Real - Time Text Reader

Jayshree R. Pansare, PhD MES College of Engineering SPPU, Pune, India

> Priyanka Karne MES College of Engineering SPPU, Pune, India

Aditi Gaikwad MES College of Engineering SPPU, Pune, India Vaishnavi Ankam MES College of Engineering SPPU, Pune, India

Shikha Sharma MES College of Engineering SPPU, Pune, India

ABSTRACT

In most character recognition systems like Optical Character Recognition(OCR), the system may not work well in case of handwritten documents, documents with poor contrast, or when the text and the background are similar in darkness. In some circumstances, the presence of the aforementioned cases leads to poor character recognition. This paper presents a Real Time Text-Reader which works for scanned images and videos. Additionally, the system also extracts text from digital comic images. The system works in 5 phases which are acquirement of the image, pre-processing on image, segmentation, feature extraction, word extraction. It then tags the words into their respective parts of speech categories.

Keywords

Optical Character Recognition (OCR), Word extraction, Parts of Speech(POS) Tagging, Deep Neural Network(DNN)

1. INTRODUCTION

Optical Character Recognition is the automated transformation of Published or handwritten manuscript to the machine editable format whether it will take from scanned document or photo of the document. OCR is a research field in artificial intelligence, pattern identification also for the vision of computer. Text categorization in videos and images can be used to extract relevant videos and frames from the text on signs of roads, billboards and so on. According to [1], a Flow of Gradient Vector based on Method to video type-set is specified. works in three steps to identify pixels, from top to bottom row of image identifying multiple least cost paths, false cuts are removed during the middle of characters. This method improves binarization results. Text extraction using two blobs extraction method in digital English Comic Images is given[2]. Initially, the balloon in the image is detected and then the horizontal text is extracted. The system correctly detects the text blob, extract the text using a median filter giving 94.82 percentage text extraction ratio. Character recognition system replicates humanly understandable manuscript into device programmed form[3]. Characters Recognition in images using shared- hidden-layer deep convolution neural network (SHL-CNN) is done[4]. A text detection method for horizontally aligned text with the complex background using Optical Character Recognition(OCR) is presented[5]. Text characterization in

videos using detection, localization, extraction methods[6]. Discovery and identification of text in videos and images using a semi-automatic system which generates ground truth values for text is proposed [7]. A standard database for recognition of handwritten text is described [8]. In offline handwritten recognition field, the performance of recognition for isolated words is previously notably subordinate. This can be achieved through syntax examination module on the basis of stochastic context-free grammar [9]. A new approach based on BP network, statistical feature and structural feature combination for polluted English character recognition are discussed[10]. Character recognition for English and Kannada using OCR method based on region features is proposed in[11]. Machine Learning with text recognition for self programming of machines according to instructions is discussed[12]. A new system of diagonal based quality removal using feed forward network and convolution neural network is given[13]. Deep neural based name entity recognizer provides recognition of name from digital documents[14].

This paper is organized into following Sections: Section 2. consists of the associated work of the Real-Time Character Recognition sys-tem (RTCR). Section 3. contains flow for detection of text from comic images. In Section 4, we present the Comparative Performance of Real-Time Text Reader System based on the performance of specific systems. We conclude with the conclusion in Section 5.

2. RELATED WORK

The methodology used in the text reader is described in the following figure which is divided into 5 phases which are used for text recognition from images, videos.

2.1 Acquirement of Image

The first stage of any recognition system is the image acquisition stage. The process begins with the acquirement of the image from the input. The user provides the input in the form of sentence or image. This sentence or image is further classified into the name of the location, person or organization on the basis of users select. The text acquisition contains animate streaming of uninterrupted images (videos), text revealing, text cropping, distinguishing and converting into a text file.

International Journal of Computer Applications (0975 – 8887) Volume 182 – No. 34, December 2018



Fig 1: Real-Time Character Recognizer System Architecture

2.2 Pre-processing on Image

Image pre-processing is necessary for the improvement of image data. It also helps to suppress the unwanted distortions present in the image. Some image features required for further processing are also enhanced.

Pre-processing has shown some positive possessions on the excellence of feature extraction and also on the results of image analysis. It deals with the improvement of the quality of the image for better recognition by the system. Image processor constantly takes the frame that is called video streaming. Text revealing box will identify the frame which consists of comprehensible text. To tie the processing speed it requires a delay between two frames. There is some basic operation of Pre-processing such as edge detection, noise decline, binarization, dilation, and filling.

