magnitude

Segmentation process involves the boundary detection region based processing i.e. analyze connected component. Usually pixel varies very rapidly along the boundaries between two regions which were usually analyzed using gradient magnitude method for boundary detection. In our process of segmentation we perform background and foreground morphological analysis before applying it into the watershed transform for segmentation is shown in the Fig.6. This morphological process involves two important process, “opening-by-reconstruction” and “closing-by-reconstruction” for cleaning up the image to avoid over segmentation and under segmentation when applying it into watershed transform [6, 9].

3.1 MORPHOLOGICAL OPENING

In morphology, the basic idea is to probe an image with a simple, pre-defined shape, drawing conclusions on how this shape fits or misses the shapes in the image. The morphology operators strongly related to Minkowski addition which is shift-invariant operator.

Morphological opening is simple erosion followed by dilation which is used to remove the unwanted structures in an image. Set theory is usually used to define the erosion (**Θ**) and dilation (⊕). In morphology we regard pixel intensities as topological highlights, which are shown in the Fig.7. Here opening by reconstruction involves erosion followed by its morphological reconstruction which helps in removing unwanted dark spot in our image using expression (3), where **A** is a binary in **E** and **B** is a structured element.

$$\begin{gathered}\mathbf{A\Theta B} = \{\mathbf{Z} \in \mathbf{E} \ \mid \ \mathbf{B_Z} \subseteq \mathbf{A}\} \\ \mathbf{A} \circ \mathbf{B} = (\mathbf{A\Theta B}) \oplus \mathbf{B} \\ \mathbf{B_Z} = \{\mathbf{b} + \mathbf{z} \ \mid \ \mathbf{b} \in \mathbf{B}\}, \forall_{\mathbf{Z}} \in \mathbf{E} \\ \mathbf{A} \oplus \mathbf{B} = \bigcap_{\mathbf{b} \in \mathbf{B}} \mathbf{A}\end{gathered} \tag{3}$$

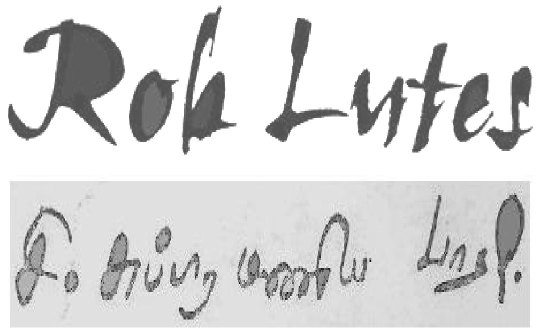

Fig.7. Opening by reconstructed image

3.2 MORPHOLOGICAL CLOSING

Morphological closing is a dilation followed by erosion which is used to merge or fill structures in an image. Closing-by-reconstruction involves dilation followed by morphological reconstruction which helps in removing the unwanted components in our background using expression (4) where A is a binary in E and B is a structured element.

$$\mathbf{A} \bullet \mathbf{B} = (\mathbf{A} \oplus \mathbf{B})\mathbf{\Theta B} \tag{4}$$

3.3 EROSION

Erosion is usually employed here to reduce objects into an image using the below expression (5) where A is a binary in E and B is a structured element.

$$\begin{gathered}\mathbf{A\Theta B} = \{\mathbf{Z} \in \mathbf{E} \ \mid \ \mathbf{B_Z} \subseteq \mathbf{A}\} \\ \mathbf{B_Z} = \{\mathbf{b} \in \mathbf{Z} \ \mid \ \mathbf{b} \subseteq \mathbf{B}\}, \forall_{\mathbf{Z}} \in \mathbf{E}\end{gathered} \tag{5}$$

where, B_Z is denoted as translation of B by a vector z. Erosion of A by B is also given by the expression (6). The output after erosion process is shown in the Fig.8.

$$\mathbf{A\Theta B} = \bigcap_{\mathbf{b} \in \mathbf{B}} \mathbf{A}_{-\mathbf{b}} \tag{6}$$

