Text Extraction In Natural Images And Converting To Speech

Soundarya D karningannavar[1], Varsha S Jadhav[2]
[1]Dept of ISE
[2]Assistant Professor, Dept of ISE
[1, 2] SDMCET Dharwad

Abstract- *The detection and extraction of text in natural scene images and the extracted text is converted into speech. Extracting text character from natural scene images is a challenging problem due to differences in text style, font, size, orientation, alignment and complex background and received a major attention in recent researches. Increased number of applications in devices like mobile phones, tablets and so on. In this paper we propose a suitable existing machine learning algorithm for text region localization and character extraction in natural scene images with complex background and noise and converting the text into the speech. Recognition of text from natural image is performed by using image processing and neural network.*

Keywords- Character extraction, Region localization, Noise.

I. INTRODUCTION

Use of smart phones and digital cameras become very common and digital cameras have become very popular, efficient and image acquisition tools, which are commonly attached with various handheld devices like mobile phones, tablets, pens, and wrist watches, and so on. Fabricators of these devices are now considering such devices for embedding numerous useful technologies. Potential technologies may include detection of texts in natural scene images, text-to-speech conversion and so on [1]. Text is a one of the important means of communication and we all understands through the text. The composite of characters is text, technically it can be said as "structured edges", "series of colour regions", "kind of strokes" or "group of strokes" [6]. The text in the image plays important role in understanding the image. The scene images captured by camera, may contain advertising boards, banners and so on, therefore scene text is embedded in the background as a part of the scene [3]. Extraction and recognition of texts in natural scene images are suitable for persons with visual impairment and travellers when they face the problem in understanding other language.

The methodology contains mainly two parts the first one is image processing and the second one is neural network. Due to increasing requirement for detection of text from images, many research work has been performed on text recognition from natural images. Many techniques will be proposed for extracting the text from an image. Recognition of text from natural image will be performed by using image processing techniques.

II. PROPOSED SYSTEM

The system is based on Python as well as Java programming language. Java GUI is used for the image retrieval. This system has three phases namely, image processing, text recognition, voice synthesis. In image processing, the image is loaded into the system and pre-processing of the image is performed where noise is removed and the image is converted into gray scale image. The text region detector is designed to detect the text region in each layer of the image. In this the colour image is converted to grayscale image [4]. The morphological operations are applied on the transformed image, to get segmented image. SVM classifier is used to classify text regions and non-text regions [5]. The non-text regions are rejected and this is done by SVM classifier [7]. In text recognition phase the text recognition of our system is based on the neural networks where each dataset is trained to recognize the text and give that as output. Voice synthesis of our system is done using java API. The java API used is Free TTS Speech Synthesis System, which is the free API available in Java. Technologies used for this system are python programming language, neural network specifically back propagation algorithm and java programming language. Python packages used are NUMPY, PILLOW (PIL), PYTESSERACT, OPEN CV. The Back propagation algorithm which is similar to Gauss-Newton algorithm. The Back propagation algorithm searches for the minimum value of the error function in weight space using a technique called the delta rule or gradient descent. The weights that minimize the error function is considered to be a solution to the learning problem. Java programming is used for the purpose of Graphical User Interface and for Voice Synthesis. NETBEANS software is used to build the GUI.

III. RESULT AND DISCUSSION

Screenshot and Explanation:

In this module image is loaded into the system by clicking load image button and by clicking voice button we can hear the voice output.

The Fig. 1. Shown below depicts the initial interface of the software application, where the image is to be loaded from the source point or location.
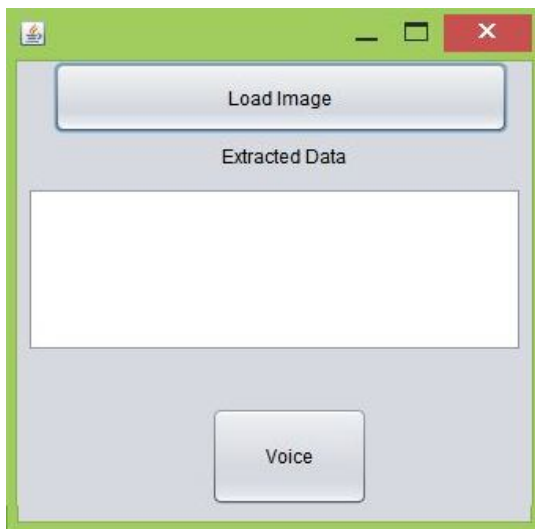

