OCR AND TEXT RECOGNITION FOR ASSISTING VISUALLY IMPAIRED PEOPLE USING ANDROID SMARTPHONE

Venkat Biyyapu[1]
School of CSE
REVA University
Bengaluru, India
venkatbiyyapu1708@gmail.com

Y Venkateswara Reddy[2]
School of CSE
REVA University
Bengaluru, India
yattapuvenkatesh122@gmail.com

T Sai Kumar Reddy[3]
School of CSE
REVA University
Bengaluru, India
tsaikumarreddy5@gmail.com

K. Thanuja Asst. Prof [4]
School of CSE
REVA University
Bengaluru, India
k.tanuja@reva.edu.in

***Abstract*— New technologies advancement of systems are providing assistance to visually impaired people. These systems aim to assist by supplying essential information about their surroundings to their users through the use of senses that they can still employ. In this project, we use Android Studios to create an app that uses existing technologies such as Text-to-Speech (TTS) and Optical Character Recognition (OCR) to identify and recognise texts from images captured by the Android phone's camera, assisting users in understanding what the image is about and also can be useful in navigation.**

Keywords— OCR, TTS, Android Studios.

I. INTRODUCTION

Braille has been used in most public areas for a long time, although it is mainly restricted to fixed signs and does not extend to temporary posting. If visually challenged persons need to read a text, they can scan it line by line with OCR technology and convert it to braille, or they may use TTS algorithms to read it. These gadgets, on the other hand, necessitate physical contact with the material being read, entails being aware that knowledge is there and having physical access to it. These two concerns are addressed by the suggested approach. The aim is to utilize a camera to automatically find information from several sources in the surroundings and tell the user of their position using TTS algorithms.

The system also reads the various sources using OCR and then uses Text-to-Speech algorithms to convey their information to the user. The system is built as a smartphone application in order to keep costs down.

Smartphones have various benefits over other devices, including the fact that they are not only inexpensive, but most models now include many cores in their primary CPU, making them quite powerful.

We begin by discussing the system design using an Android phone in this project. The text is then recognized from the picture, and the OCR engine is discussed. Finally, we go through the application's preliminary findings.

II. LITERATURE SURVEY

1. "OpenCV Based Implementation of Zhang-Suen Thinning Algorithm Using Java for Arabic Text Recognition" [2019]

   In this paper, they have used OpenCV library in order to perform text recognition, the algorithm they used in order to recognize is way complexed.

2. "Text recognition and face detection aid for visually impaired person using Raspberry PI" [2020]

This research presents a camera-based assistive text reading system to aid visually impaired individuals in reading text recorded on a captured image. Faces can also be identified by the mode control when a person enters the frame.

3. "Implementation of Text Recognition and Text Extraction on Formatted Bills using Deep Learning" [2020]

   In this paper, they employed the East method to convert letters and words from an image or scanned page into machine readable form, which is time consuming and also, they are just extracting the text.

4. "Machine Learning Tensor Flow Based Platform for Recognition of Hand Written Text" [2021]

   In this paper, recognition of text is based on artificial NN (Neural Networks), and they have determined the probability of getting the expected text.

5. "Offline Handwritten Quranic Text Recognition" [2020]

   The purpose of this work is to describe the offline handwritten Quranic text recognition system.

III. METHODOLOGY AND IMPLEMENTATION

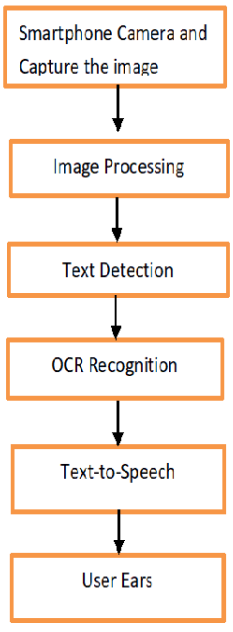

Fig1: Flow Chart of Proposed System.

In this project we have used Google Vision API which help us to achieve the OCR.

A. Optical Character Recognition (OCR)

OCR technology is a business solution for extracting data from written or printed text in an image file or scanned page and converting the text into a machine-readable format for data processing such as

searching and editing.

B. Text-To-Speech (TTS)

To produce a vocal version of the written document, the Text-to-Speech Engine technology (also known as TTS) is employed. TTS is gaining traction as the usage of digital devices grows, as is our reliance on speech recognition and other comparable technologies.

Above Fig1 flowchart, shows the design flow of app created. First of all, the app opens the camera and there would be surface area shown how much can the camera capture an image. Once the image comes under the surface area it gets captured.

