Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 31 No. 6s (2024)

Multilingual Handwritten Recognition using DNN

Thenarasi V[1], Sharath Kumar Y H[2]

[1,2]Government First Grade College Siddhartha layout Mysore -570011

Department of ISE, MITMysore

Article History:

Received: *15-06-2024*

Revised: *13-07-2024*

Accepted: *04-08-2024*

Abstract:

Handwriting has continued to persist as a means of communication and recording information in day-to-day life even with the introduction of new technologies. Given its ubiquity in human transactions, machine recognition of handwriting has practical significance, as in reading handwritten notes in a Personal Digital Assistant (PDA), in postal addresses on envelopes, in amounts in bank cheques, in handwritten fields in forms, etc. However the script-independent methodology for multilingual Offline Handwriting Recognition (OHR) becomes a very difficult task, since the multilingual methods have different characters and words. Prediction of script-independent methodology reduces accuracy rate of the OHR method. To overcome this problem, new OHR of Tamil and English is transduced into electronic data. It majorly focuses on the removal of noises and word, character segmentation methods with higher recognition rate. The images which are scanned may also contain noises. Image denoising steps consists of binarization, noise elimination, and size normalization. Words and characters segmentation are performed by using Particle Swarm Optimization (PSO) algorithm. Then those segmented samples are used for the next step which is feature extraction. Finally, word recognition is performed by using the deep neural network classifier.

Keywords: PSO, DNN, OHR, PDA.

1. Introduction

In the context of our country, script recognition is an important task because more number of prevailing scripts and many of the official documents are found in single scripts as well as in multi-scripts in nature. English is being commonly used along with regional script. Hence, it is pre-requisite to identify the script in the documentand then feed the document to respective OCR for further processing. In computerized processing of handwritten or printed document images, wheremany documents are multi-script in nature. Script identification plays a significant roleas prerequisite to OCR system. In script identification, the document is subjected to analysis for identifying the script type of the text in image. A document containing text information in more than one script is called a multi-script document. In many countries, it is common to communicate in different languages and that many documents are multilingual or multi-script.

2. Related Work

Research in script identification is not a new one. Past many years, methods have been proposed for script identification from scanned document images. More research is done in identification of script from printed documents compared to researchin script identification from handwritten documents. Handwritten documents are common in developing countries like India. Moreover, many scripts are

used in India compared to any other country in the world [1]. Given a document image, script identification task can be carried out at line level ,word level and block level. Block level script identification identifies the script of the given document in a mixture of various script documents. If document appears in single script, a block of the document is sufficient to identify the script type [2]. Hence, input document to the block level identification is a mono script document. In case of multi-script document, line level or word level identification is preferred. Many of the reported studies in the literature accomplish script recognition either at the line level or at the word level. In line based script identification, a documentimage can contain more than one script but it requires the same script on a single line. Word level script identification allows the document to contain more than one script and the script of every word is identified. However, these components are applicable only after line and word segmentation of the underlying document image. However, such a fine segmentation is not required at block level and consequently simplifies the script classification task. Generally, script identification task comprises of three stages viz, pre- processing, feature extraction and recognition. Pre-processing eliminates unwanted objects in the document image (noise, background, etc.) and extracts the words or lines or blocks from the image for feature -extraction. The feature extraction method extracts the relevant data (usually a reduced representation of the input object),called features vector, which is then fed to a classifier for script recognition. A brief overview of different methods available in the Literature for script identification is presented below. Emphasis is given on the recognition methods for Indian script as domain of our study confines to Indian scripts. Using fractal based features, S. Ben Moussa et al., in [3] have reported script identification for Arabic andRoman scripts at line level of the text document. Profile based script identification technique is presented by M.C. Padma et al.,in [4]. Distinct features at top and bottom of the text line are considered for feature extraction. Features are extracted by calculating density ratio, pixel distribution, max pixel density at the bottom row of text line and density of connected components at thetop of the text line. Classification is performed by using KNN. Mallikarjun Hangarge et al., in [5] have proposed a visual texture based methodology for handwritten script recognition at line level and block level. Horizontalprojection profile is used to extract non-touching text line from document image. Morphological filters are utilized to extract 13 spatial spread features. KNN classifier has been used to classify the script Rajput et al., in [6] presented line level script recognition using Gabor filter combined with DCT and wavelets. Nine Indian scripts were used for performing experiment, SVM classifier reported better recognition accuracy compared to the performance by KNN classifier. Prasanthkumar et al., in [7] have proposed an automatic separation of text line followed by word segmentation and computed morphological and Gabor features. Classification was carried out by using MLP, SVMand KNN classifiers with the observation that MLP classifier yielded better results overSVM and KNN. Long Short Term Memory (LSTM) architecture based script identification is reported by Adnan Ul-Hasan in [8]. DCT and distance transform based script identification from handwritten document is proposed by S.K. Md. Obaidullah et al., in[9]. Bangla, Roman, Devanagari and Oriya scripts are considered for performing experiments. Feature vector based on texture (BRT, BDCT, BFFT and BDT) are used to classify bi-script and tri-scripts. Gabor filter and morphological re-construction is proposed by Nibaran Das et al., in [10] to identify the Indic scripts by using MLP classifier. B. Shi et al., in [11] have incorporated deep learning based script identificationfrom natural images. They have presented a two stage features

