



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Real Time Text Detection & Recognition Using OCR System to Assist Blind People

Nikita Inamke¹, Shraddha Sanas², Sanket Mungase³, Viraj Nikam⁴

Department of Computer Engineering, NESGOI, Pune, India^{1,2,3,4}

ABSTRACT: Nowadays text recognition is an essential problem for visually impaired peoples. Text-to-speech (TTS) is the generation of synthesized speech from the text. This paper presents a survey on text detection and the conversion of text to speech. There have been many methods for detecting text in past work. This paper reviews the techniques of different methods for text recognition. This synopsis proposes a laptop camera-based assistive text reading framework to help blind persons to read text labels and product packaging from hand-held objects in their daily lives. The system framework consists of three functional components: First, scene capture-using a system camera, the text which the user needs to read gets captured as an image and has to be sent to the image or data processing platform., second, data processing -where text will be filtered from the surrounding and will be recognized by optical character recognition (OCR) software, and finally, Speech output - A filtered text will be passed into this system to get an audio output. In addition to text recognition using OCR, the project also includes template matching technique as a separate method for recognizing certain objects like currency notes. Since the project should be portable to assist the blind persons, the entire application is based on laptop based camera system.

KEYWORDS: Blind, Object, Detection,CNN (Convolutional Neural Network), OCR (Optical character Recognition), Deep Learning.

I. INTRODUCTION

Text recognition is the ability of a machine to receive and interpret handwritten input from multiple sources like paper documents, photographs, touch screen devices etc. Recognition of text on object and machine characters is an emerging area of research and finds extensive applications in banks, offices and industries. Optical Character Recognition (OCR) is a process by which text characters can be input to a computer by providing the computer with an image. The computer

uses an OCR Engine a computer program with the specific function of making a guess which letter (recognizable to a computer) an image (recognizable to a human) represents. The main objective of our system is to detect, extract and recognize text from object, image or hand gesture as per system. To provide an easy user interface to input the object image. User should be able to upload the image. System should be able to pre-process the given input to suppress the background. System should detect text regions present in the image. System should retrieve text present in the image and provide audio as an output to the user.

II. RELATED WORK

There exist multiple tools to use computer vision technologies to assist blind people. The mobile app Tap See uses computer vision and crowd sourcing to describe a picture captured by blind users in about 10 seconds. The Blindsight offers a mobile app Text Detective featuring optical character recognition (OCR) technology to detect and read text from pictures captured from the camera. Facebook is developing image captioning technology to help blind users engaging in conversations with other users about pictures.

Baidu recently released a demo video of a DuLight project. No further details of the product is available at the moment. However, the product video suggests concepts of describing scenes and recognizing people, money bills, merchandises, and crosswalk signal. However, these products were not focusing on enabling general visual sense for blind people and did not use the spatial sound techniques to further enhance the user experience. Some works exist in the general scope of sensory substitution.



Daniel Kish, who are totally blind, developed accurate echolocation ability using “mouth clicks” for navigation tasks including biking and hiking independently .Colorblind artist Neil Harbisson developed a device to transform color information into sound frequencies. An extreme attempt of converting visual sense to sound is introduced by the voice technology.

The voice system scans each camera snapshot from left to right, while associating height with pitch and brightness with loudness. However, all these attempts on sensory substitution are reported with very difficult learning process. In contrast, we utilize visual recognition algorithms which lead to more direct ways of understanding objects from a visual scene.

The use of 3D sound technology for providing useful information and assisting blind people has also been investigated by researchers introduced a system that uses spatial audio to facilitate discovery of points of interest in large, unfamiliar indoor environments (e.g. shopping mall) tries to integrate 3D sound into GPS-based outdoor navigation product. However, no visual recognition has been used in those works. The use of object detection techniques can open up new possibilities in assisting indoor navigation for blind and visually impaired people.

