Segmentation and Identification of Bilingual Offline Handwritten Scripts (Devanagari and Roman)

Priyanka Nirwan, Gurpreet Singh

Abstract: Hand written text acknowledgment field has been considered as one of the hardest issues in the digital word. The multifaceted nature dimension of this exploration zone is high due to the reasons like diverse method for writing pursued by the clients, auxiliary independences, age elements of people and so on. This paper shows a novel procedure for the recognition of handwritten scripts, for example division of words and characters. In this paper, we have used two different scripts :"Devanagari" and "Roman" scripts. For which three Convolution Neural Networks(CNN) models are applied on different types of classification: one for language classification for which we have achieved 98% accuracy, second one for Devanagari character classification for which we have achieved 89% and third one for Roman character classification for which have achieved 97% respectively.

Index Terms: Convolution Neural Network;Character segmentation;CharacterIdentification;Deeplearning;Handwritten text recognition; Devanagari script;Roman script.

I. INTRODUCTION

Handwritten character recognition is an exploration area in AI, Computer vision and Image processing. A handwriting recognizer acquires and distinguishes characters in documents and other sources and converts them into machine-readable form. Such innovation discovers its application in fields like optical character recognition (OCR).[1]. OCR most often includes three procedures, text localization, character segmentation and recognition[2][3]. The recognition is a significant zone of image processing, which is on experienced stage for machine printed content. However, for unify text in multilingual condition, remains a demanding issue[4]. Some of the reasons, responsible for the challenging nature of handwriting recognition application considered are: different scripts from different regions, different writing styles, aging effects, different style of holding the pen etc[5].

A concise description of Devanagari and Roman scripts are given below:

I. Devanagari script comprises of 11 vowels , 14 modifiers and 33 consonants shown in fig.1(a) and (b)[6].

Fig.1.(a) Vowels And Modifiers

Fig.(b) Consonants

A	E	I	O	U

Fig.2.(a) Vowels

B	C	D	F	G	H	I
J	K	L	M	N	P	R
S	T	V	W	X	Y	Z

Fig.(b) Consonants

II. Roman Script comprises of 5 vowels and 21 consonants as shown in fig.2 (a)and (b).

It is noticed that most of the segmentation and identification strategies revealed in the literature are created for a particular text or script(either handwritten or printed). Hence, the focus of this proposed research work is to segment and identify the character from scanned bilingual handwritten text.

II. RELATED WORK

Preeti Sharma and Manoj Kumar Sachan, 2017 [7]performed the research work on Hindi Language. OCR was used to convert handeritten text into editable form. A hybrid approach was used for character segmentation. An accuracy of 96% was achieved for the dataset of distinct categories. This algorithm provides good results for broken, isolated and broken characters, but results for skewed and overlapping characters are not good.

Retrieval Number: B11780782S619/19©BEIESP
DOI: 10.35940/ijrte.B1178.0782S619

603

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

Priya et al., 2016[8] decribed the two approaches,i.e. On-line and Off-line approach for handwritten text recognition. On-line approach takes input at run time while Off-line takes scanned images as an input and also purposed that results of on-line approach were better than off-line approach.

Nhat et al.,2016[9] proposed a line segmentation approach for hand written images. They used FCN (Fully Convolutional Network) classifier for the purpose of segmentation. Authors also per posed line adjacency graph technique to deal with the complexity factor of touching lines. They considered the dataset of 200 document images, 150 for testing and 50 for training. The overall accuracy gained by their proposed system was claimed to be 95% approximately.

Wu et al.,2015[10] devised a model for multiple instance learning application using the concepts of deep networks. The main concept was based on image classification and auto-annotation scheme. They tried to improve the learning by considering the direct relation between the object present in the image and its proposed or possible annotation. They worked on the dictionary level approach by considering a dataset of 981 nouns. They claimed good results from the proposed system, but only the restriction was the fixed number of objects. Their system needed a new vocabulary to handle unknown objects.