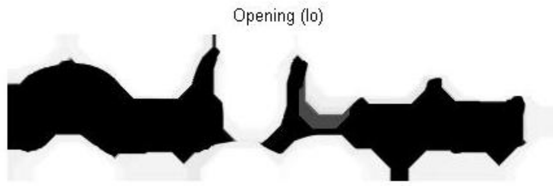

Fig.8. Opening or Erosion of the image

3.4 DILATION

Dilation is employed here to increase the object into an image using following expression (7) where A is a binary in E and B is a structured element.

$$\mathbf{A} \oplus \mathbf{B} = \bigcap_{\mathbf{b} \in \mathbf{B}} \mathbf{A_b} \tag{7}$$

The dilation is also obtained by the expression (8) and its corresponding output image is shown in the Fig.9.

$$\begin{aligned} \mathbf{A} \oplus \mathbf{B} &= \{\mathbf{z} \in \mathbf{E} \quad | \quad (\mathbf{B}^{5})_{\mathbf{z}} \cap \mathbf{A} \neq \mathbf{\Phi}\} \\ \mathbf{B}^{5} &= \{\mathbf{x} \in \mathbf{E} \quad | \quad - \in \mathbf{B}\} \end{aligned} \tag{8}$$

where, $\mathbf{B}^{5}$ denotes the symmetry of B

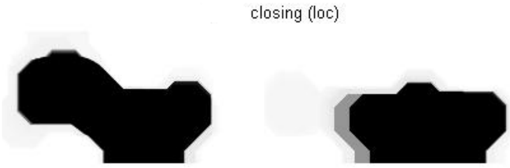

Fig.9. Closing or Dilation of the image

Further apply the watershed transform to segment the handwritten image into letter and words based on watershed lines obtained from the above process are shown in the Fig.10 and Fig.11.

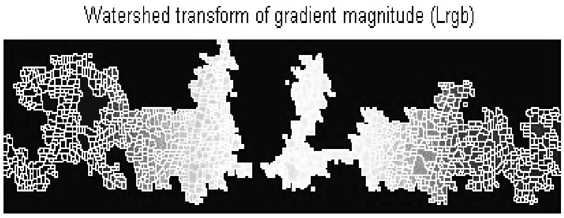

Fig.10. Morphological Watershed transformed image

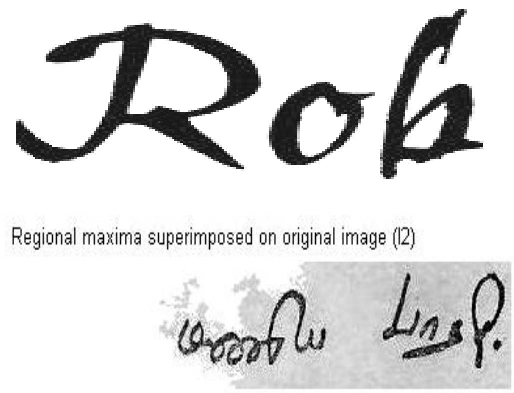

Fig.11. Segmented image using watersheds transform

4. FEATURE EXTRACTION

Feature extraction is the process of transforming input data into set of feature. Our handwriting image is now applied for feature extraction which helps in reducing the training complexity of neural networks and for reducing computation time before training process. In our analysis we use daubechies'5/3 integer to integer wavelet transform for feature extraction. Feature extraction is to get most relevant information from original data and represent that information in low density region. After the handwriting image is transformed by integer to integer wavelet transform, the coefficient cannot be used directly to describe the characteristics. The main advantage of using daubechies'5/3 integer to integer wavelet transform is that, it is easily reversible and lossless when compared with other wavelet transform [3].

4.1 MORPHOLOGICAL CLOSING DAUBECHIES' 5/3 WAVELET TRANSFORMS

Daubechies' 5/3 biorthogonal wavelet is used for 3 level decomposition and the equation is given by the Eq.(9),

$$\begin{aligned} d_i^1 &= d_i^0 - \left\lfloor \tfrac{1}{2}\left(s_i^0 + s_{i-1}^0\right)\right\rfloor \\ s_i^1 &= s_i^0 + \left\lfloor \tfrac{1}{4}\left(d_i^1 + d_{i-1}^1\right) + \tfrac{1}{2}\right\rfloor \end{aligned} \tag{9}$$

where, d_i high pass is sub band signal and s_i is low pass sub band signal [10], [14], [17].