The captured image will undergo image processing which would filter out the text. Then the detection of text happens with the google vision api which uses OCR software to perform the character recognition.

Once the character recognition is done the text is converted to speech by the Text-To-Speech Engine. The speech usually is converted to the locale language such as English or British.

After conversion the displayed on the app and also it read aloud so the visual impaired or any other user can hear it.

IV. RESULT

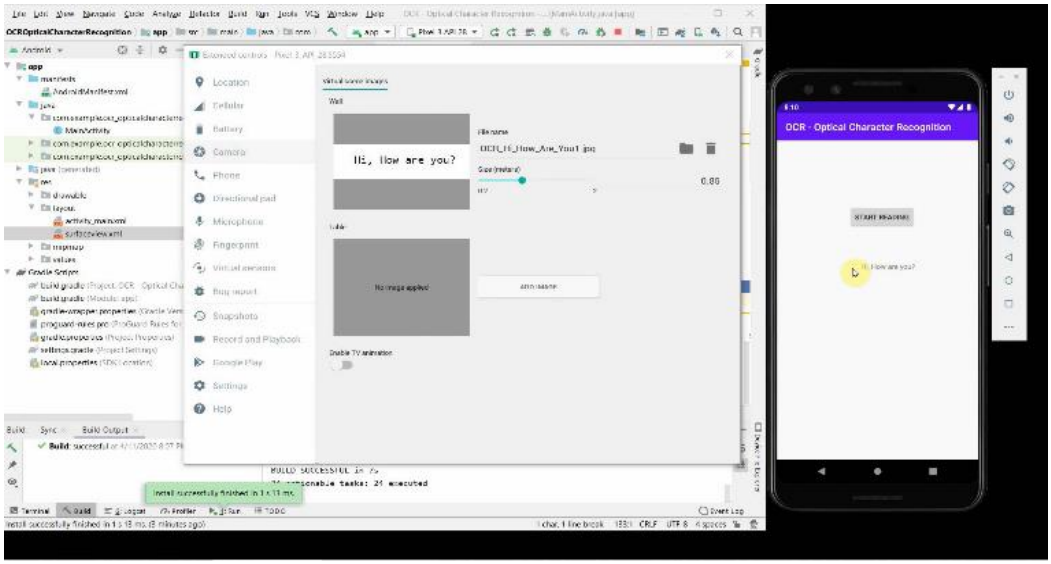

Fig 2: HI How Are You Message

The above fig shows the extraction of text form the printed image captured and it is read aloud correctly.

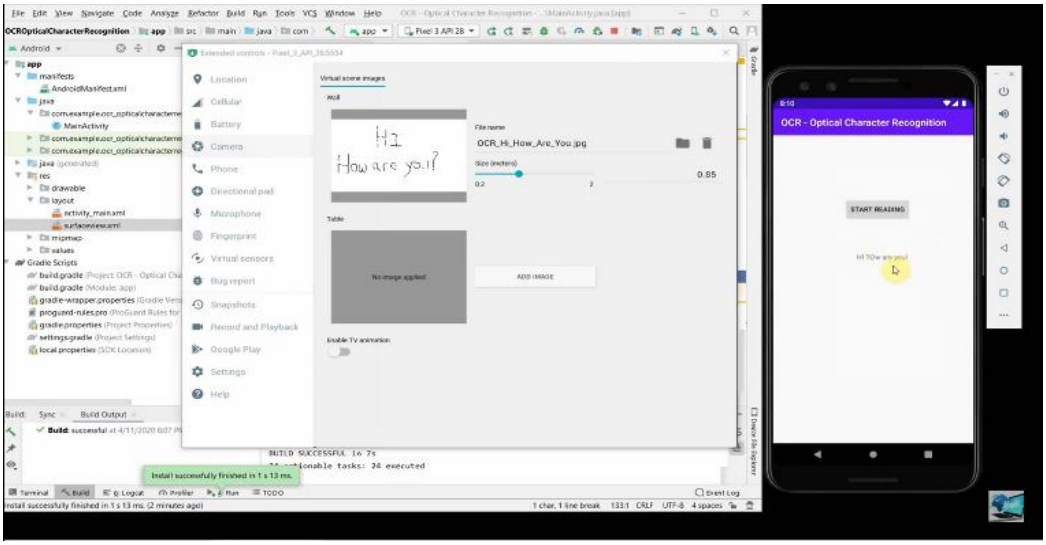

Fig 3: HI How Are You Message

The above fig shows the extraction of text from the handwritten image captured and it is read aloud.