2.2.1 Noise Decline

Some algorithms are used to blur some part of the image which is not required. Noise declining involves morphological and filtering operations.

2.2.2 Edge Detection

In images, the edge detection algorithm is used for identification of borders of objects. This contains the Sobel method, canny technique, Prewitt method to perform edge identification.

2.2.3 Binarization

Binarization is nothing but a conversion of grey scale image into a binary image. For binary image there are two colours are used which are black and white. In binarization thresholding method is used in which the pixels which are com-parable in grey scale are dumped into one group.

2.3 Segmentation

It is one of the main phases of the proposed system model. It includes complex text is divided into sentences. In many languages, punctuations are used to divide the text but this technique is not sufficient for all the text. Some examples are an exception for this technique. Consider an example, "Ms." is not considered as a sentence but because of this, it is divided into individual sentences. So this problem can be avoided by providing the list of abbreviations. Segmentation is required to characterize the characters.

2.4 Feature Mining

Feature mining is related to dimensionality reduction. In Image processing, various portions of digitized images or

video streams are detected using algorithms. Optical character recognition is also particularly based on the idea of feature mining.

2.5 Word Extraction

Word extraction is an option to extract individual words when the text consists of multiple words. Each word is stored and processed individually. A sentence is tested for a new word every time there is pen-up. At this point, the following word extraction parameters are evaluated to decide whether a new word has started and whether the collection of sampling points since the previous pen-up or start of the recording could be considered a new word. There are three aspects of word extraction.

- Mysterious colour precedence; that means text is dark or light.
- Various stroke widths
- Complex background

The words are also known as tokens. The process which is used to divide the text into words is also known as tokenization. This can be achieved with the help of String Tokenizer () function in java. Each tokenized word is referred to as an entity. The further POS-tagging technique is used to detect if the word is an entity or not.

2.6 Part-Of-Speech Tagging

For categorizing the tokens into different sets we use cosine similarity approach of the deep neural network (DNN). Tokens are classified into the name of person, organization, place, and department. Consider the example as "Savia works in Minority Department, Gurugram, Harayana." So this system will classify the sentence as follows: Savia as name, Minority Department as an organization, Gurugram and Haryana as the name of a place.

3. HOW TO DETECT TEXT FROM COMIC IMAGES?

Specially, the flow adopted in Real-Time Character Recognizer from Comic Image has been shown in Fig. 2 is fragmented in six consequent parts applied to comic images; Comic image as Input, Pre-processing, Detection of Balloon, Extraction of Text Blob, Recognition of Text, Extraction of Text.



Fig 2: Work Flow of Real-Time Character Recognizer from Comic Image

3.1 Input as Comic Image

Comic Image represents the different scenes, with dialog by using balloons in an art form. For research, the comic image is considered as input which contains English character.

3.2 Pre-Processing on Comic Images

The RGB image is converted into binary by applying the threshold values between 0 to 1. If noise containing an image is taken as input, then pre-processing steps are applied to remove the noise for improving the efficiency of text extraction. The median filter is used for noise reduction.

3.3 Detection of Balloon

Balloon detection is necessary for text extraction from comic images. CCL algorithm is applied on noise-removed image for detecting the connected components in images. First regions boundaries are detected using CCL algorithm. Balloon detection is needed for separating the text blob and non-text blob detection.

3.4 Extraction of Text Blob

After applying the CCL algorithm we get the script blob and non-script blob. Blob extraction is needed for categorizing the script blob from non-script blobs. By using features of blob size identification is done. If the part is 10 percentage(8 percentage) of the unique image then it is categorized as a script blob and another is a non-script blob. By using the equation (a) the part of the blobs are planned.

X.TB[i] =TB[j].Width*TB[j].Height (a)

3.5 Recognition of Text

The OCR is executed to know the text from the balloon after text blob extraction. The OCR process involves various phases like classification, feature extraction, and segmentation. Then lastly text characters are kept in a text file.

4. COMPARATIVE PERFORMANCE OF REAL-TIME TEXT READER SYSTEM

This section emphasizes the comparative performance of a diversified system applied for detection of various characters as shown in Graph 1. It is observed that for recognizing English character SVM classifier achieves the highest accuracy of 96.13 percentage. More-over, Independent text recognizer achieved the lowest recognition rate of 49.1 percentage.