extraction methodologyi.e., feature extraction and discriminative clustering at mid-level representation , globalfine tuning at second stage by modeling feature extraction and classification into one neural network by transferring learned parameters at first stage.Back propagation is used to train the network .Script dependent and script independent feature based classification of script at block level, line level and word level is proposed by S.K. Obaidullah in [12]. Classification of script is carried out by using MLP and random forest classifiers. Md. Obaidullah et al., in [13] have proposed a method for handwritten script identification based on structural, directional and texture based features. The features were extracted from document image in binary and input to MLP for text identification.Judith Hochberg et al., in [14] have reported recognition of six Indic and non-Indic scripts, viz, Arabic, Chinese, Cyrillic, Devanagari, Japanese, and Latin,using features like horizontal and vertical centroids, sphericity, aspect ratio, white holes, and so forth. The experiments have been carried out at document level. V Singhal et al., in [15] proposed a technique to identify four Indic scripts, viz,Devanagari, Bangla, Telugu, and Latin. They have used rotation invariant texture features using multi-channel Gabor filtering and gray level co-occurrence matrix. It is observed that many of the approaches are sensitive to scaling and rotation. Very recently, Scale Invariant Feature Transform (SIFT) technique has been proved to be a promising image features, for object recognition, that provides a set of features of an object that are not affected by many of the complications experienced in other methodsincluding object scaling and rotation proposed by David G. Lowe in [16]. Recognition of Bangla and English scripts has be presented by Lijun Zhou et al., in [17] using connected component profile based features. Experiments are performed on the document text at line, word and character level. Multi-script identification at word level based upon features extracted using Discrete CosineTransform (DCT) and Wavelets of Daubechies family have been proposed by G .G Rajput et al., in [18]. The recognition has been carried out by using neural network forhandwritten documents of nine Indian scripts including English script. K. Roy et al., in [19] have reported recognition of six popular Indic-scripts, viz,Bangla, Devanagari, Malayalam, Urdu, Oriya and Roman by using component based features, fractal dimension based features, circularity based features and so forth. This is the first kind of work involving six Indic scripts altogether. Mallikarjun Hangarge etal., in [20] have proposed a word level scheme based on directional DCT based featureto recognize six Indic-scripts, viz, Roman, Devanagari, Kannada, Telugu, Tamil, and Malayalam.

3. **Proposed Method**

The proposed work focuses on removing noises and image segmentation methods with recognition methods. The proposed flow diagram is shownin Figure 1.

