Tesseract Aid for Blind People Using Optical Character Recognition (OCR) Algorithm

M. Lilly Florence¹, Arishiya Thabasum², Ashwini A³, Kamesh P⁴, Mahmood Hameed Majeed⁵

¹Professor, Department of Computer Science and Engineering Adhiyamaan College of Engineering, Tamil Nadu, India. drlilly2011@gmail.com
²Department of Computer Science and Engineering Adhiyamaan College of Engineering, Tamil Nadu, India. emailofarishiya@gmail.com
³Department of Computer Science and Engineering Adhiyamaan College of Engineering, Tamil Nadu, India. ashwiniannaduarai@gmail.com
⁴Department of Computer Science and Engineering Adhiyamaan College of Engineering, Tamil Nadu, India. kamesh1711476@gmail.com
⁵Al-Bayan University, College of Health and Medical Techniques, Iraq. Mahmood.h@albayan.edu.iq

Received 04th Feb 2022, Accepted 08th Mar 2022, Online 7th May 2022

Abstract: Great vision is a valuable gift, yet tragically, vision misfortune is becoming normal these days. Outwardly impeded individuals report countless hardships in their everyday life. One of the primary and most significant challenges is understanding texts. To assist with blinding individuals, the visual world must be changed into the sound world with the possibility to illuminate them about texts they are running over. In this venture, we will quite often give them a gadget that could help them in their consistently exercises by assisting them with further developing perusing and advancing by changing over visual texts into sound signs. This gadget catches the picture when pointed by the client and finds the message present in the picture. The text is then separated from the picture and changed over into sound to give the client an explained result. This task assists us with distinguishing different hardships in identifying and perceiving text progressively by a normal outwardly hindered individual and foster answers for help them. In our methodology, we have utilized OCR (Optical Character Recognition) for text-level forecasts, and afterward we get the boxed math result of the relative multitude of texts in the pictures. Then, at that point, for acknowledgment of the text, we give it to tesseract OCR to get the separated text, and afterward we convert the text to discourse for the result. The primary inspiration driving our undertaking is to assist outwardly debilitated individuals with bettering perceive every one of the texts before them and assist them with carrying on with their everyday life very much like some other ordinary individual.

Keywords: Tesseract Aid; Blind People; Optical Character Recognition (OCR) Algorithm

I. INTRODUCTION

Perusing difficulties can influence outwardly weakened individuals, those with acquiring handicaps or low education abilities, and the people who experience issues holding books or documents [1]-[5]. These people might profit from the utilization of different understanding innovations and procedures.
One arrangement is to utilize a perusing framework that utilizes optical person acknowledgment innovation to make an electronic duplicate of a record and afterward use text to discourse innovation to peruse the data to the client [6]. There are three fundamental components for text understanding innovation. They are checking, perceiving and understanding text. At first, a printed record is filtered by a Camera. OCR then, at that point, changes over pictures into perceived characters. And afterward stands up the perceived text [7]. Optical Character Recognition (OCR) changes over the pixel Representation of a letter got by filtering a message or a report into its identical person portrayal [8-11]. OCR has been being developed for right around 80 Years. A cutting edge form of OCR showed up in the center of the 1940s, with the advancement of computerized PCs. From that point forward, a few person acknowledgment frameworks Has been proposed. The primary patent for an OCR machine was documented by a German named Gustav Tauscher in 1929 [12]. There are assortments of programming based arrangements Available for OCR. Nonetheless, there is little work done in the equipment execution of OCR [13].

A text-to-discourse framework changes over ordinary language text into discourse. Text to discourse change frameworks has a tremendous scope of uses [14]-[17]. Their first genuine use was in perusing frameworks for the visually impaired, where a framework would peruse a few text from a book and convert it into discourse. Programmed text acknowledgment from pictures has as of late gotten a developing interest in light of its utilization in picture recovery, video ordering, portable mechanical technology, and canny vehicle frameworks, to give some examples key applications [18]-[21]. Specifically, strategies have been created to peruse the text experienced in regular scenes, help outwardly weakened people, or distinguish the text inside street signs from normal scene video [22]. Notwithstanding, the creators by and large concede that current outcomes are deficient for functional use. Different methodologies of interest incorporate a camera-prepared PDA based wayfinding framework and a wearable camera framework that consequently finds and track text areas in encompassing scenes [23].

