

# CNN Based Approach for Offline Text Detection

Anju Bala, Surinder Kaur, Gurpreet Singh

**Abstract:** Offline text detection is a research area where the scanned images containing printed text have been considered as input and then processed to convert the text into digital form that can be understandable and further manipulated by the computer system. In these applications the major difficulty is to segment the area containing text and then after segmentation extracting the meaning of symbols present in the segmented area. The later step is known as classification of characters and words. In this paper a system has been proposed and implemented based on convolution neural network (CNN) classifier to extract the handwritten text from the scanned images. The targeted text has been considered to be written with the help of Roman script means English language. Classifier performed well and produced results at character level up to 97% perfection.

**Keywords:** Roman, CNN, English, Text Detection, Classifier

## I. INTRODUCTION

Offline text detection and extraction has been considered as an area where scanned images have been considered as the input. These images contain printed text. The major processing performed by these systems is to convert the printed text to the digital text. This digital text further can be manipulated or edited in any way. The main reasons for the conversion of text from scanned images to digital is to produce the raw text to perform operations like searching, insertion, deletion, update etc. The alternate to these types of offline systems is the use of human as typist to produce the same kind of output. But this kind of process may contain human error as typographic mistakes, own perception of typist about the readability of the text etc. To avoid these kind of issues an automated system is required which can produce the same output by ignoring the human error. Here, there is a need to provide some learning to the machine. So that the machine can follow the steps as per the learning provided and produce results based on the logics fed to the system. The major issues related to text identification and recognition systems are the existence of multiple languages around the world the presence of multiple scripts to follow these languages at the time of writing. These scripts have different character sets also. So, there is a need to follow a specific script engine while designing a text recognition system. In this paper the focus has been given to the Roman script. As the system has been designed to convert the printed text written using English language. During the classification phase Convolution Neural Network (CNN) has been used. The strength of CNN to work with machine learning applications has been the reason for using CNN as classifier.

**Manuscript published on 30 June 2019.**

\* Correspondence Author (s)

**Anju Bala**, Assistant Professor in Chandigarh University, Computer Science Engineering, Gharuan

**Surinder Kaur**, Assistant Professor, Department of Computer Science & Engineering, Chandigarh University, Gharuan, Punjab

**Gurpreet Singh**, Assistant Professor, Department of Computer Science & Engineering, Punjab Technical University, Jalandhar

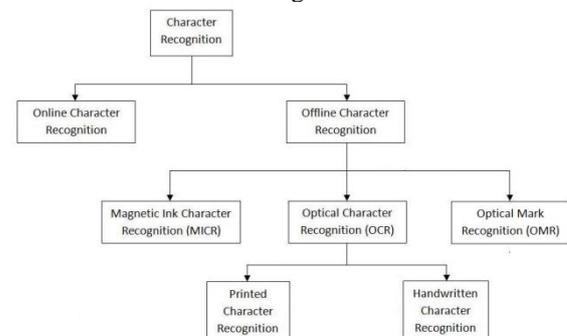
© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

The next phase of the paper will cover the basic approach to be followed during automatic character recognition applications. After that the related work in this field of text recognition has been discussed to find the state of the art work done the specific field. Then it is followed by a section explaining the type of CNN architecture considered. After that, Results and discussion section will present the output generated by the system. Conclusion part will cover the aspect of the overall work done to consolidate the idea taken under consideration.

## II. CHARACTER RECOGNITION

Mainly the character recognition process has been classified into following two categories:

- Offline character recognition
- Online character recognition



**Fig. 1. Types of character recognition system**

Both categories of the system can be differentiated on the basis of the way in which the input is supplied to the automated system. The input in case of offline systems are always in the form of digital text images and the input in case of Online systems are always in the form of tablet pc strokes samples considered with the help of stylus or with the use of digital pens. Fig.1. also shows the further categories where offline processing can be helpful. It shows that in case of offline processing systems we can consider both printed text or handwritten text for the purpose of automatic text detection. In this paper an offline handwritten text detection approach has been addressed. For writing the content hide by the system Roman script is used. The reason of normal script usage is the consideration of English text.

## III. RELATED WORK

Apurva Srivastav and Jayant kumar [1] proposed a character recognition system. They focused to capture the edges used in the formation of text regions and the ROI where text is present in the image. The overall recognition accuracy achieved by the authors has been considered as 96.38% . Junsik at el. [2] also worked over an offline system. They generated and algorithm to trace the text written on wine images during the execution of digital processing application.



Fuzzy means clustering classifier had been used by the researcher, that performed considerably.

