

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 5, May 2019

## Read Assist, Device that Read Text from Images

Anita Kumari Khareshiya<sup>1</sup>, Gaurav Dhande<sup>1</sup>, Ghazala Mirani<sup>1</sup>, G. B. Aochar<sup>2</sup>

U.G. Student, Department of Computer Engineering, MES College of Engineering College, Pune, India<sup>1</sup>

Assistant Professor, Department of Computer Engineering, MES College of Engineering College, Pune, India<sup>2</sup>

**ABSTRACT:** The problems faced by Visually Challenged individual lies in different areas which needs to be addressed. One such difficulty is gaining information from text. Although Braille dialect is present, but all the resources of information in text is not present in this format. Hence a system needs to be designed which can help a person understand the text in front of them. The proposed system helps in solving the problem by clicking the image, containing text, and give audio as the output. The Image is converted into text by Optical Character Recognition Technique (OCR), using Tesseract OCR Engine. After which the text is converted into speech using a Python library.

**KEYWORDS:** Optical Character Recognition, TesseractOCR, Python Library, Braille Dialect

### I. INTRODUCTION

Reading is an activity that we need to perform in our everyday lives. Visually Challenged individuals face this as an issue among many that they face every waking minute of their life. [7] It is estimated that approximately 1.3 billion people live with some form of vision impairment. Emerging technologies and research in taking place round the globe in this domain. [9] There are systems like Eye-Scan Reader that scans a document and user can hear the content on the document, but this can only be used for documents that can be handy and scanned. For reading the text on signboards or computer this system is not feasible. Hence a viable, robust, efficient and cost effective solution is needed. The main goal is to assist these individual in reading which can recognize the text in front of them and give the output in the form of speech. The proposed system is designed to click the image using Raspberry Pi 3 Camera Module and convert the text, in image, to speech.

The work in the paper is divided into two sections 1) Image to text, which contains the pre-processing of the image and then converting into text using Tesseract OCR. 2) Text to Speech which implements the Python Library to convert the text obtained from previous step to speech.

#### Project Scope

1. This project is focused on visually challenged students who wish to pursue higher education.
2. The project has exponential possibilities from reading of normal day to day articles in the form of printed information to reading of large hoardings and public transport route information.

#### Assumptions

1. The users of the system are assumed to have no other disabilities apart from being visually impaired.
2. The user should have the physical ability to wear it and move around in the external environment without any constraints.
3. The user should have complete knowledge about the language (i.e. English) in which the documents are represented in.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 5, May 2019

## II. RELATED WORK

[1] Technology is key for helping visually challenged students studying at university. This reference proposes a method to develop software which performs text-to-speech and speech-to-text conversion which can be used for examination of Visually Challenged individual. The desired result is fully operational web based software for challenged student and teachers. But this is only specific to the computer which has this software installed in it.

[2] Along with this they also face difficulty in understanding text in front of them to obtain information from goods, computer, books, etc. Limited resources are available in Braille. So to provide a viable solution this reference proposes a method in which the image of the text is clicked which is then separated using Optical Character Recognition (OCR) Technique. The text extracted is then processed to get the output as sound.

[3] Although mobile application has also been developed but it is difficult for these people to exactly find the location of text in the environment. To improvise these factors, this reference paper proposes a real-time tracking method by helping them locate text by sound signals.

[4] Along with this ignoring the non-text region is also important while retrieving text from image. This research paper elaborates methods of ignoring the non-text area from the image. This reference identifies the problem faced by the Visually Challenged people and also the difficulty they face while handling the devices made for them. [5] This method proposes a system which implements scene-text detector and tracking it by capturing in video-frames hence providing a robust solution for individuals. The output is provided using speech synthesizer. Along with this avoiding duplicate images and obtaining single result is vital for optimistic result proposed.

## III. PROPOSED WORK

The algorithm consists of two steps, first, Image to text and second, Text to Speech

### Step 1: Image to text

In this step, the proposed device captures the image connected to Raspberry Pi (Rpi) 3. Rpi is used to execute the program using Raspbian OS and Python, OpenCV for implementing the desired processing. The image captured is initially pre-processed. This step involves rotating the image if the image is tilted, applying threshold and performing the final segmentation of whole text.

This image is now fed into Tesseract OCR to obtain Text as the output. Tesseract OCR is an Optical Character Recognition Engine used to convert the image, containing text, to text. The output can be either be displayed on the console or stored in a file for further processing. In our proposed system, we have saved it in a file for further processing

### Step 2: Text to Speech

The text obtained from Step 1 is now converted into audio using Python, pyttsx3 library. This library works offline and helps the user dynamically. This is a feature of the implemented device, that is, it doesn't require Internet connectivity