The inverse transform to recover lossless the original samples is given by the expression (10),

$$\begin{aligned} d_i^0 &= d_i^1 + \left\lfloor \tfrac{1}{2}\left(s_i^0 + s_{i-1}^0\right)\right\rfloor \\ s_i^0 &= s_i^1 - \left\lfloor \tfrac{1}{4}\left(d_i^1 + d_{i-1}^1\right) + \tfrac{1}{2}\right\rfloor \end{aligned} \tag{10}$$

Above decomposition is performed by decomposing the rows, then by columns for one wavelet decomposition of image and the same process is repeated for multi-scale wavelet decomposition and are shown in the Fig.10. Usually in order to identify the characteristics of handwriting image we use statistical method for analyzing handwriting image by calculating mean and variance of each level of decomposition using the following Eq.(11).

Let $f(x, y)$ be the segmented image, $c_1(x, y)$, $c_2(x, y)$, $c_3(x, y)$ are the coefficients of image at each level of decomposition, mean and variance equations are given for first level of decomposition [1].

$$\begin{aligned} M_{ij} &= mean\left(c_1(x, y)\right) \\ V_{ij} &= \mathrm{var}\left(c_1(x, y)\right) \end{aligned} \tag{11}$$

Fig.12. Triple level decomposition using 5/3 ITI WT

5. NEURAL NETWORK

Neural networks are usually trained for conditions to solve problems, which are difficult for conventional computer and human beings. After feature extraction the restored image is compared with our original image for writer identification which is trained by using neural network using back propagation. Initially we have created 2 layer back propagation network using neural tool [14, 16].

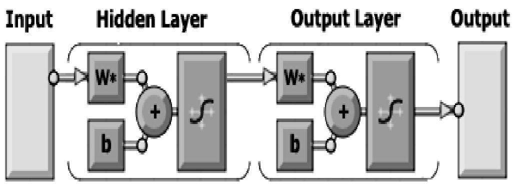

Fig.13. Created 2 Layer Neural Network

The created 2 layer network requires input and target image for training the network, which helps for the writer identification. After the repeated training with input and different target image, writer identification is done through the obtained result from training process [7, 9]. Once the network weights and biases are initialized, the network is ready for training. The multilayered feed forward network can be trained for function approximation (nonlinear regression) or pattern recognition is shown in the Fig.13. The training process requires a set of examples of proper network behavior—network inputs *p* and target outputs *t*. The default performance function for feed forward networks is the mean square error. The average squared error between the networks outputs *a* and the target outputs *t* is defined as follows:

$$F = mse = \frac{1}{N}\sum_{i-1}^{n}(e_i)^2 = \frac{1}{N}\sum_{i-1}^{n}(t_i - a_i)^2 \tag{12}$$

Network training function that updates the weight and bias values is done according to Levenberg-Marquardt optimization algorithm. One iteration of this algorithm can be written as,

$$X_{k+1} = X_{k\text{-}ak}\, g_k \tag{13}$$

where, X_k is a vector of current weights and biases, g_k is the current gradient, and αk is the learning rate. This equation is iterated until the network converges.

6. EXPERIMENTAL RESULTS

The performance plot shows the value of the performance function versus the iteration number. It plots the training, validation and test performances. The training state plot shows the progress of other training variables, such as the gradient magnitude, the number of validation checks, etc. The regression plot shows a regression between network outputs and network targets and to validate the network performance. The validation and test results also shows that the R values are greater than 0.9. The scatter plot is helpful in showing that certain data points have poor fits.

The three axes represent the training, validation and testing data. The dashed line in each axis represents the perfect result – outputs = targets. The solid line represents the best fit linear regression line between outputs and targets. The R value is an indication of the relationship between the outputs and targets. If R = 1, this indicates that there is an exact linear relationship between outputs and targets. If R is close to zero, then there is no linear relationship between outputs and targets.