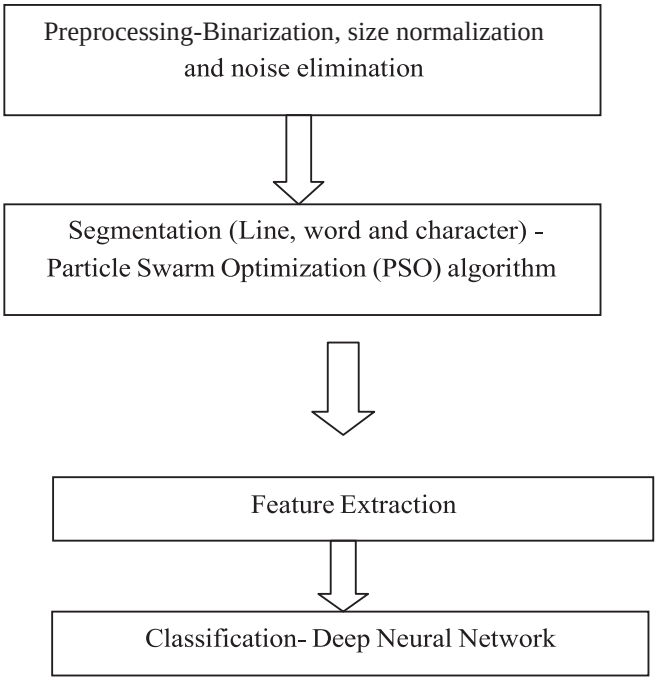


Figure 1: Proposed Flow Diagram for Handwritten Recognition

3.1.1 PRE-PROOCESSING

Optical Character Recognition is one of the leading approaches used in handwriting recognition. Optical character recognition shortly OCR performs, text analyses to recognize the character. It deals with various input sources such as documents that were scanned, paper based content, etc. This obtained input could be converted to a form that allows for editing. The process of OCR involves various stagessuch as recognition, segmentation, post processing and pre-processing. Firstly, pre- processing step should be carried out and it acts as a difficult stage too.

Noise removal by Modified Median Filter

Pre-processing deals with reduction or removal of noise in images. There are different types of noise are available namely, noise, Salt, Shot, Quantization noise (uniform noise), Film grain, Anisotropic noise and Periodic noise. To obtain efficient results, a modified mechanism is used in research areas namely K-Algorithm. This approach includes two different stages namely, binarization and filtering. Re-sampling algorithm has been used to perform filtering process in noise elimination steps. Filtering always refers to the various functions that are predefined in image to assign value to pixel which is consider as function of values.Bit patterns which are unwanted are diminished using this methodology. It would remove the images that are textured slightly or background with colour and sharper it. Due to thisprocess, much of possible noise is reduced by make retaining only relevant information.Filter is of two categories: non-linear and linear filters. Since linear type has some disadvantages, the non-linear type has been used to overcome those disadvantages suchas blurring edges, blurring details and destruction in lines.

Binarization

Binarization is one among the approaches used for image denoising. Binarizationstep could be carried out after a filtering process. These filtering and binarization methodologies are adapted by modified mechanisms for noise reduction. The name of the modified technology is K-Algorithm which would actively involve in performing removal of noise presence in image. This binarization (Ntogas et al 2013) approach could be applicable for images in order to separate the text from background. This process is purely based on thresholding and the filtering which is combined with algorithms of image processing. The procedures of binarization involve 5 sets of discrete stages which depend on various classes of images. It acts as refinement methodology inorder to improve image quality. The result obtained in filtering stage might still possess coloured backgrounds slightly that lead to interference in functioning of next stages. In order to avoid and deal with these said issues, binarization has been introduced.Binarization step is involved in conversion of filtering image into digital image which means binary. From this, the value of threshold could be calculated and finally based oncolours, the processes are carried out. That is, if the intensity value of pixel is above threshold value, it is set as white (0) and if it is below the thresholding value, it is set tobe black (1). Thus by using average of overall pixel intensities in document, thethreshold value could be obtained.

Size Normalization

In image processing, normalization is one of the important steps that could obtainbetter image recognition. Normalization is the process of changing the ranges of intensityvalue available for every image. For instance, it is used in application of photography tooin order to avoid poor contrast. It is also called contrast based stretching or histogram based. The purpose of this approach is to bring out images with good visual sense. Thissize normalization aims to remove the noise. In addition to that, it could improve the image into high range quality based on intensity result. This means the quality depends on value in the term of intensity. This process of normalization is a linear process. For instance, if the image has intensity ranges between 50 and 180 then the range desired is between 0 and 255. Eachand every pixel based intensity is multiplied by using 255/180 and obtains a range of 0 to 255.Automatic normalization typically normalize the image in any file format. It leads toproduction of image in constant dimension. It aims to reduce the variations which occurduring writing of data. For Instance, the size normalization (Kumar et al 2013) has beenused to adjust the size of character in a form of standard. Segmentation Using Particle Swarm Optimization (PSO). The phenomenon widely involved around the world is problem solving. Adomain where emerges from specific behaviours of particular particle during interactions. Because of topology structure of communication, the populations have been organized. This is carried out in social network. In the research of PSO, coordinates havebeen tracking in solution space. This could be associated along with fitness which meansbest solution. The value obtained from this process is said to be as pbest which stands forpersonal best. Alternate value at best level which is tracked by PSO and this isconsidered as best value which is obtained from particle present in the neighbourhood ofparticle.