## IV. EXPERIMENTAL RESULTS

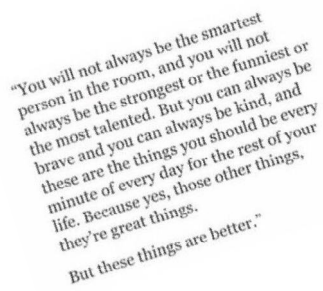
As the image captured by Visually challenged individuals will always be not straight, hence the implemented system rotates the tilted image. Figure shows the results after each step. Fig(a) is the original tilted Image which is rotated to Fig (b). This Figure is then used to apply threshold and further the pixel value is calculated to perform segmentation. This segmented image is then passed to Tesseract OCR to convert it to image.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

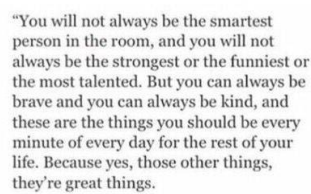
Vol. 8, Issue 5, May 2019



"You will not always be the smartest person in the room, and you will not always be the strongest or the funniest or the most talented. But you can always be brave and you can always be kind, and these are the things you should be every minute of every day for the rest of your life. Because yes, those other things, they're great things.  
But these things are better."

Figure (a)

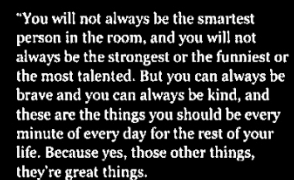
Tilted Image



"You will not always be the smartest person in the room, and you will not always be the strongest or the funniest or the most talented. But you can always be brave and you can always be kind, and these are the things you should be every minute of every day for the rest of your life. Because yes, those other things, they're great things.  
But these things are better."

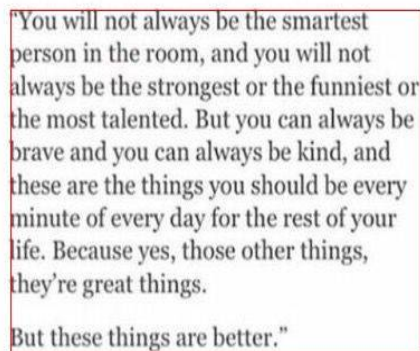
Figure (b)

Image rotated for  
thresholding



"You will not always be the smartest person in the room, and you will not always be the strongest or the funniest or the most talented. But you can always be brave and you can always be kind, and these are the things you should be every minute of every day for the rest of your life. Because yes, those other things, they're great things.  
But these things are better."

Figure (c)



"You will not always be the smartest person in the room, and you will not always be the strongest or the funniest or the most talented. But you can always be brave and you can always be kind, and these are the things you should be every minute of every day for the rest of your life. Because yes, those other things, they're great things.  
But these things are better."

Figure (d)

Whole Image segmented for further  
processing

Figure1. Pre-processing of the image before passing it to Tesseract OCR

# International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 8, Issue 5, May 2019

## VI. CONCLUSION

We have implemented a system that can be used by Visually Challenged Individuals for reading the text in an image captured by him. This is done by applying pre-processing steps to the image and passing it to Tesseract OCR. The text received from OCR Engine is then passed on to the Python Library which converts it into audio.

## REFERENCES

1. Carlos Sterling, 'Web Based Speech to Text and Text to Speech Application for Providing Online Academic Test for Blind Students', [http://icee.usm.edu/ICEE/Projects/P-1901-04-S\\_\(Carlos\\_Sterling\)/graduateProposal/Carlos\\_Sterling\\_Final\\_Proposal\\_9-20-04.htm](http://icee.usm.edu/ICEE/Projects/P-1901-04-S_(Carlos_Sterling)/graduateProposal/Carlos_Sterling_Final_Proposal_9-20-04.htm), 2004
2. Mrs. Shilpa Reddy K, Mounika S.K, Pooja K and Sahana N, 'Text to Speech for Visually Impaired', International Research Journal of Computer Science (IRJCS), Vol. 4, Issue 5, pp. 220-225, 2017
3. Hideaki Goto and Takuma Hoda, 'Real-Time Text Tracking for Text-to-Speech Translation Camera for the Blind', Springer International Publishing Switzerland, pp. 658-661, 2014
4. Hyung Il Koo and Duck Hoon Kim, 'Scene Text detection via Connected Component Clustering and Nontext Filtering', IEEE Transactions on Image Processing, Vol. 22, Issue 6, pp. 2296-2305, 2013
5. Faustin P'egeot and Hideaki Goto, 'Scene Text Detection and Tracking for a Camera-Equipped Wearable Reading Assistant for the Blind', Springer-Verlag Berlin Heidelberg 2013, pp. 454-463, 2013
6. Text Cleaning, Text Normalization: <https://machinelearningmastery.com/clean-text-machine-learning-python/>
7. World Health Organization, 'Blindness and vision impairment', <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>, 2018
8. Wikipedia, Speech Synthesis, [https://en.wikipedia.org/wiki/Speech\\_synthesis](https://en.wikipedia.org/wiki/Speech_synthesis)
9. Eye-Scan-Reader, <https://www.maxiaids.com/eye-scan-reader-plus>
10. Optical Character Recognition, [https://en.wikipedia.org/wiki/Optical\\_character\\_recognition](https://en.wikipedia.org/wiki/Optical_character_recognition)