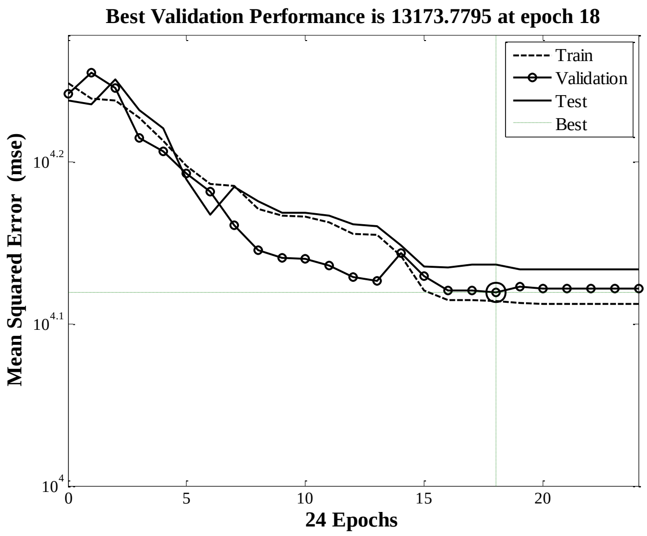

Fig.14. Performance validation

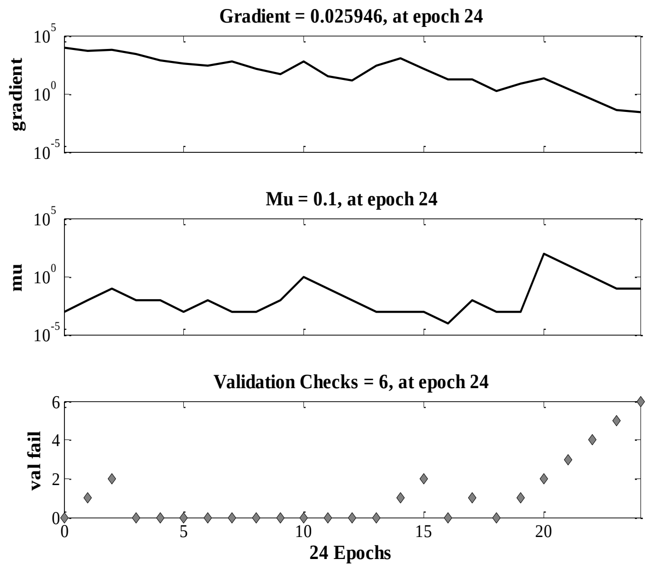

Fig.15. Mean Square Error & Gradient for Created Network

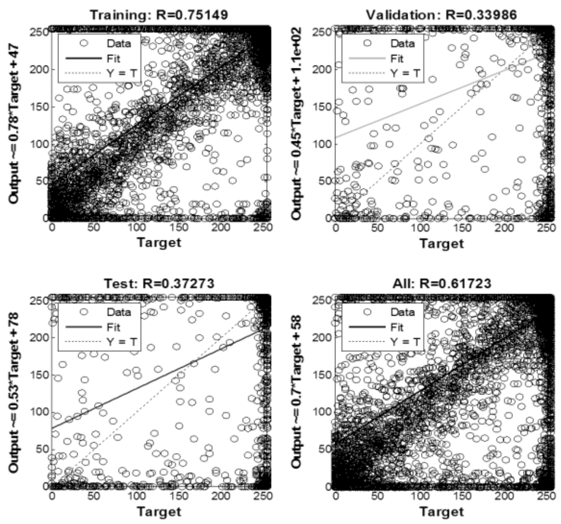

Fig.16. Regression plot for 2 layer Neural Network

Table.1. Performance variation in the output is explained by the targets

Output Image	Train(R)	Validation (R)	Testing (R)
Rob lutes	0.76514	0.33548	0.29856
Subramanya Bharathi	0.69892	0.75481	0.56481
David Cameron	0.85261	0.65482	0.58942

The output results for 50 different input handwriting images are analyzed and its output values of R listed in the Table.1 as shown above.

7. CONCLUSION

The proposed neural network writer identification system is to justify our input image matches with the writer. It has been greatly supported by slanting removal which helps in segmentation and feature extraction. This writer identification system uses simplest and efficient watershed transform to segment handwritten document which makes this system for a lossless feature extraction process. At the same time it also faces some of the problems like under and over segmentation which must be carefully handled in our segmentation process by employing “opening-by-reconstruction" and "closing-by-reconstruction” and in 2 layer neural network each time we have to give input to the network if our input image size is too large.

