Text Recognition and Medicine Identification by Visually Impaired People

Snigdha Kesh
MTech Scholar, Dept. of CSE,
AMC College of Engineering, Bangalore, India,

Ananthanagu U
Assistant Professor, Dept. of CSE,
AMC College of Engineering, Bangalore, India

Abstract- **Now-a-day digital technology and image capturing devices are growing rapidly. This advanced technology reached to some corner of the society like visually impaired people and older people who are struggling with their daily activities like reading, identifying known and unknown people and specially identifying daily consumable medicine. Due to perspective distortion, aspect ratio, font size, unique style, detecting text from scene image is complex job compared to the printed document. There should be some hardware device which can help the visually impaired people to detect text and identify the medicines. But such devices are costlier and need to carry such devices would make things more complex. The proposed system will help visually impaired people to detect text by converting it to voice message and identifying medicines using in-built camera and application of the smart phone.**

Keywords: Android Application, Medicine Identification, Open CV, Smart phone, Text Recognition.

I. INTRODUCTION

Visually impaired means vision less whether it is totally or partially blindness. Blindness cannot completely solve by eyeglass or contact lenses. Loss of vision is another way of loss of independence.

Visually impaired people are suffering their daily life activities such as walking, reading document, identifying object, recognizing known or unknown person, recognizing medicine, doing shopping independently, even when they go for restaurant, need third person help to read restaurant menu card. According to WHO (World Health Organization) report [1] 285 million are visually impaired, among them 39 million blind, 246 million having low vision. Now-a-days Digital technology, Artificial Intelligent, Computer Vision and economical image capturing devices are become very popular and powerful. Using this advance technology and device researchers are making economical and potable system to assist Visually impaired people.

Reading is essential part in our society. Reading text provide huge knowledge about our enlivenment and help to navigation. Text can be anything like report, receipt, product package, even restaurant menu card etc. Reading text document is one of the difficult job for Visually impaired people. Sometimes document need to maintain privacy like bank document, Tax invoice, letter etc., which people not like to give other to read. But because of loss of vision people need to take other help to read those personal documents.

The need to read textual and symbolic information become essential in case of blind or Visually challenged persons. With this point of view, the system need to detect text from textual and symbolic information and recognize the text character from the captured scene text image and last text or symbolic information convert to the voice message. Some algorithms are used to extract text information from image and text detection. Algorithm follow some key thing: Cluttered backgrounds with noise and text outliers and diverse text patterns such as character types, front and sizes. Convert scenic image to text structure is difficult job due to lack of discriminative pixel level, non-text background outliner, and different front style and size. So, scene text extraction process having two part such as text detection and text recognition.

Some other daily life activities, identifying medicine is one of the difficult job for Visually impaired people. They need every time somebody's help to identify medicine. Taking wrong medicine is dangerous for health. Image capturing base economical device like smart mobile help Visually challenged people to identifying the medicine, which is already stored by pharmacist. It also gives voice alert to blind person.

II. RELATED WORK

Nobuo Ezaki, Marius Bulacu , Lambert Schomaker [2] exhibited a framework for outwardly tested individuals ,can read content from common scenes and Novel (little text).Kumar J.A.V ,Visu A,Raj M.S actualized a computerized content to sound changing over pen . On the off chance that a man might want to peruse/see any bit of content that content is changed over to a sound flag. This sound flag is transmitted to the individual's ears through remote innovation, for example, ZigBee [3]. Oi-Mean Foong, Nurul Safwanah Bt Mohd Razali displayed a signage acknowledgment structure for Malaysian Visually Impaired People. Their proposed structure catches a picture of an open signage and changes it into a content record utilizing Otsu's OCR strategy. The content document peruses by a discourse synthesizer that tells the outwardly hindered individuals what the picture is. This structure does not require enormous database of the signage but rather just the character database [4]. In 1974, Ray Kurzweil began the organization Kurzweil Computer Products, Inc. what's more, proceeded with improvement of omni-textual style OCR [5], which could perceive content imprinted in basically any text style (Kurzweil is regularly credited with developing omni-textual style OCR, yet it was being used