3.1.2 DEEP NEURAL NETWORK CLASSIFIER

Deep neural network is called to be learning approach, where it could be used forhumans to obtain knowledge. In another words, it would be considered as an approach for automation of predictive base analysis. Deep Neural Network shortly DNN is an Artificial Neural Network with various hidden layers. These layers are placed between output and input layers. It is used to develop complex relationships that are non-linear. This model involves generation of compositional approach where objects can be expressed in a layered based composition. In addition, the layers that are available extrawould enable the features from the layer which is lower. Since DNN is a feed forward network, the data can flow from input to output layer by one direction which means notloop towards back. This deep neural classifier has been used in Image recognition due tothe presence of noise. Classifier has been designed for dealing with noise in images in order to remove it.

4. Results

In order to measure the results of the methods 12000 samples of 30 Kannada and English characters are collected from different handwritten documents. 9000 samples (300 samples * 30 characters) are gathered for training purpose in DNN, rest of them 3000 samples (100 samples * 30 characters) are used for testing purpose.

F- measure

It is the harmonic implication of recall and precision. A higher value of FM shows the efficiency of the perfect binarization.

$$F_{score} = \frac{2tp}{(2tp+fn+fp)} \quad (1)$$

Recall

It is the fraction of relevant characters, which is retrieved over the entire amount of the characters present in the detected images.

$$Recall = \frac{tp}{(tp+fn)} \quad (2)$$

Precision

It indicates the amount of information retrieval for the user.

$$Precision = \frac{tp}{(tp+fp)} \quad (3)$$

Accuracy

It is called the highest intuitive overall performance measure and it shows a ratio of efficiently expected statements to total observations.

$$Accuracy = \frac{tp+tn}{(tp+tn+fn+fp)} * 100 \quad (4)$$

where,

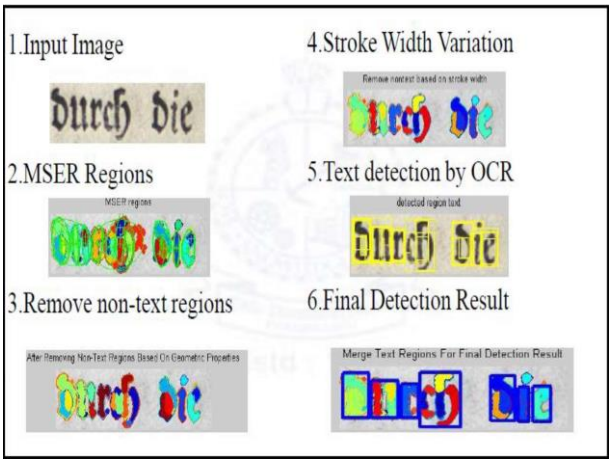
$t_{p=}$ True positive $f_{n=}$ False Negative

False-positive f_p=

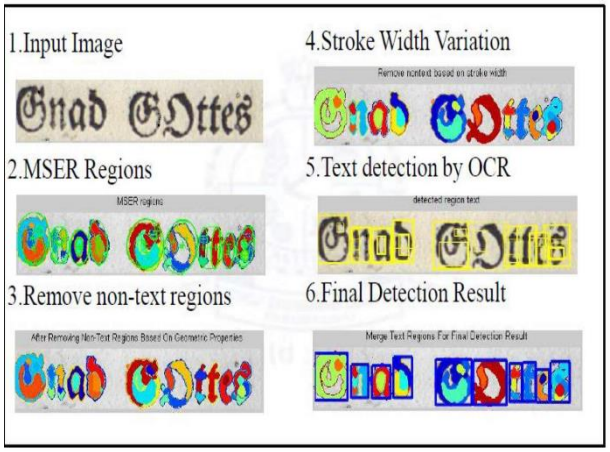
t_n= True Negative

Simulation Result for the Proposed Method

The simulation results are depicted in Figures 2 (a)-(d) for a few samples.