Fig. 1: Image Loading

Image Retrieving: Here the image can be searched and retrieved from remote place of the system.
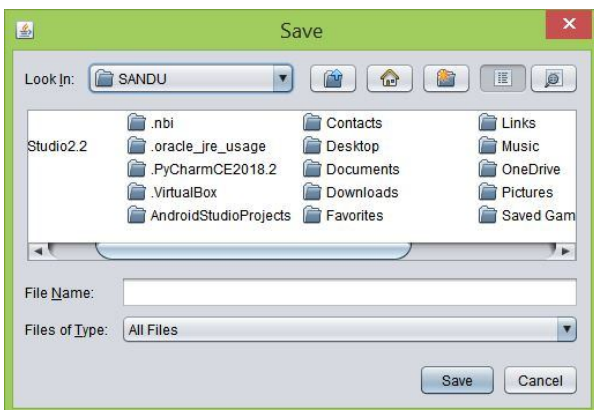

Fig. 2: Image Retrieving

The Fig. 3 is about the Image which may contain all/any format of contents such as Text Alphabets, Numerical symbols etc.
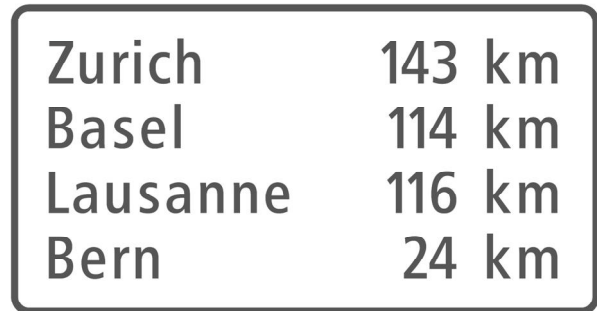

Fig. 3: Input Data to the Software Application

Output: The detected text in the given image is given as output which is shown below figure.
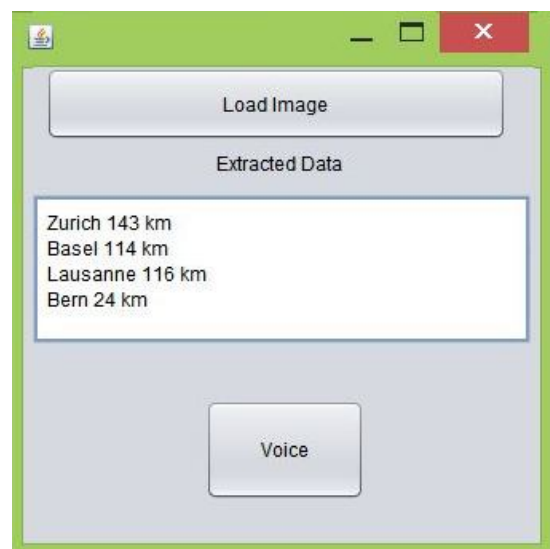

Fig. 4: Output screen

IV. CONCLUSION

In this project the natural scene images will be uploaded into the system and is the image meets all the requirements, it has to scan whole image and detect the texts in the image and convert the same into the voice output. Convolution Neural Network techniques are used to detect the text which uses back propagation algorithm. This approach presents a framework to extract text region in a scene image with heterogeneous background and text with orientation.

REFERENCES

[1] Sivananda V. Seeri, J. D. Pujari, P. S. Hiremath, Text Localization and Character Extraction In Natural Scene Images Using Contourlet Transform and SVM Classifier, I J Image Graphics and Signal Processing.

[2] Taufin M Jeeralbhavi, Dr. Jagadeesh D. Pujari, Sivananda V. Seeri, Text Extraction and Localization from Captured

Images, International Journal on Recent and Innovation Trends in.

[3] Jian Zhang, Kai Wang,Renbong Cheng, Hong Zhao ,Research on the text detection and extration from complex images, IEEE, 2013.

[4] Amruta Pise, S.D.Ruikar, Text Detection and Recognition in Natural Scene Images International Conference on Communication and Signal Processing, Apri 2014.

[5] Chunmei Liu, Text Extraction from Natural Images Based on Stroke Width Map, IEEE, 2013.

[6] Qixiang Ye, David Doermann, Text Detection and Recognition in Imagery: A Survey, IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 37, NO. 7, JULY 2015.

[7] Shabana M A, Alphy Jose, Ani Sunny, TEXT DETECTION AND RECOGNITION IN NATURAL IMAGES, IRJET Jan-2018.

[8] Cong Yao, Xiang Bai, "A Unified Framework for Multioriented Text Detection and Recognition", IEEE Transactions On Image Processing, Vol. 23, No. 11, November 2014.

[9] Hyung Il Koo, "Scene Text Detection via Connected Component Clustering and Nontext Filtering", IEEE Transactions On Image Processing, Vol. 22, No. 6, June 2013

[10] Y.-F. Pan, X. W. Hou, and c.-L. Liu, "A robust system to detect and localize texts in natural scene images," in Proc. 8th IAPR Workshop on Document Analysis Syetems (DAS'08), Nara, Japan, 2008, pp. 35-42.

[11] S. T. Deepa, S. P. Victor, Tamil Text Extraction, International Journal of Engineering Science and Technology (IJEST), ISSN: 0975-5462, Vol. 4 No.05 May 2012, pp. 2176-2179