Santos et al.,2014[11] proposed DCNN to perform sentimental analysis on short texts by using character to sentence level information. A method was applied for corpora of two distinguish areas: the Stanford Sentiment Treebank(SST), for movie reviews and the Stanford Twitter Sentiment corpus(STS) for twitter messages. For the SSTb collection, sentimental prediction accuracy for positive and negative classification was 85.7% accuracy, fine-grained classification, was 48.3% accuracy and for the STS collection, accuracy was 86.4%.

T. Bluche et al.,2013[12] explore the mix of concealed markov model and Convolutional neural systems for handwritten text acknowledgment. The Convolutional neural systems works effectively connected to different PC vision undertakings, along with manually written character acknowledgment. This experiment demonstrate that supplant Gaussian blends to register emanation prospect in shrouded markov models (half and half mix), or fill in as a highlight extractor for a standard Gaussian HMM framework. The proposed frameworks outflank a fundamental HMM dependent on either de-correlated pixels or handmade highlights. Two freely accessible databases, were used and reported up to 60% (Rimes) and 35% (IAM) improvement contrasted with a Gaussian HMM dependent on pixel esteems.

T.Bluche et al.,2013[13] performed an experiment, shows that learning feature with CNN is superior to anything utilizing hand made features for handwritten word recolonization. Two types of framework were used, A

graphene based division and a sliding window division. The examination was directed on the rimes database. The framework dependent on graphene division yields lower acknowledgment rate yet is excepting quick, which is reasonable for a specific app, like the record to classification.

III. SYSTEM DESCRIPTION

A. Overview Of The System

Architecture of Deep CNN model comprises of two convolution layers (CL1,CL2) and two inspecting layers (SL1,SL2). In the initial phase, by using three prepared channels the first information input image is processed, called convolution part window estimate 3*3. Three feature maps are produced in the CL1 layer, and after that, for every feature map The restricted districts are averaged and weighted, and three new feature maps are introduced in the SL1 layer through a RELU initiation function. These highlight maps are then processed with the three prepared channels of the CL2 layer, and three element maps are yield through the SL2 layer. The last yield of SL2 layer is vectorized and afterward contribution to the conventional neural system for training.

B. Convolution Neural Network

The architecture of CNN(Convolution Neural Network) is designed using multi layer feed forward network, which is tailored to minimize the sensitivity, translation, distortions and rotations of the input images[14]. Insensitivity toward nearby changes is incorporated with the system engineering by compelling arrangements of every units situated at better places to utilize indistinguishable weight vectors, in this way constraining them to distinguish a similar component on various pieces of the input[1]. The yields of the every units at indistinguishable areas in various element maps can be all in all idea of as a nearby element vector. The element maps are at that point sub sampled to additionally improve the stability of the yield as for the movements and disfigurements. Our model consists of two Convolutional layers with 32, 64 feature map respectively with convolution filters 3*3, 2*2px channels trailed by max-pooling activities. This architecture is fully connected and consist of hidden layer of 256 concealed units and a yield softmax layer. These models are trained on a GPU using Tensor Flow – GPU (Deep Learning Library). We arbitrarily split the first preparing sets of words into a system, preparing set(90%) and a system advancement set(10%).

C. Dataset

We have used Open source dataset, which contains 2900 handwritten images of every character .i.e. 71,890 total images for Roman script and 2000 handwritten images for each character .i.e. 72,000 total images for Devanagari script which means approximately 1,43,890 images are used for testing and training the network each image is of 28*28px.

IV. METHODOLOGY

The main aim of this experiment is to train the network to segment and identify the handwritten

script. To perform the work first of all, we have done segmentation of character[15][16], for which we have followed following steps fig.3[6].

Image is scanned, pre-processing of the image is performed to eliminate inconsistency by reshaping and resizing the image.

By applying Vertical Projection Line segmentation is performed after that Horizontal Projection is applied for Word segmentation, then by applying headline removal algorithm character is segmented[15][17][18].