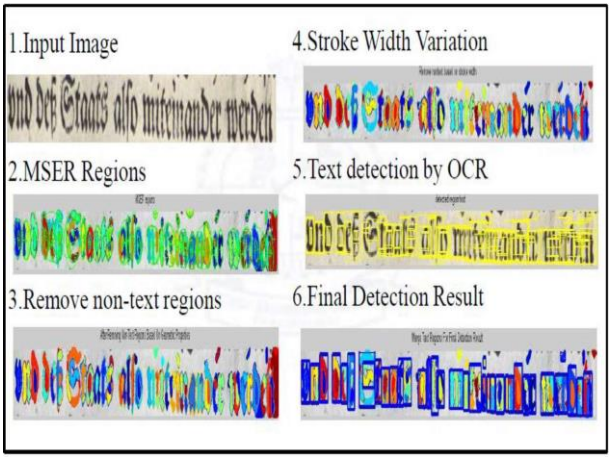


(a).

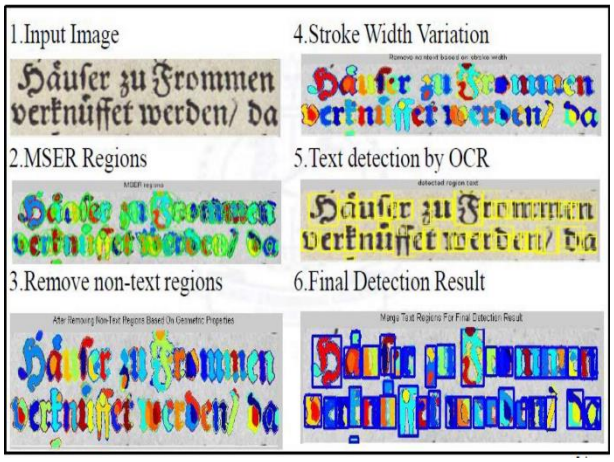


(b)

Figure 2: (Continued)



(c)



(d)

Figure 3 (a)-(d) Enhanced MSER based segmentation results

Similarly, the remaining samples present in datasets have been simulated. The corresponding line detection rate, word detection rate and character detection rate is calculated. Table 1 depicts the performance of detection rate in terms of line, word and character segmentation for 10 samples.

Table 1: Segmentation rate for different samples

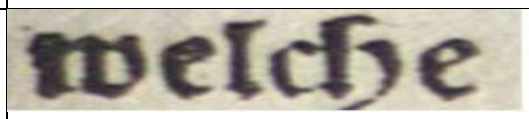
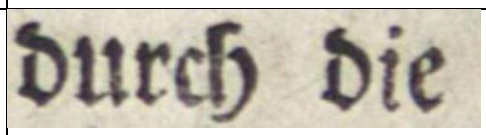
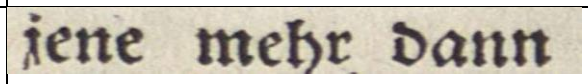
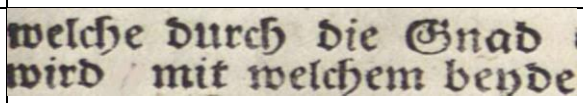
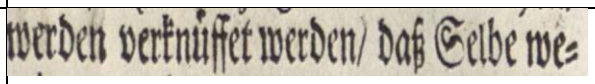
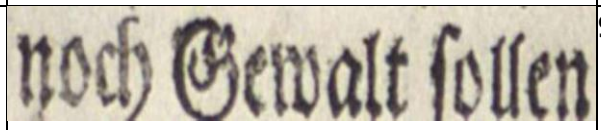
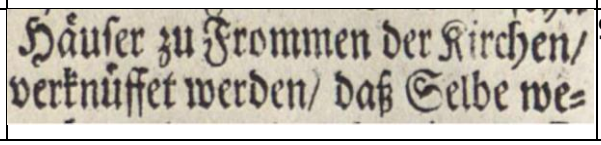
Sample	Input Image	Line detection rate(%)	Word detection rate(%)	Character detection rate(%)
1		-	86.27	89.32
2		96.36	88.27	92.45
3		94.67	87.63	90.26
4		97.81	86.47	89.35
5		98.63	88.72	95.54
6		88.27	89.42	92.73
7		91.63	95.46	87.56
8		98.76	86.59	91.49
9		93.72	96.02	87.15
10		92.53	97.32	86.41

Figure 4 shows the output obtained by stroke width filtering and DNN classifier for a typical sample in the segmented image. Similarly all samples are simulated and evaluated in terms of Precision, Recall, F-measure and Accuracy.