II. OBJECTIVE

To help blind individuals in perusing texts which will be useful for them to perceive texts progressively circumstances like knowing the texts in transport vehicles (transports) [24]-[29]. The principle inspiration driving our undertaking is to assist outwardly weakened individuals with bettering perceive every one of the texts before them and assist them with carrying on with their everyday life very much like some other typical individual. To assist outwardly hindered individuals with including in their examinations by perusing texts of their own without somebody’s assistance. To foster a text perusing help utilizing the video handling method [30].

III. EXISTING SYSTEM

Printed archives can be immediately changed over into computerized text documents through optical person acknowledgment and altered by the client [31]. Thusly, negligible time is expected to digitize reports, especially accommodating while filing volumes of written words. This study exhibits how picture handling advancements can be joined with optical person acknowledgment to further develop acknowledgment exactness and the productivity of extricating text from pictures [32].
IV. PROPOSED SYSTEM

This caught picture will go about as contribution to our OCR module, checking the picture and perceiving its text. The perceived text from the OCR module will be utilized as contribution for our text-to-discourse module, giving clients discourse yield [33]. OCR is a system that converts pictures of composed or printed text into machine-encoded text [34]-[39]. This framework will catch live web based video and concentrate pictures as edges by changing over them from 24fps to 1fps. Then, at that point, the extricated outline/picture will be handled in OCR to separate text from those pictures [40]. The reason for conveying the result as voice/discourse is to serve the data present on the live video to the outwardly hindered so they can know about the messages around them without the assistance of another person (figure 1) [41].

V. DIAGRAM

![Diagram of Optical Character Recognition](image1)

VI. MODULES

There are 5 modules used in this study

- Input camera module
- Conversion module
- Noise removal module
- Text recognition module
- Audio module

**INPUT CAMERA MODULE**

This module will capture the live streaming video using a camera, and the input video is considered an input source for text recognition [42]-[55].

**Conversion Module**

This module will capture the live streaming video using a camera, and the input video is considered an input source for text recognition [42]-[55].
In this module, the captured live streaming video extracts images from the video as frames by converting it from 24fps to 1fps [56].

**Noise Removal Module**

The converting video from the input image can be extracted, and the images can be segmented; after that, the background noise can be removed in this module [57].

**Text Recognition Module**

The extracted images can be processed using the OCR object recognition algorithm in the text recognition module, and the text character and numbers can be recognized. After the text is recognized, the text is converted into voice [58].

**Voice Alert Module**

In this module, the text can be converted as audio. The voice alert is used for blind easily understanding. Optical Character Recognition. Optical Character Recognition (OCR) fills in as an apparatus to distinguish data from regular pictures and move them into machine-coded texts, for example, words, images and numbers [59]-[65]. It is as yet a hot continuous hunt region, and a few novel calculations are distributed now and again [98-124]. It is really fascinating and fundamental to perceive the characters in the picture since it could help incredibly in a specific region [125-149]: auto plate number acknowledgment, books and records checking, assistive innovation for blind and outwardly hindered clients, postal district acknowledgment required for mail depots and considerably more [66].

**Feature Extraction**

In the principal strategy, the calculation for highlight discovery characterizes a person by assessing its lines and strokes [67]-[75]. In the subsequent strategy, design acknowledgment works by distinguishing the whole person [76]-[79]. We can perceive a line of text via looking for white pixel pushes that have dark pixels in the middle [80]. Also, we can perceive where a person starts and wraps up [150-175]. The following pictures show the visualization of these methods respectively (figures 2 to 4)

\[
\text{i} + \text{i} = \text{V}
\]

Fig. 2. Feature detection
Next, we convert the image of the character into a binary matrix where white pixels are 0s, and black pixels are 1s, as shown in the following image figure 5:

![Sample of binary matrix](image_url)

**Fig. 5. Sample of binary matrix**

Then, by using the distance formula, we can find the distance from the centre of the matrix to the farthest 1.