This generated Output is in the form of segmented characters, pre-processing of these characters are performed again and are feed to the CNN for real time prediction of trained model. For which we have trained our model using our dataset. Our target is to predict or identify the language of segmented character.
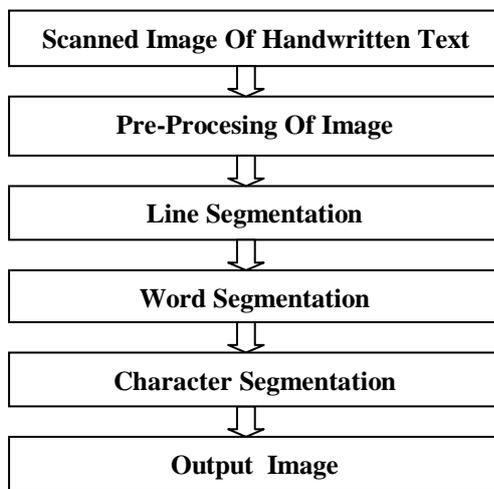
Fig.3. Process For Character Segmentation
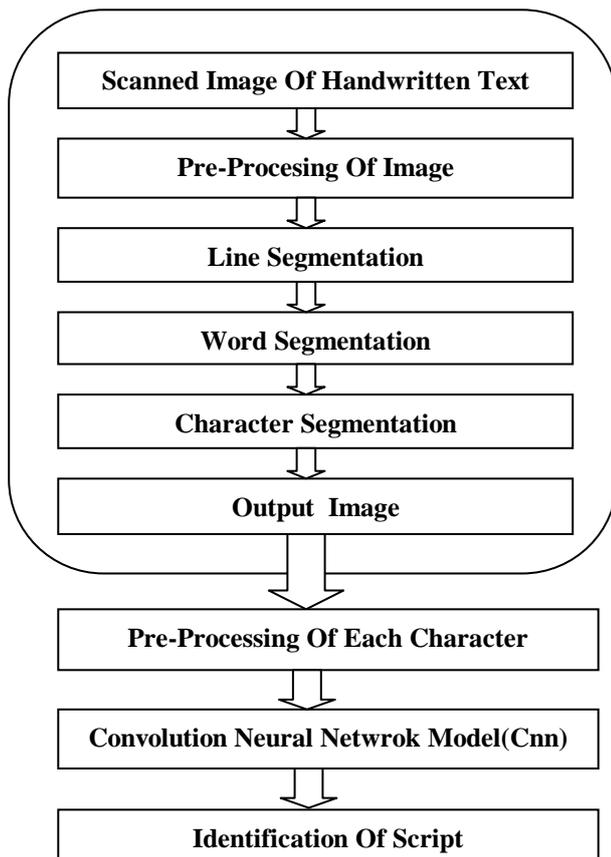
Fig.4. Process For Identification Of Script

In our approach, we need to train three CNN models- one for language classification , one for Devanagari classification and one for Roman classification.

For language classification model, we have combined both the datasets(Devanagari and Roman Dataset) to train our model. We have used 90% data for training purpose and 10% for testing purpose. Now our target variables are Roman and Devanagari. We trained our Multilayered Deep CNN using approximately 1,20,000 images and tested it for approximately 2000 images. Our next step is to identify a character from the segmented image. As a result of our previous model, Output can be either Roman or Devanagari, depending upon the output further trained classification model will be used.