.

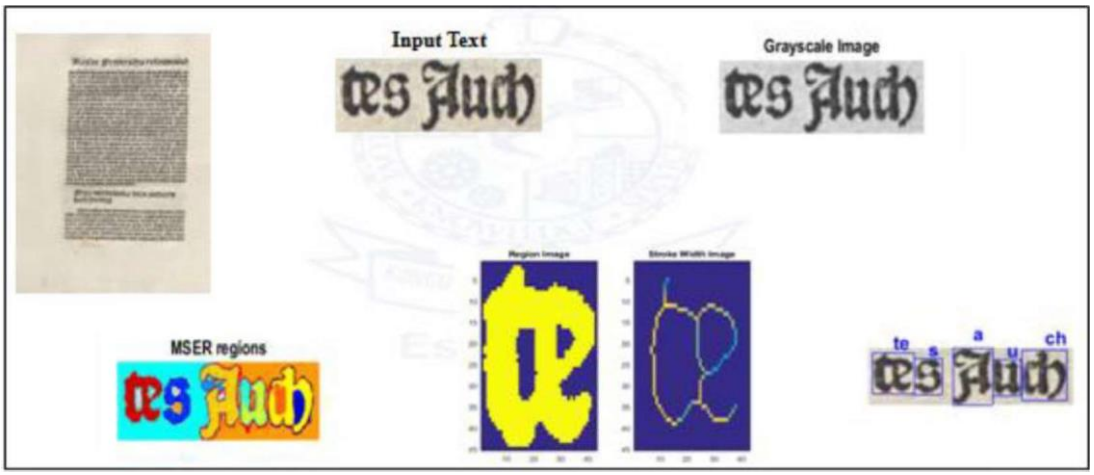


Figure 4: SWF and RF classifier output

Performance Metric

Table 2: Performance Metric Calculation

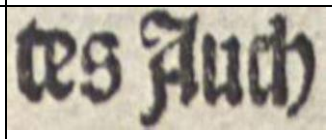
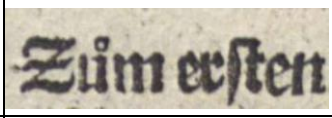
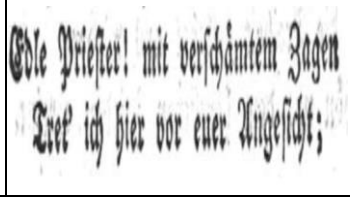
S.NO	Samples of HDLA 2011 dataset	Recall	Precision	F-Measure	Accuracy
1		1	1	1	100%
2		1	1	1	100%
3		1	0.90	0.95	90%
4		0.95	1	0.97	94.6%

Table 2 (Continued)

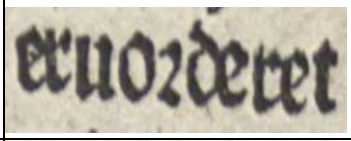
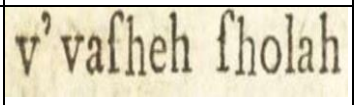
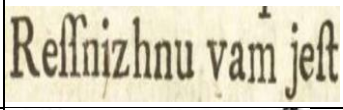
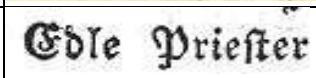
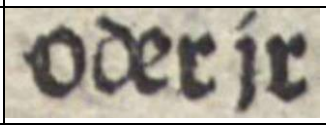
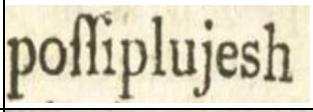
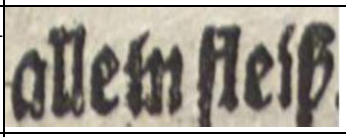
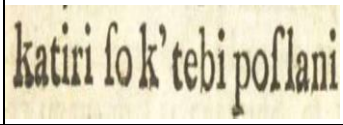
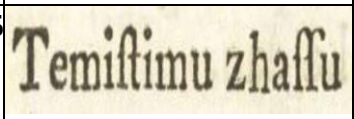
5		1	1	1	100%
6		1	1	1	100%
7		0.94	1	0.97	94.1%
8		1	1	1	100%
9		1	1	1	100%
10		0.91	1	0.96	91.7%
11		1	1	1	100%
12		1	1	1	100%
13		1	1	1	100%
14		1	1	1	100%
15		0.87	1	0.93	86.6%
16		1	1	1	100%