by organizations, including CompuScan, in the late 1970s). Kurzweil chose that the best utilization of this innovation is make a perusing machine for the visually impaired, which would enable visually impaired individuals to have a PC read content to them so anyone can hear. This gadget required the creation of two empowering advances – the CCD flatbed scanner and the content to-discourse synthesizer. On January 13, 1976, the effective completed item was revealed amid a generally announced news meeting headed by Kurzweil and the pioneers of the National Federation of the Blind. In 1978, Kurzweil Computer Products started offering a business variant of the optical character acknowledgment PC program. LexisNexis was one of the primary clients, and purchased the program to transfer lawful paper and news archives onto its early online databases. After two years, Kurzweil sold his organization to Xerox, which had an enthusiasm for further commercializing paper-to-PC content transformation. Xerox inevitably spun it off as Scan-delicate, which converged with Nuance Communications The examination assemble headed by A. G. Ramakrishnan at the Medical insight and dialect designing lab, Indian Institute of Science, has created Print To Braille instrument, an open source GUI frontend that can be utilized by any OCR to change over examined pictures of printed books to Braille books.

There is some electronic gadget that helps outwardly tested individuals to recognize the solution. AccessaMed [6] is a gadget that can be connected to solutions, regardless of whether they're in a jug or a container. It is recorded by the drug specialist through discourse amalgamation and an uncommon program that enables him or her to print the data – precisely as it shows up on the remedy name – to the gadget. Digit-Eyes[6] is an application for iPhone that permits to peruse numerous things, including physician recommended drug marks, if drug specialist will print a "QR" code that is lucid by Digit-Eyes. A QR code is a code fundamentally the same as a scanner tag, just it's not molded like a standardized tag and it implants electronic data. The application permits to utilize telephone's camera to peruse the QR code and after that the data is talked so anyone might hear by the iPhone. The QR code data can be the data that is on the remedy name. It can likewise be extra data, for example, notices and precautionary measures to be brought with specific prescriptions.

All current framework helping outwardly debilitated individuals yet there are a few impediments. The majority of the framework required equipment that need to convey. Complex equipment builds costs of the gadget.

III. PROPOSED METHODOLOGY

Computer vision is one of the emerging technology that can be used to aid visually impaired people for navigation (both indoor and outdoor), accessing printed material, etc. This proposed system describes an approach to extract and recognize text from scene images effectively using computer vision technology and to convert recognized text into speech so that it can be incorporated with hardware to develop Electronic Travel aid for visually impaired people. This proposed system detects and recognizes text from surrounding and converting into speech to guide blind people in day-to-day activities. It supports blind people in terms of mobility, pathway finding and printed information access. It also helps to verifying the medicine details. This all implemented through smart phone and android application. Blind person need to shake the phone, the application start on automatically. This system capture the sense image and extract the text from that. System will convert the printed text information to voice message. Visually impaired people can aware of the information of the text document from voice command. Blind person can also get help from this system to identified their medicine. All the medicines are required by blind person are stored by pharmacist along with the doses and time in the server. Application will give alert of the medicine to blind person. System will capture the image of the medicine and identify the name of the medicine. After that system will provide voice command the name of the medicine.
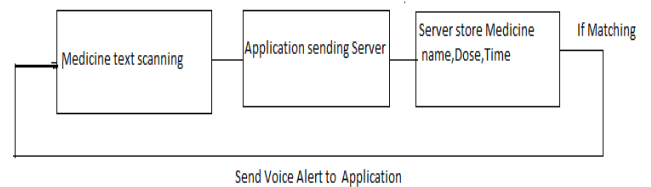

Figure 1: Text Recognition

Figure 2: Medicine Recognition

IV. IMPLEMENTATION

Execution is the stage it will begin after plan stage is finished. This proposed framework utilize two modules one is Text Recognition and Medicine distinguishing proof.

Content Recognition utilize some calculation and utilize two strategy - Pre - preparing and Post - handling.