The distance formula:

\[ d = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} \]
We then, at that point, make a circle of that range and split it up into more granular areas.
At this stage, the calculation looks at every subsection to a data set of frameworks addressing characters
with various text styles to distinguish the person it shares most practically speaking measurably [81]-[97].
It makes it simple to carry printed media into the computerized world for each line and character (figure 6).

![Fig. 6. Compare each subsection against the matrix database](image)

**Tesseract**

Tesseract is an optical person acknowledgment (OCR) framework. It is an open-source programming go
through a Command-Line Interface (CLI) [176-189]. Tesseract is viewed as perhaps the most reliable
open-source OCR motor presently accessible, and Google has supported its turn of events. Its abilities can
be more restricted than business programming like Adobe Acrobat Pro and ABBYY Fine Reader [190-
195]. Nonetheless, on the grounds that it is open-source programming, anybody with programming
information can alter the code behind tesseract and assist it with realizing what you really want to do. It
tends to be utilized on Mac, Windows, and Linux machines. Essential OCR Operations in Tesseract:
Image design (JPG, TIF, PNG, and so forth) to PDF, Microsoft Word. The new record shows up in a
similar registry as the underlying report. Go through your Command-Line Interface.

**VII. CONCLUSION**

The voice helped text perusing framework for the outwardly hindered is examined. The result is displayed
for the different info informational collection like just text inputs, text with pictures combined and so
forth Optical Character Recognition predicts the information text with a pre-stacked data set layout. On
the off chance that the characters are thought about, discourse yield is created utilizing text to a discourse
synthesizer. The work is reproduced utilizing iota programming, creating the discourse yield. With the
assistance of the proposed module, the client feels more straightforward to peruse the text as discourse
utilizing Optical Character Recognition and discourse incorporating. The proposed work is tried with
various info sets in printed text design where clamor parts are taken out and text is extricated to foresee
the text aurally. It is additionally upgraded by recognizing the neighboring person acknowledgment for
successful perusing to stay away from irregularity. A result is delivered as a sound result to peruse the
relating input, which assists blind individuals with perusing any printed text in vocal structure.

**Conflicts of Interest:** The authors declare that they have no conflicts of interest to report regarding the
present study.
REFERENCES


26. V. Kumar, S. Arora, S. Kumar, T. W. Kang and H. C. Jeon, “Annealing led conversion from polypyrrole to carbon nitride nanowires and the fabrication of highly efficient ammonia sensing


42. H. Chhabra, V. Mohan, A. Rani, and V. Singh, "Multi objective PSO tuned fractional order PID


55. Shalini Stalin, Vandana Roy, Prashant Kumar Shukla, Atef Zaguia, Mohammad Monirujjaman Khan, Piyush Kumar Shukla, Anurag Jain, "A Machine Learning-Based Big EEG Data Artifact Detection


61. Janarthanan Ramadoss, Maheshwari Uma, Shukla Prashant Kumar, Shukla Piyush Kumar, Mirjalili Seyedali, Kumar Manoj (2021). Intelligent Detection of the PV Faults Based on Artificial Neural Network and Type 2 Fuzzy Systems. Energies 2021, 14(20), 6584,


2021.


105. Maninder Singh, Hardeep Singh Saini and Dinesh Arora, “Bit error rate minimization in OFDM-MIMO system”,2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), held on 5-7 March 2015,Coimbatore, Tamil Nadhu-India. IEEE.


155. Subramanian Vasantharaj, Indragandhi Vairavasundaram, Subramaniyaswamy Vairavasundaram, Yuvaraja Teekaraman, Ramya Kuppusamy, Nikolovski Srete, Efficient Control of DC Microgrid with Hybrid PV—Fuel Cell and Energy Storage Systems” Energies 2021, 14(06), 3234 (1-18);


160. Hariprasath Manoharan, Yuvaraja Teekaraman, Pravin R Kshirsagar, Shanmugam Sundaramurthy, © 2022, CAJOTAS, Central Asian Studies, All Rights Reserved