For Roman character classification again we have trained our DCNN model using open source dataset consisting of 71,000app. Handwritten images of characters and corresponding for Devanagari character classification dataset consists of 72,000app. Handwritten images of characters. Our DCNN model is trained with 3 different datasets for each dataset we hyper tuned it to achieved maximum accuracy for each classification process.
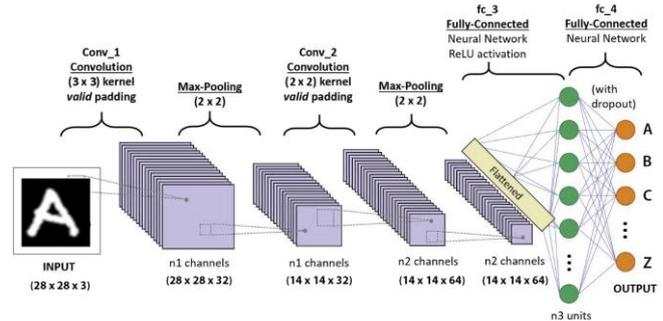

Fig.5. A CNN Sequence To Classify Handwritten Roman Character
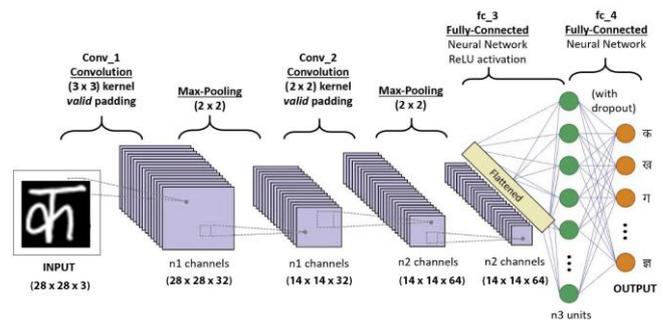

Fig.6. A CNN Sequence To Classify Handwritten Devanagari Character

V. RESULTS

In order to segment and identify the character from handwritten bilingual scripts. We have trained DCNN models. In very first phase, the model is

trained to identify language of handwritten text and achieved accuracy of 98%. The training summary of the model is shown in fig.7 .which represents training loss ,validation accuracy and validation loss. After identification of languages, our next step is to identify character from the image and depending upon the result of the previous step further model will be loaded (if identified language is Roman script then input image will be referred to the trained English classification model and Hindi classification model for Devanagari script).For this purpose our DCNN is trained again with Roman character dataset and Devanagari character dataset and accuracy of 93% and 87% is achieved. The training summary of both the models are shown in fig.8 and fig.9 .
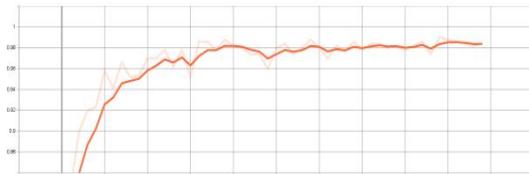
Fig.(a) Language Accuracy

Fig.(b) Language Validation Accuracy
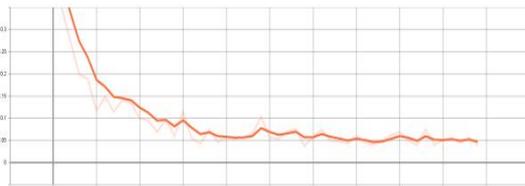
Fig.(c) Language Loss
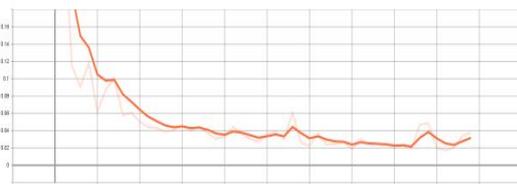
Fig.(d) Language Validation Loss