Table 3 (Continued)

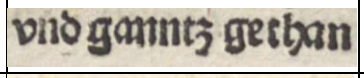
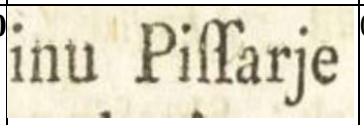
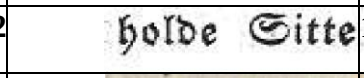
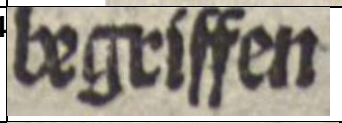
17		1	1	1	100%
18		1	1	1	100%
19		1	1	1	100%
20		0.73	1	0.84	72.7%
21		1	1	1	100%
22		1	1	1	100%
23		1	1	1	100%
24		1	1	1	100%
25		0.77	1	0.87	76.9%

Table 3 shows the performance metric values obtained for 25 samples through simulation. Similarly, all samples are simulated and their corresponding performance metric values are calculated .

5. Summary

The handwritten recognition is generally of two types namely, online recognition and offline recognition. This model deals with the off-line recognition in which the pre-processing of image is performed in which binarization, noise elimination and normalization are performed. The trailing process involved is segmentation in which thePSO algorithm is employed. The feature extraction phase helps to obtain the set of features and then the classification phase is where the deep neural network classifier is proposed for handwritten extraction. It is an iterative model. The final phase is where post-processing is done for efficient identification in which the errors occurred in classification and recognition of words is eliminated.

References

[1] Sahare, P., & Dhok, S. B. (2017). Script identification algorithms: a survey. International Journal of Multimedia Information Retrieval, 6(3), 211–232. doi:10.1007/s 13735-017-0130-2.

[2] K. Ubul, G. Tursun, A. Aysa, D. Impedovo, G. Pirlo, T. Yibulayin , " Script Identification of Multi-Script Documents: a Survey " , IEEE Access DOI 10.1109/ACCESS.2017.2689159 volume 5 2017, pp. 6546-6559.

[3] S. Ben Moussa, A. Zahour, A. Benabdelhafid and A.M. Alimi, "Fractal-Based Systemfor Arabic/Latin, Printed/Handwritten Script Identification, " 978-1-4244-2175- 6/08/$25.00 ©2008 IEEE.

[4] M. C. Padma, P. A. Vijaya "Script Identification From Trilingual Documents Using Profile Based Features" International Journal Of Computer Science And Applications, Technomathematics Research Foundation Vol. 7no. 4, Pp. 16 – 33, 2010.

[5] Mallikarjun Hangarge and B.V.Dhandra "Offline Handwritten Script Identification in Document Images" , InternationalJournal of Computer Applications (0975 – 8887) Volume 4 – No.6, July 2010.

[6] Dr. G.G. Rajput, Anita H.B "Handwritten Script Recognition at Line Level – A Multiple Feature Based Approach" International Journal of Engineering andInnovative Technology (IJEIT) Volume 3, Issue 4, October 2013. Pp 90-95.

[7] Prasanthkumar P V and Dileesh E D "Word level Script and Language identification for Unconstrained handwritten document images" 2014 3rd International Conference on Eco-friendly Computing and Communication Systems , 978-1-4799- 7002-5/14 $31.00 © 2014 IEEE DOI 10.1109/Eco-friendly.2014.78,

[8] Adnan Ul-Hasan, Muhammad Zeshan Afzal, Faisal Shafaitt, Marcus Liwicki, and Thomas M. Breuel "A Sequence Learning Approach for Multiple Script Identification". 978-1-4799-1805-8/15/$31.00 ©2015 IEEE pp 1046-1050.