1)   Pre-handling

OCR programming frequently "pre-forms" pictures to enhance the odds of effective acknowledgment. Strategies include:

De-skew – If the report was not adjusted appropriately when checked, it might should be tilted a couple of degrees clockwise or counterclockwise to make lines of content superbly level or vertical.

a)   Despeckle [7] – evacuate positive and negative spots, smoothing edges

b)   Binarization – Convert a picture from shading or greyscale to highly contrasting (called a "paired picture" in light of the fact that there are two hues). The undertaking of binarization [8] is executed as a straightforward method for isolating the content (or some other sought picture segment) from the foundation. The errand of binarization itself is vital since most business acknowledgment calculations work just on twofold pictures since it ends up being less complex to do as such. Likewise, the adequacy of the binarization step impacts to a critical degree the nature of the character acknowledgment arrange and the cautious choices are settled on in the decision of the binarization utilized for a given information picture sort; since the nature of the binarization strategy utilized to get the parallel outcome relies on upon the kind of the info picture (checked report, scene content picture, recorded debased archive and so on.).

c)   Line evacuation [9] – Cleans up non-glyph boxes and lines

d)   Layout examination or "zoning" – Identifies sections, passages, inscriptions, and so on as unmistakable pieces. Particularly essential in multi-segment formats and tables.

e)   Line and word recognition – Establishes standard for word and character shapes, isolates words if essential.

f)   Script acknowledgment – In multilingual records, the script may change at the level of the words and henceforth, distinguishing proof of the script is fundamental, before the privilege OCR can be conjured to deal with the particular script.

g)   Character segregation or "division" – For per-character OCR, different characters that are associated because of picture ancient rarities must be isolated; single characters that are broken into various pieces because of antiques must be associated.

Standardize perspective proportion and scale-Segmentation of settled pitch textual styles is proficient moderately basically by adjusting the picture to a uniform matrix in view of where vertical network lines will minimum regularly converge dark regions. For relative text styles, more refined methods are required on the grounds that whitespace between letters can here and there be more noteworthy than that amongst words, and vertical lines can meet more than one character.

h)   Character acknowledgment There are two essential sorts of center OCR calculation, which may deliver a positioned rundown of competitor characters.

Network coordinating includes contrasting a picture with a put away glyph on a pixel-by-pixel premise; it is otherwise called "design coordinating", "design acknowledgment", or "picture connection". This depends on the info glyph being effectively secluded from whatever is left of the picture, and on the put away glyph being in a comparative text style and at a similar scale. This system works best with typewritten message and does not function admirably when new text styles are experienced. This is the strategy the early physical photocell-based OCR executed, rather specifically.

Highlight extraction deteriorates glyphs into "elements" like lines, shut circles, line bearing, and line crossing points. These are contrasted and a unique vector-like portrayal of a character, which may diminish to at least one glyph models. General procedures of highlight identification in PC vision are material to this sort of OCR, which is normally observed in "canny" penmanship acknowledgment and to be sure most current OCR programming. Closest neighbor classifiers, for example, the k-closest neighbor's calculation are utilized to contrast picture includes and put away glyph highlights and pick the closest match.

Programming, for example, Cuneiform and Tesseract utilize a two-pass way to deal with character acknowledgment. The second pass is known as "versatile acknowledgment" and utilizations the letter shapes perceived with high certainty on the principal go to perceive better the rest of the letters on the second pass. This is worthwhile for bizarre textual styles or low-quality outputs where the text style is twisted (e.g. obscured or blurred). The OCR result can be put away in the institutionalized ALTO configuration, a devoted XML blueprint kept up by the United States Library of Congress.

2)   Post-handling

OCR precision can be expanded if the yield is compelled by a vocabulary – a rundown of words that can happen in an archive. This may be, for instance, every one of the words in the English dialect, or a more specialized vocabulary for a particular field. This strategy can be dangerous if the report contains words not in the dictionary, as formal people, places or things. Tesseract utilizes its word reference to impact the character division venture, for enhanced precision.

The yield stream might be a plain content stream or document of characters, however more complex OCR frameworks can protect the first format of the page and create, for instance, an explained PDF that incorporates both the first picture of the page and a searchable printed portrayal.