REFERENCES

[1] S.K. Jayanthi and D. Rajalakshmi, “Writer identification for offline Tamil handwriting based on gray-level co-occurrence matrices”, *Third International Conference on Advanced Computing*, pp. 187-192, 2011.

[2] C.M. Travieso, J.B. Alonso, O. Santana and M.A. Ferrer, “Writer identification based on graphology techniques”, *IEEE Aerospace and Electronic Systems Magazine,* Vol. 25, No. 6, pp. 35 - 42, 2010.

[3] Yali Yang and Shunhua Shen, “A handwriting image texture characteristics extraction algorithm based on integer-to-integer wavelet transform”, *Third International Congress on Image and Signal Processing*, Vol. 5, pp. 2438 - 2440, 2010.

[4] K. Ubul, D. Tursun, A. Hamdulla and A. Aysa, “A Feature Selection and Extraction Method for Uyghur Handwriting-Based Writer identification”, *International Conference on Computational Intelligence and Natural Computing*, Vol. 2, pp. 345 - 348, 2009.

[5] Bei-Bei Zhu, Zhao-Wei Shang, Feng Zhang and Bo Yuan, “Chinese handwriting-based writer identification with PDTDFB transform”, *International Conference on Wavelet Analysis and Pattern Recognition*, pp. 205 - 210, 2009.

[6] G. Louloudis, B. Gatos, I. Pratikakis and C. Halatsis, “Text line and word segmentation of handwritten documents”, *Elsevier Pattern recognition*, Vol. 42, No. 12, pp. 3169 - 3183, 2009.

[7] Yi Li, Yefeng Zheng, D. Doermann and S. Jaeger, “Script-Independent Text Line Segmentation in Freestyle Handwritten Documents” , *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 8, pp. 1313 - 1329, 2008.

[8] M. Bulacu and L. Schomaker, “Text-Independent Writer Identification and Verification Using Textural and Allographic Features”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, No. 4, pp. 701 - 717, 2007.

[9] V. Singhal, N. Navin and D. Ghosh, “Script-based classification of hand-written text documents in a multilingual environment”, *Proceedings of Thirteenth International Workshop on Research Issues in Data Engineering: Multi-lingual Information Management*, pp. 47 - 54, 2003.

[10] M.D. Adams and R.K. Ward, “Symmetric-extension-compatible reversible integer-to-integer wavelet transforms”, *IEEE Transactions on Signal Processing*, Vol. 51, No. 10, pp. 2624 - 2636, 2003.

[11] L. Schomaker and M. Bulacu, “Automatic writer identification using connected-component contours and edge-based features of uppercase Western script”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 26 , No. 6, pp. 787 - 798, 2004.

[12] E. Kavallieratou, N. Fakotakis and G. Kokkinakis, “Slant estimation algorithm for OCR system”, *Pattern Recognition*, Vol. 34, No. 12, pp. 2515 - 2522, 2001.

[13] H.E.S. Said, T.N. Tan and K.D. Baker, “Writer Identification Based on Handwriting”, *Proceedings of IEE Third European workshop on Handwriting Analysis and Recognition*, Vol. 33, No.1, pp. 133 - 148, 2000.

[14] H.J. Kim and C.C. Li, “Lossless and lossy image compression using biorthogonal wavelet transforms with multiplierless operations”, *IEEE Transactions on Circuit and Systems-II: Analog and Digital Signal Processing*, Vol. 45, No. 8, pp. 1113 - 1118, 1998.

[15] R.M. Bozinovic and S.N. Srihari, “Off-line Cursive Script Word Recognition”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 11, No. 1, pp. 68 - 83, 1989.

[16] M. Yasuhara, “Identification and decomposition of fast handwriting system”, *IEEE Transactions on Circuits and Systems*, Vol. 30, No. 11, pp. 828 - 832, 1983.

[17] M.D. Adams, “Reversible integer to integer wavelet transforms for image coding”, PhD thesis submitted to the Department of electrical and computer engineering, University of British Columbia, Vancouver, BC, Canada, 2002.

[18] E. Kavallieratou, N. Fakotakis and G. Kokkinakis, “New algorithms for skewing correction and slant removal on word-level [OCR]”, *Proceedings of the Sixth IEEE International Conference on Electronics, Circuits and Systems*, Vol. 2, pp. 1159 - 11622, 1999.