Jing Song at el. [3]. worked on a system of capturing the text written in image CAPTCHA. The difficulty of the task was the multiple back ground images. This complexity was too high but still the system performed considerably well. Shangqin at el. [4]. proposed an offline system used to detect the text written on license plates. They used the approach of wavelet transform to complete this task. They also considered the use to various color models based on the type of requirement faced. Apart from the above mentioned application models of offline handwriting recognition systems many other papers have been considered under the literature review to give shape to the work presented in this paper[5]-[21]. This proposed work also falls in to the category of offline handwriting recognition.

IV. CNN MODEL

During the classification process, a deep net model has been used. This model has been based of the convolution neural network (CNN). This model has been explained with the help of Fig. 2. This represents the entire model and the use of CNN classifier. Input provided to the system is a scanned image containing English text. This image has been pre-processed by using the functions Resizing and normalization to provide images in the better quality for further process. Then a CNN classifier with six hidden layers and ReLU function has been used for the purpose of text identification. Equation 1. represents the ReLU activation function. The CNN worked in two different steps in the first one, first three hidden layers of the CNN are using 5X5 filter mask for the purpose of convolution and in the next step three other layers of CNN are using 3X3 convolution mask to process the image and to provide idea the idea about the text hidden in the image or to convert the handwritten text to the digital form.

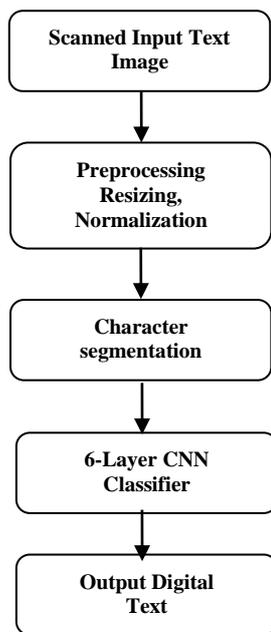


Fig2. Proposed system with CNN Classifier

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}$$

Equation1.

Digital scanned image has been considered as input to the system. To improve the quality of the system or the quality of the input of the system pre-processing functions like resizing and Normalization of image coordinates have been performed. After this all, the next part is segmentation to find out the proper region of interest where the text representing character can be located. This has to be done during the segmentation phase. For the purpose of extracting ROI, horizontal and vertical projection techniques have been used. The output of the segmentation phase is the production of various sub images containing character impressions. These sub images further have been assigned to the proposed CNN model for classification and final recognition process.

V. RESULTS AND DISCUSSION

Table1. shows the output of segmentation phase of the proposed system. Firstly, the image containing handwritten text has been displayed. The next row contains the extracted characters from the input image by using horizontal and vertical projection approach. The last row of Table1. presented the output formed with the help of CNN Classifier in the form of Digital text. This digital text is the final representation of the text present in the input image.

Table1. Segmentation results sample text

<b>Input Sample</b>
E X T R A
<b>Segmented Characters</b>

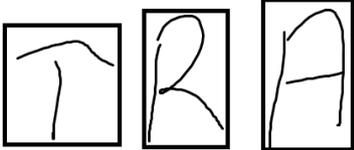

<b>Digital Output</b>
EXTRA

Table2. shows the accuracy chart of segmentation phase and the recognition phase of the proposed system at the level of different characters present in the roman script character set.

Table2. Recognition accuracies

S.No	Character	Segmentation Accuracy	Recognition Accuracy
1	A	100	100
2	B	100	95.5
3	C	98	96
4	D	98	97.8
5	E	97	99
6	F	99	98
7	G	100	100
8	H	99	100
9	I	98	100
10	J	96	97
11	K	100	96
12	L	100	98
13	M	100	99
14	N	99	95
15	O	100	100



16	P	99	100
17	Q	97	97
18	R	98	95
19	S	100	99
20	T	100	100
21	U	99	98
22	V	100	97
23	W	100	94
24	X	100	98
25	Y	98	98
26	Z	97	94

The segmentation phase overall contains the accuracy of 98.92% and the recognition phase shows the accuracy level of 97.64%. These levels of accuracies are considerable of the valuable applications.

## VI. CONCLUSION

The application of handwriting recognition system have been addressed. For writing text or collecting handwriting samples Roman script has been considered. This script is used to write the very popular language "English". The system has been considered as offline system. This means, scanned images containing text has been assigned to the system as input. Pre-processing techniques used for improving the quality of the input were resizing and normalization. Then horizontal and vertical projection techniques have been used for line detection and word detection as ROI. After this, the active regions have been passed as images to the CNN classifier to extract the character present in each image. The overall accuracy observed from the system is 97.64%.