III. PROPOSED METHODOLOGY

In this work, a novel text detection and recognition algorithm is proposed. The system describes about reading of printed letters on the objects for helping the blind people. The problem for the blind persons in reading is solved by this proposed method. This system will distinguish the object from the captured image and detect the texts from the image. The proposed method is demonstrated for localizing text in natural images.

MSER is used to detect the text information from the image. This will extract the text patterns from different backgrounds. This is the robust algorithm for text detection. OCR is used for text recognition on the localized text patterns. The system will transform the localized text to the audio output for blind people. The results point out that it is be possible to achieve high performance using MSER and OCR for text detection and recognition. Two techniques are used to detect the text and recognized that text in this proposed system.

The results are evaluated on the dataset of captured images. This system is instructing the blind users to read the hand held objects. Using this algorithm we are correctly detecting the text in proper area. After the localization of text, these text are recognized using OCR tool. This system is actually based on the reading text from camera based images. The efficiency of the system will be improved by parallel processing of text detection and recognition. The output is giving text as the speech output to blind people.

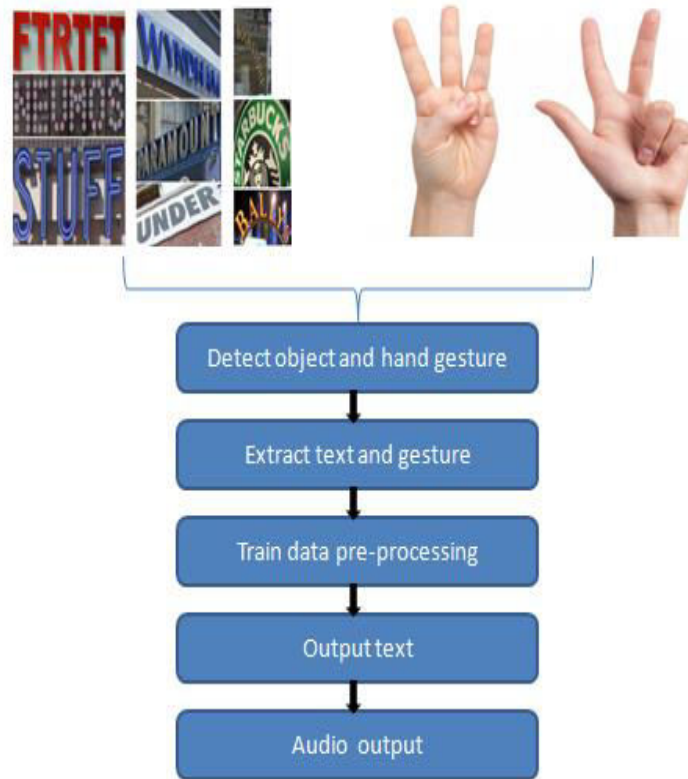


Fig 1. Proposed System Architecture

Optical character Recognition:

Optical character Recognition (NLP) is the process that converts an image of text into a machine-readable text format. For example, if you scan a form or a receipt, your computer saves the scan as an image file. You cannot use a text editor to edit, search, or count the words in the image file. However, you can use OCR to convert the image into a text document with its contents stored as text data.

The OCR engine or OCR software works by using the following steps:

- **Image acquisition**

A scanner reads documents and converts them to binary data. The OCR software analyzes the scanned image and classifies the light areas as background and the dark areas as text.

- **Pre-processing**

The OCR software first cleans the image and removes errors to prepare it for reading. These are some of its cleaning techniques:

1. Deskewing or tilting the scanned document slightly to fix alignment issues during the scan.
2. Despeckling or removing any digital image spots or smoothing the edges of text images.
3. Cleaning up boxes and lines in the image.
4. Script recognition for multi-language OCR technology

- **Text recognition**

The two main types of OCR algorithms or software processes that an OCR software uses for text recognition are called pattern matching and feature extraction.

- **Pattern matching**

Pattern matching works by isolating a character image, called a glyph, and comparing it with a similarly stored glyph. Pattern recognition works only if the stored glyph has a similar font and scale to the input glyph. This method works well with scanned images of documents that have been typed in a known font.