Special Issue - 2017

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
ICPCN - 2017 Conference Proceedings

"Close neighbor investigation" can make utilization of co-event frequencies to right mistakes, by noticing that specific words are frequently observed together. For instance, "Washington, D.C." is for the most part significantly more typical in English than "Washington DOC" [10].

Learning of the sentence structure of the dialect being examined can likewise help decide whether a word is probably going to be a verb or a thing, for instance, permitting more noteworthy precision.

The Levenshtein Distance calculation has additionally been utilized as a part of OCR post-handling to additionally advance outcomes from an OCR API.

In Medicine Identification Android studio is used to develop mobile application and data can be store in the server. Doctor will give only medicine prescription to the visually impaired people. Pharmacist will store the name medicine details, time and doses in the server from prescription. The details will store in the TOM CAT server and server will store the data into the MYSQL database. Blind person at home shake the phone and mobile application will start automatically and camera of the phone start scanning, when the mobile is placed above the medicine, scanning will start and text detector detect the me medicine name and send it to the server. Server check whether the medicines are matching or not and check user should take at that time. If the medicine need to take at that time, server send the notification to the user, the mobile application once receives the information from the then convert the text information to speech and plays the audio.

V. RESULT ANALYSIS

This project will start when Visually Impaired people shake mobile the application will start working. Mobile inbuild camera start capturing image and application will extract text from scene image then application will again convert text data to voice message. To identify medicine Visually impaired people shake phone, camera will on and take medicine image convert to text and send to server, server will match with store data., give voice alert to the blind person. Using this application, they will never miss medicine and no need third person help.

Figure 3: Shaking Phone

VI. CONCLUSION

In this proposed system used to detect the text from any document, it can recognize any signboard, recognize medicine, and can detect phone has stolen or not recognize known and unknown person, can read restaurant menu to help visually impaired people. These people can solve they're some of daily life problem without taking any third person help. After combining different techniques for Text detection and extraction, it is found that system works faster and better than using single technique for overall system. Text detection followed by recognition using supervised pattern recognition algorithm not only improves accuracy but also increases speed of the system. After successful recognition, text is converted into audio output. This project will encourage Visually impaired people to long live .No need take third person help to perform their daily activities and can maintain privacy in their life.

The areas of improvement are:
1. Smart Phone can be replaced by any other device if available.
2. Need to use better algorithm so low light image also can recognize.
3.Better algorithm for text recognition so any front style can accept.

ACKNOWLEDGEMENT

REFERENCE

[1] www.who.in
[2] Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, "Improved text-detection methods for a camera-based text reading system for blind persons", IEEE in Proceedings of Eighth International Conference on Document Analysis and Recognition, pp 257 - 261 Vol. 1 ISSN: 1520-5263, 2005.
[3] C. Yi and Y. Tian, "Text string detection from natural scenes by structure-based partition and grouping," IEEE Trans. Image Process., vol. 20, no. 9, pp. 2594–2605, Sep. 2011.
[4] P. Blenkhorn, D.G. Evans "Using speech and touch to enable blind people to access schematic diagrams" science direct, Journal of Network and Computer Applications,1998.
[5] www.edubilla.com
[6] www.visionaware.org -.
[7] https://books.google.co.in 4 % End for p end % End for l xmit = (1.0/9.0) *xmit; var=0.0; pk=0.0; pk4=0.0; % Detect the position of the mask for which the C of the gray levels is minimum
[8] $f$elixniklas.com/imageprocessing/binarization An attempt of a simple introduction to Image Processing
[9] www.leptonica.com/line-removal
[10] self.gutenberg.org/articles/eng/Character_recognition "Washington, D.C." is generally far more common in English than "Washington DOC".
[11] Diego López-de-Ipiña, Tania Lorido, and Unai López, "Indoor Navigation and Product Recognition for Blind People Assisted Shopping ", In proceedings of IWAAL 2011, LNCS 6693, pp. 33–40, 2011 .
[12] https://sandysview1.wordpress.com/2015/04/16/how-do-people-who-are-blind-or-visually-impaired-shop-independently/