## VII. FUTURE SCOPE

The work done in the considered system is offline work where, scanned images have been used. This work further can be extended to the Online handwriting recognition task by considering the input in the form of digital pen strokes. That can be a totally different task. But in those kind of applications again the approach like CNN is required. Further the structure of Deep nets can be extended by changing the model of the artificial neural network used. This can be done by increasing the number of hidden layers or by using hybrid nets in the model like CNN in combination with RNN (Recurrent Neural Network) etc. This kind of new challenges will definitely increase the efficiency of the system along with the Online ability.

## REFERENCES

1. Srivastav,A., Kumar,J.,2010, "Text Detection in Scene Images using Stroke Width and Nearest-Neighbor Constraints"
2. Lim,J., Kim,S., Park,J., Lee,G., Yang,H., Lee,C.,2009, "Recognition of Text in Wine Label images"
3. Cui,J., Wang,L., Mei,J., Zhang,D., Wang,X., Peng,Y., Zhang,W.,2010, "CAPTCHA Design Based on Moving Object Recognition Problem"
4. Mao,S., Huang,X., Wang,M.,2010,"An Adaptive Method for Chinese License Plate Location"
5. Kumar,M., Kim,Y.C., Lee,G.S.,2010, "Text Detection using Multilayer Separation in Real Scene Images"
6. Al-Shamma,S.D., Fathi,S.,2010,"Arabic Braille Recognition and Transcription into Text and Voice"
7. Ikica,A., Peer,P.,2011,"An improved edge profile based method for text detection in images of natural scenes"
8. Raj,S.B.E., Devassy,D., Jagannivas,J.,2011, "A New Architecture for the Generation of Picture Based CAPTCHA"
9. Wazalwar,D., Oruklu,E., Saniie,J.,2011,"Design Flow for Robust License Plate Localization"
10. Shivakumara,P., Phan,T.Q., Shijian Lu, Tan.C.L.,2011,"Video Character Recognition Through Hierarchical Classification"

11. Chang,R.C.,2011,"Intelligent Text Detection and Extraction from Natural Scene Images"
12. Zhang,S., Li,H.,2011, "Video Semantic Mining based on Dense Sub Graph Finding"
13. Oi-Mean Foong, Hairuman,I.F.B., 2011, "OCR Signage Recognition with Skew & Slant Correction For Visually Impaired People"
14. Wang,X., Lai,W., 2010, "Edge Detection for Chinese Text Image Based on Novel Differential Operator"
15. Huang,X., 2012, "Automatic Video Text Detection and Localization Based on Coarseness Texture"
16. Bo Bai, Fei Yin, Cheng-Lin Liu ,2012,"A Fast Stroke-Based Method for Text Detection in Video"
17. Shih-Chung Chen, Chung-Min Wu, Shih-Bin Su, 2012,"Image Morse Code Text Input System"
18. Mello,C.A.B., Costa, D.C., Santos,T.J.D., 2012, "Automatic Image Segmentation of Old Topographic Maps and Floor Plans"
19. Zhang,Y., Wang,C., Xiao,B., Shi,C., 2012, "A New Text Extraction Method Incorporating Local Information"
20. Yang,Z., Shi,P.,2012,"Caption Detection and Text Recognition in News Video"
21. Rajeshbaba, M., Anitha, T., 2013, "Detect and Separate localization Text in Various Complicated-Colour Image"

## AUTHOR'S PROFILE



**Anju Bala** received B. Tech degree in computer science Engineering from punjab technical University in 2006 and M.Tech degree in computer science from punjab university chandigarh in 2014. she is working as assistant professor in chandigarh university gharuan since 2012. Her research area is image processing and optimization of algorithm.



**Surinder Kaur** received B.E. Degree in Computer Science & Engineering from Punjab Technical University, Jalandhar in 2000 and M.Tech. Degree in Computer Science & Engineering from GNDEC in 2008. She worked as Lecturer in the department of Computer Science & Engineering ,IET Bhaddal ,Punjab from 2001 to 2008. She worked as Assistant Professor in the department of Computer Science & Engineering , IET,Bhaddal, Punjab from 2008 to 2009. Currently, she is working as Assistant Professor in the department of Computer Science & Engineering , Chandigarh University, Gharuan ,Punjab from 2009. Her research areas include Digital Image Processing and Network Security.



**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, Punjab from 2007 to 2010. He worked as Assistant Professor in the Department of Computer Science & Engineering, IET, Bhaddal from 2010 to 2014. He was a Research Scholar in CSE Department at SLIET, Longowal from 2014 to 2018. Currently, he is working as an Assistant Professor in Chandigarh University, Punjab. His research areas include Digital Image Processing, Handwriting recognition and Natural Language Processing. He is a lifetime member of Indian Society for Technical Education.