Accuracy in percentage

Fig 3: Comparative Study For Text Reader

5. CONCLUSION

In this paper, an innovative method for real-time character recognition using Deep Learning Algorithm has been presented. The pro-posed system acquires the text from the input, pre-processes it using the 5 phases mentioned in the methodology and then tags it into their parts of speech category. The inspection time required by traditional methods it reduces. Further efforts will be dedicated to improving their liability and exactness of the proposed system. The actual extra efforts will be required to develop a similar system for other regions as well as foreign languages other than the English language.

6. REFERENCES

- [1] T. Q. Phan et.al., "A Gradient Vector Flow-Based Method for Video Character Segmentation", International Conference on Document Analysis and Recognition, pp:-1520-5363. doi:-10.1109/ICDAR.2011.207.
- [2] Dr.M.Sundaresan, et.al., "Text Extraction from Digital English Comic Image Using Two Blobs Extraction Method", Proceedings of the International Conference on Recognition, Informatics and Medical Pattern Engineering (PRIME), pp.449-452, 2012.
- [3] H. Mehta, et.al., "Optical Character Recognition (OCR) Sys-tem for Roman Script and English Language using Artificial Neural Network (ANN) Classifier", IEEE International Conference on Research advances in Integrated Navigation Systems (RAINS- 2016), R. L. Jalappa Institute of Technology Doddaballapur, Bangalore, India, 2016.
- [4] J. Bai, et.al., "Image Character Recognition Using Deep Convolutional Neural Network Learned From Different Languages", IEEE, pp.2560-2564, 2014.
- [5] M. R. Lyu, et.al., "A Comprehensive Method For Multilingual Video Text Detection, Localization, and Extraction", IEEE Transaction On Circuits And Systems For Video Technology, vol.15, no.2, pp.243-255, 2005. doi:10.1109/TCSVT.2004.841653
- T. Q. Phan, et.al., "Semiautomatic Ground Truth Gen-[6] eration for Text Detection and Recognition in Video Images", IEEE Transaction On Circuits And Systems For

Video Technology, vol.24, no.8, pp.1277-1287,2014. doi:10.1109/TCSVT.2014.2305515

- [7] J. J. Hull "A Database for Handwritten Text Recognition Research", IEEE Transactions on Pattern Analysis and Ma-chine Intelligence, Vol.16, no.5. pp:-550-554,2017.doi:0162-8828/94
- [8] M. Zimmermann, et.al. , "Offline Grammar-Based Recognition of Handwritten Sentences", IEEE Transaction On Pattern Analysis And Machine Intelligence, Vol. 28, No. 5. pp. 818 ? 821, 2006.
- [9] Y. Yang, et.al., "English Character Recognition Based on Feature Combination", 2011 International Conference Advances in Engineering, pp.159-164,2011. on doi:10.1016/j.proeng.2011.11.2619
- [10] G. Mukarambi, et.al. , "A Zone-Based Character Recognition Engine For Kannada And English Scripts", pp.3292 ? 3299, 2012.
- [11] S. L. Wasankar, et.al., "Machine Learning with Text Recognition", doi:10.1109/ICCIC.2010.5705811
- [12] U. Yadav, et.al., "A Deep Learning Based Character Recognition System From Multimedia Document", International Conference on Innovations on Power and Advanced Computing Technologies.pp.1-7,2017
- [13] S. P. Singh, et.al., "Deep Neural Based Name Entity Recognizer and Classifier for English Language", Proceeding of Second International Conference on Circuits, Controls and Communications, pp. 241 ? 246, 2017.
- [14] A. D. Amensisa, et.al., "A Survey on Text Document Categorization Using Enhanced Sentence Vector Space Model and Bi-Gram Text Representation Model based on Novel Fusion Techniques", IEEE Xplore Proceedings of the Second Inter-national Conference on Inventive Systems and Control(ICISC 2018), pp.218-225,2018.
- [15] J. R. Pansare and M. Ingle, "Comprehensive Performance Study of Existing Techniques in Hand Gesture Recognition System for Sign Languages" Int. J. Computer Sci. Inf. Tech-nol., vol. 7, no. 3, pp. 1343?1347, 2016.