We have installed an android emulator (Nokia 5G and Android above version 6.0), upon successful build the emulator starts running and after successful run it install the OCR App that we designed. Once OCR App is up and running, we have inputted printed image that. Once, the image comes under surface view, the text gets extracted. Extraction of the text is explained in the Methodology. After the extraction it is converted to speech and read aloud. Similarly, in case of handwritten the text is extracted and read aloud.

Figures from 2 to 3 shows the text detection and how it is read aloud. In case of printed image, the app could read it out loud accurately. But in case of handwritten image the app couldn't identify the text accurately. Instead "HI How are you" it detected "HI Tow are
you".

After trying with many images, we have found that the app developed was pretty useful in case of printed images
that had more clarity in extraction but it fails in case of handwritten images, as they don't have enough clarity of characters.

V. CONCLUSION

The design method for implementing a TTS algorithm that can help visually challenged persons to "read" the text in photos or documents was provided in our study. The trials indicate that the idea can be implemented on an Android smartphone using still photos, and that it can eventually be extended to a real-time version. But the drawback of the App developed is it fails to recognize the text and convert it to speech accurately in case of handwritten images. So, there is an enhancement going on to improve text recognition, by using Artificial Intelligence (AI) and Natural Language Programming (NLP).

REFERENCES

[1] D'souza, Lyzandra, and Maruska Mascarenhas. "Offline handwritten mathematical expression recognition using Convolutional Neural Network." 2019 International Conference on Information, Communication, Engineering and Technology (ICICET). IEEE, 2019.

[2] World Health Organization, "Visual impairment and blindness - Fact Sheet N°282," August 2019. Available: http://www.who.int/mediacentre/factsheets/fs282/en/. [Accessed January 2019].

[3] B. N. K. Sai and T. Sasikala,"Object Detection and Count of Objects in Image using Tensor Flow Object Detection API," 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, 2019

[4] Obaid, Ahmed & El-Bakry, Hazem & Eldosuky, M.A. & Shehab, Abdulaziz. (2020). Handwritten Text Recognition System based on Neural Network. International Journal of Advanced Research in Computer Science & Technology. 4. 72-77

[5] Noor, Rouhan, Kazi Mejbaul Islam, and Md Jakaria Rahimi. "Handwritten Bangla Numeral Recognition Using Ensembling of Convolutional Neural Network." 2019 21st International Conference of Computer and Information Technology (ICCIT). IEEE, 2019

[6] Malakar, Samir, et al. "Text line extraction from handwritten document pages using spiral run length smearing algorithm," IEEE International Conference on Communications, Devices and Intelligent Systems (CODIS), 2020.

[7] R. Smith, "An Overview of the Tesseract OCR Engine," Ninth International Conference on Document Analysis andRecognition (ICDAR 2020), vol. 2, pp. 629-633, 2020.

[8] X. Chen and A. L. Yuille, Detecting and reading text in natural scenes, in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2019, pp. 366373.

[9] K. C. Jung, K. I. Kim, and A. K. Jain,Text information extraction in images and video: A survey, Pattern Recognit., vol. 5, pp. 977997, May 2020.

[10] A.Nichols,"Why Use The Long White Cane?," 2020. Available: https://web.archive.org/web/20100330050804/http://www.blind.net /g42w0001.htm.

[11] M. W. A. Kesiman, S. Prum, J. C. Burie and J. M. Ogier, "Study on Feature Extraction Methods for Character Recognition of Balinese Script on Palm Leaf Manuscript Images, Proceedings of 23rd International Conference on Pattern Recognition," Cancun, Mexico, pp. 4006-4011, Dec 2021.

[12] Basha, S. M., Poluru, R. K., & Ahmed, S. T. (2022, April). A Comprehensive Study on Learning Strategies of Optimization Algorithms and its Applications. In 2022 8th International Conference on Smart Structures and Systems (ICSSS) (pp. 1-4). IEEE.

[13] American Foundation for the Blind, "Technology Resources for People with Vision Loss,". Available: http://www.afb.org/info/living-with-vision-loss/usingtechnology/12. [Accessed January 2019].

[14] Y. M. Y. Hasan and L. J. Karam, Morphological text extraction from images, IEEE Trans. Image Process., vol. 9, no. 11, pp. 19781983, Nov. 2019

[15] T. Kasar, J. Kumar, and A. G. Ramakrishnan, Font and background color independent text binarization, in Proc. 2nd Int. Workshop Camera-Based Document Anal. Recognit., 2007, pp. 3 9.