- **Feature extraction**

Feature extraction breaks down or decomposes the glyphs into features such as lines, closed loops, line direction, and line intersections. It then uses these features to find the best match or the nearest neighbor among its various stored glyphs.

- **Post processing**

After analysis, the system converts the extracted text data into a computerized file. Some OCR systems can create annotated PDF files that include both the before and after versions of the scanned document.

IV. WORKING MODULE

The system identifies the hand written English word and generate its corresponding output in English word and the output English word is converted into audio, by using different techniques. The techniques involved are CNN (Convolutional Neural Network), OCR (Optical character Recognition), Deep Learning. In this work, a novel text detection and recognition algorithm is proposed. The system describes about reading of printed letters on the objects for helping the blind people.

The problem for the blind persons in reading is solved by this proposed method. This system will distinguish the object from the captured image and detect the texts from the image. The proposed method is demonstrated for localizing text in natural images. MSER is used to detect the text information from the image. This will extracts the text patterns from different backgrounds. This is the robust algorithm for text detection. OCR is used for text recognition on the localized text patterns.

The system will transform the localized text to the audio output for blind people. The results point out that it is be possible to achieve high performance using **MSER** and **OCR** for text detection and recognition. Two techniques are used to detect the text and recognized that text in this proposed system. The results are evaluated on the dataset of captured images. This system is instructing the blind users to read the hand-held objects. Using this algorithm we are correctly detecting the text in proper area.

After the localization of text, these text are recognized using OCR tool. This system is actually based on the reading text from camera based images. The efficiency of the system will be improved by parallel processing of text detection and recognition. The output is giving text as the speech output to blind people

V. CONCLUSION

In recent years, deep learning has been increasingly popular and it has begun to be applied to various areas ranging from facial recognition to target detection, OCR included. Researchers continuously apply target detection networks of different types such as Faster and RCNN to the OCR area, in order to detect texts. Compared with traditional text recognition, text recognition embedded by deep learning is more accurate with stable noise immunity and robustness. It is able to resist influences such as changes in backgrounds. The end-to-end network of deep learning is more accurate when used in the area of text detection, compared with traditional image processing and cutting, and character classification enabled by machine learning. Moreover, as research progresses, researchers have shifted their focus from everyday document processing to more complex text detection and text recognition in natural scenes, and developed excellent structures such as CTPN and CRNN.

REFERENCES

1. Sumita A. Jakate "Object Recognition App For Visually Impaired,(2020)
2. Yuming He, Research on Text Detection and Recognition Based on OCR Recognition Technology. (2020)
3. ChallaYaswanath, SreeHarshKVS.Object and Currency Detection with Audio Feedback for Visually Impaired. (2020)
4. Rashika Joshi, MinakshiTripathi,Amit Kumar, Object Recognition and Classification System for Visually Impaired. (2020)
5. Joe YaunMambu, Elisa Anderson.BlindReader : An Object Identification Mobile - based Application for the Blind using Augmented Reality Detection. (2019)
6. Sunit Vaidya, Naisha Shah, NaiShah.Real – Time Object Detection for Visually Challenged People. (2020)
7. Zhou, Xinyu, et al. "EAST: An Efficient and Accurate Scene Text Detector." (2017).



|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205215 |

8. Baoguang Shi, Xiang Bai, and Serge Belongie. "Detecting Oriented Text in Natural Images by Linking Segments." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE, 2017.
9. Graves, A., Fernández, S., & Gomez, F. (2006). Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks.
10. Shi, Baoguang, X. Bai, and C. Yao . "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition." IEEE Transactions on Pattern Analysis & Machine Intelligence 39.11(2015):2298-2304.
11. B. Shi, X. Wang, P. Lyu, C. Yao, and X. Bai. Robust scene text recognition with automatic rectification. In CVPR, pages 4168– 4176, 2016



Impact Factor: 8.423



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details