Fig. 7. Language Classification Mode
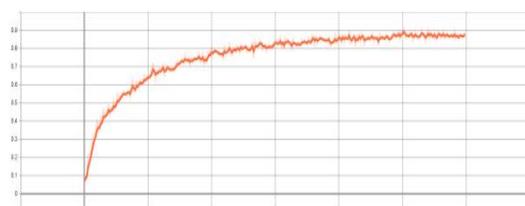
Fig.(a). English Accuracy
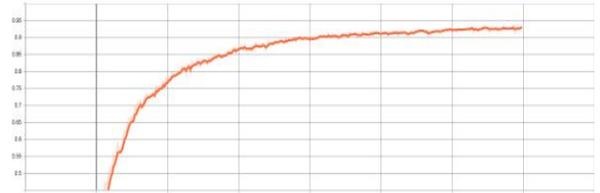
Fig.(b). English Validation Accuracy
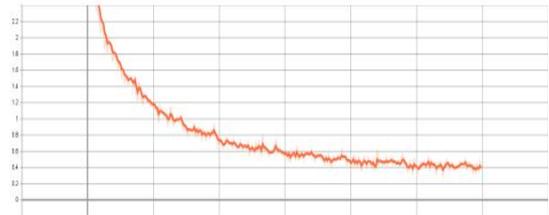
Fig.(c). English Loss Accuracy
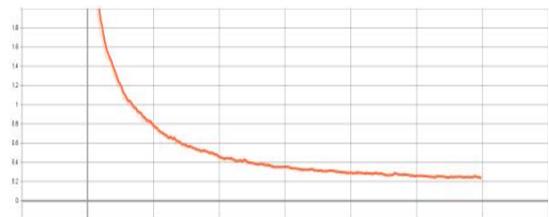
Fig.(d). English Validation Loss

Fig. 8. English Character Classification

Fig.(a). Hindi Accuracy
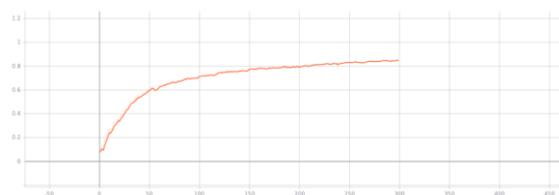
Fig.(b). Hndi Validation Accuracy
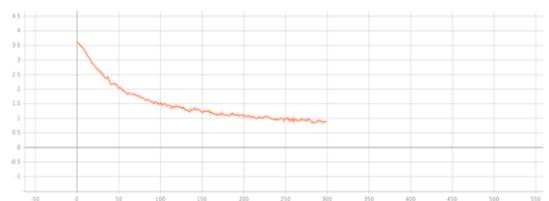
Fig.(c). Hindi Loss
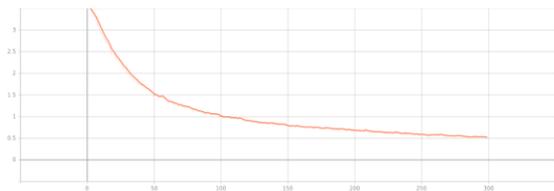
Fig.(d). Hindi Validation Loss

Fig.9. Hindi Classification Model

VI. CONCLUSION AND FUTURE WORK

In this paper, we represents a novel approach for the recognition of handwritten characters that implements Image processing, Multi-layer Deep Convolution Neural Networks with some regularization techniques, namely dropout and batch normalization and apply it using keras with backend Tensorfow framework. It contains approximately 1,40,000 images with which will be the largest publicly available dataset English and Hindi text in natural images. We train powerful and robust modules. The additional regularization parameters lead to better outcomes with regard to the accuracy. The implementation was fast using keras functions and it took only 90s for each epoch on a PC with Intel i7 processor, 16GB RAM and GPU.

In future, it will be possible to implement this approach, For multiple languages from different regions of India for both Handwritten and Printed Text .

ACKNOWLEDGMENT

REFERENCES

[1] S. Srivastava, J. Priyadarshini, S. Gopal, and S. Gupta, *Optical Character Recognition on Bank Cheques Using 2D Convolution Neural Network*. Springer Singapore, 2019.

[2] A. Choudhary, R. Rishi, and S. Ahlawat, "New character segmentation approach for off-line cursive handwritten words," *Procedia Comput. Sci.*, vol. 17, pp. 88–95, 2013.

[3] U. Pal and B. B. Chaudhuri, "Indian script character recognition : a survey," vol. 37, pp. 1887–1899, 2004.

[4] S. B. Dhok, "Multilingual Character Segmentation and Recognition Schemes for Indian Document Images," *IEEE Access*, vol. 6, no. ii, pp. 10603–10617, 2018.