[9] Sk Md Obaidullah, Rownaqul Karim, Sujal Shaikh, Chayan Halder, Nibaran Das, KaushikRoy "Transform Based Approach for Indic Script Identification from Handwritten Document Images" 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN), 978-1-4673-6823-0/15/$31.00 ©2015 IEEE.

[10] Sk Md Obaidullah, Nibaran Das, Kaushik Roy "Convolution Based Technique for Indic Script Identification from Handwritten Document Images" I.J. Image, Graphics and Signal Processing, 2015, 5, 49-57 Published Online April 2015 in 12. MECS (http://www.mecs-press.org/) DOI: 10.5815/ijigsp.2015.05.06.

[11] Baoguang Shi. Xiang Bai, Cong Yao "Script identification in the wild via discriminative convolutional neural network" Pattern Recognition 52 (2016)448–458.

[12] Sk Md Obaidullah, K. C. Santosh, Chayan Halder, Nibaran Das, Kaushik Roy"Automatic Indic script identification from handwritten documents: page, block, line and word-level approach" Int. J. Mach. Learn. & Cyber. DOI 10.1007/s13042-017- 0702-8. Published online 2017.

[13] Sk Md Obaidullah, Chayan Halder2, K. C. Santosh, Nibaran Das, Kaushik Roy"Automatic Line-Level Script Identification From Handwritten Document Images - ARegion-Wise Classification Framework For Indian Subcontinent"Malaysian Journal of Computer Science. Vol. 31(1), 2018, pp 63-84.

[14] Judith Hochberg, Kevin Bowers, Michael Cannon and Patrick Kelly, "Script and language identi_cation for handwrittendocument images," International Journal on Document Analysis and Recognition ©Springer-Verlag 1999. IJDAR (1999) 2: 45-52.

[15] V. Singhal, N. Navin and D. Ghosh, "Script-based Classification of Hand-written Text Documents in a Multilingual Environment," 0-7803-7868-7/03/$17.00 ©2003 IEEE.

[16] DAVID G. LOWE, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision 60(2), 91–110, 2004 © 2004 Kluwer Academic Publishers. Manufactured in The Netherlands.

[17] Lijun Zhou, Yue Lu and Chew Lim Tan, "Bangla/English script identification based on analysis of connected component profiles," H. Bunke and A.L. Spitz (Eds.): DAS 2006, LNCS 3872, pp. 243–254, 2006. ©Springer-Verlag BerlinHeidelberg 2006.

[18] G. G. Rajput and Anita H. B, "Handwritten Script Recognition using DCT and Wavelet Features at Block Level,"IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition RTIPPR, 2010.pp 158-163.

[19] K. Roy, S. Kundu Das and Sk Md Obaidullah, "Script Identification from Handwritten Document," 2011 Third National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics, 978-0-7695-4599-8/11 $26.00 © 2011 IEEE DOI 10.1109/NCVPRIPG.2011.22.

[20] Mallikarjun Hangarge, K.C. Santosh and Rajmohan Pardeshi," Directional Discrete Cosine Transform for Handwritten Script Identification," 2013 12th InternationalConference on Document Analysis and Recognition, 1520-5363/13 $26.00 © 2013 IEEE, DOI 10.1109/ICDAR.2013.76.

[21] Rajneesh Rani, Renu Dhir and Gurpreet Singh Lehal, “Script dentification of Pre- Segmented Multi-Font Characters and Digits,” 2013 12th International Conference onDocument Analysis and Recognition, 1520-5363/13 $26.00 © 2013 IEEE, DOI10.1109/ICDAR.2013.233.

[22] Abdel Belad, Santosh K.C and Vincent Poulain d'Andecy, “Handwritten and Printed Text Separation in Real Document,” The Thirteenth IAPR International Conference on Machine Vision Applications - 2013, May 2013, Kyoto,Japan. 2013.

[23] S Basavaraj Patil, N. V. Subbareddy, "Neural network-based system for script identification in Indian documents", Sadhana Vol. 27, Part 1, February 2002, pp. 83–97.

[24] D Dhanya, A G Ramakrishnan_ And Peeta Basa Pati “Script identification in printedbilingual documents” ,Sadhana Vol. 27, Part 1, February 2002, pp. 73–82.