[5] A. A. M. Al-saffar, H. Tao, and M. A. Talab, "Review of Deep Convolution Neural Network in Image Classification," pp. 26–31, 2017.

[6] P. Nirwan, "A Survey Of Segmentation Of Characters In Handwritten Devanagari Script," vol. 5, no. 8, pp. 710–712, 2018.

[7] P. Sharma and M. K. Sachan, "A Technique for Character Segmentation in Middle zone of Handwritten Hindi words using Hybrid Approach," no. July, pp. 1–10, 2017.

[8] A. Priya, S. Mishra, S. Raj, S. Mandal, and S. Datta, "Online and offline character recognition: A survey," in *International Conference on Communication and Signal Processing, ICCSP 2016*, 2016, pp. 967–970.

[9] D. Q. G. X. Hh *et al.*, "Dense Prediction for Text Line Segmentation in Handwritten Document Images," *Icip*, 2016.

[10] J. Wu, Y. Yu, C. Huang, and K. Yu, "Deep multiple instance learning for image classification and auto-annotation," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol.

07–12–June, pp. 3460–3469, 2015.

[11] "Deep Convolutional Neural Networks for Sentiment Analysis of Short Texts," pp. 69–78, 2014.

[12] T. Bluche, H. Ney, and C. Kermorvant, "TANDEM HMM WITH CONVOLUTIONAL NEURAL NETWORK FOR HANDWRITTEN WORD RECOGNITION CNRS , Spoken Language Processing Group c RWTH Aachen University , Human Language Technology and Pattern Recognition b LIMSI a A2iA," pp. 2390–2394, 2013.

[13] H. Ney and C. Kermorvant, "Feature extraction with convolutional neural networks for handwritten word recognition," 2013.

[14] G. Singh and M. Sachan, "Multi-layer perceptron (MLP) neural network technique for offline handwritten Gurmukhi character recognition," in *2014 IEEE International Conference on Computational Intelligence and Computing Research, IEEE ICCIC 2014*, 2015, pp. 1–5.

[15] M. Kumar, M. K. Jindal, and R. K. Sharma, "Segmentation of Isolated and Touching Characters in Offline Handwritten Gurmukhi Script Recognition," *Int. J. Inf. Technol. Comput. Sci.*, vol. 6, no. 2, pp. 58–63, 2014.

[16] T. B. Laboratories and U. De Montrhal, "Word-Level Training of a Handritten Word Recognizer Based on Convolutional Neural Networks," pp. 88–92, 1994.

[17] V. Kumar and P. K. Sengar, "Segmentation of Printed Text in Devanagari Script and Gurmukhi Script," *Int. J. Comput. Appl.*, vol. 3, no. 8, pp. 24–29, 2010.

[18] M. S. Nehra, N. Nain, and M. Ahmed, "Benchmarking of text segmentation in devnagari handwritten document," in *2016 IEEE 7th Power India International Conference, PIICON 2016*, 2017, pp. 0–3.

AUTHORS PROFILE

Priyanka Nirwan received B.Tech Degree in Computer Science & Engineering from DAV University, Jalandhar in 2016 and pursuing M.Tech from Chandigarh University. Her research ares are Natural Language Processing, Image Procesing and Handwriting Recognition.

Gurpreet Singh received B.Tech. Degree in Computer Science & Engineering from Punjab Technical University, Jalandhar in 2007 and M.Tech Degree in Computer Science & Engineering from BBSBEC in 2011. He worked as Lecturer in the Department of Computer Science & Engineering, IET, Bhaddal from 2007 to 2010. He worked as Assistant Professor in the Department of Computer Science & Engineering, IET, Bhaddal from 2010 to 2014. He worked as a Research Scholar in the CSE Department at SLIET, Longowal from 2014 to 2018. Currently, He is working As AP in CSE department at Chandigarh University since 2018. His research areas include Digital Image Processing, Natural Language Processing and Handwriting recognition. He is a lifetime member of Indian Society for Technical Education