International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 04 | April - 2021 ISSN: 2582-3930

AI-OCR IN HEALTHCARE

Abarna M*, Srinivasini*, Gayatri Rupaa B*, Ashok Kumar S**

**UG Scholar, Department of Information Technology, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India*

***Assistant Professor, Department of Information Technology, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India*

Abstract - The different methods to report digitization are discussed in this article. The bulk of written works released do not contain Braille or audio copies, and digital versions are also in the minority. Optical character recognition technology allows for the recognition of texts from image files, as well as their translation and recording in audio format. Speech synthesis is a technique that converts a digital text into a human voice and plays it via an audio device. There is a massive market these days for storing information in paper record format on a data storage disc and then reusing the information through a search process. Blind individuals are more likely to read their audio recording, which will aid them with interpreting the right report regarding themselves.

***Key Words*:** Optical Character Recognition, Text to Speech, Audio system, Storage disk

1. INTRODUCTION

Since blind individuals are unable to execute visual activities, this device may assist them in understanding the content of a study or text. Digital formats are not used with the bulk of written works, and braille and audio versions are also in the minority. Optical Character Recognition (OCR) is a technology that helps you to read text from scanned image images. It's used to identify characters in documents as well as to distinguish scanned files. To recognize the character in the scanned image of the text or book, Optical Character Recognition is used.

The speech synthesis (TTS) technology converts a digital text into a human voice and plays it via an audio device. TTS aims to turn sentences into spoken discourse in a very natural language instantly and without restrictions. This technology is very useful in combining the parts of a character and forming certain characters into an expression that sounds very much like a real human voice. Speech synthesis research has progressed, resulting in a rise in the importance of many new applications. Speech Synthesis is a tool for transforming scanned image data into human speech using audio. It is extremely beneficial for blind people to be aware of the material or substance included in a text or article.

This paper demonstrates a revolutionary concept and low-cost method for hearing the contents of a text picture without having to translate them. As a result, it incorporates the concepts of OCR and Text to Speech synthesizer (TTS). This device uses a voice interface to aid blind people in successfully communicating with machines.

Computers have a difficult time removing text from colored or non-colored images. The text-to-speech device reads the English alphabets and numbers in an image using optical character recognition (OCR) and converts the text into a voice format. The planning, deployment, and experimental results of the system are presented in this article.

The computer is made up of two parts: an image processing module and a voice processing module. The method of converting scanned or typed text images into text format for further processing is known as optical character recognition (OCR). This paper discusses a basic method for extracting text and converting it to words. Device testing is carried out using the Visual Studio module. The text-to-speech (TTS) device creates a more normal sound that closely resembles that of a human. Voice-enabled e-mail and texting are two examples of speech synthesis.

2. LITERATURE SURVEY

Character recognition is not a recent problem, but its origins can be traced back to programs before computers were invented. The first optical character recognition systems were mechanical machines that could identify characters but only at a very slow speed and with poor precision. M. Sheppard invented the GISMO reading and robot in 1951, which is considered the first work on modern Optical Character Recognition.

A novel adaptive binarization approach based on wavelet filter has been proposed by Yang et al. [2]. This method was developed at a high rate to make it more suitable for real-time processing and use on mobile devices. They tested this adaptive approach on ICDAR 2005 database complex scene pictures. On highly corrupted Indian language text photos, Sankaran et al. suggested a novel recognition method that resulted in a 16 per cent reduction in word error rate.

Extraction of text lines from paper images is one of the essential steps in the process of an Optical Character Recognition method, according to Malakar et al. [3]. The inclusion of bent, touching, or overlapping text line(s) in handwritten paper images or pictures makes this process a real problem for the researcher.

3. PROPOSED PLAN

To fix the limitations of the current method, we created a project for blind people focused on neural Optical Character Recognition in OpenCV. The suggested method is mostly used to help blind people interpret the text from a ch pattern and context to read a document or article. Our system's key goal is to classify the text in the texts. The object image is obtained using a camera first, and then image processing is performed. To run an automatic machine that checks a document and reads its contents to a user at the touch of a button. The vocal is conveyed with the aid of a speaker to assist the user in reading the text of the scan paper or article. Our system makes it possible for blind people to read without taking up a lot of space.

Our scheme assists blind people in reading text from difficult patterns and backgrounds without taking up a lot of space, and the new system is to assist blind people in reading text from challenging patterns and backgrounds for the sake of reading.

A. *Image capturing*

The initial stage in which a computer is passed over a written page and an inbuilt camera takes images of the document's text. The nature of the picture captured would be high, making for fast and simple identification.

B. *Pre-processing*

Skew Correction, Noise Reduction, and other pre-processing tasks were completed, and the captured image was tested for skewing. There are a variety of outcomes that can occur when an image is tilted to the left or right. The image has been brightened and binarized in this version. The skew recognition process primarily looks for a point of orientation about 15 degrees and, if one is found, a minor image rotation is performed before the lines align with the real horizontal axis, resulting in a skew corrected image. Before proceeding with the editing, any noise that could have occurred during the capture or as a result of the page's poor quality must be removed.

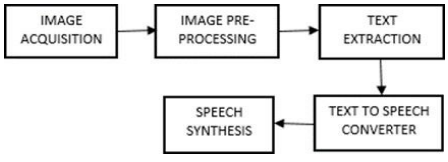

Fig. 1. Flow of process

C. *Text extraction*

Text extraction is a way of removing text from electronic records that have text associated with them. The interface is identical to using the Laser fiche Client's "Generate Searchable Text" order on an electronic record. Once the text has been extracted from a document or article, it can be submitted to the archive along with the document, rendering it full-text searchable in Laser fiche. Text Extraction can be combined with other processes in Fast Fields, such as Pattern Matching, to categorize documents and populate metadata. Optical Character Recognition engines are used to distinguish the interpreted material in the scanned image at this time. The tesseract Optical Character Recognition engine is used to distinguish the known characters in this scenario.

D. *Feature extraction*

The attribute extraction over each scanned image glyph is taken into account and characteristics are extracted. The width and height of the symbol, the number of horizontal and vertical lines present (short and long), the number of vertically and horizontally set arcs, the number of circles present, the centroid of the graphic, and the location of the different features and pixels in the various zones all contribute to the definition of a character glyph.

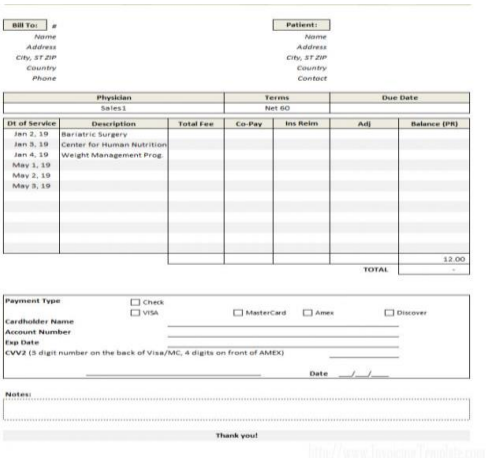

Fig. 2. Captured Image

E. *Text to Speech*

The Android operating system's Text-to-Speech functionality is similar to a screen reader. It is used to control applications that read aloud (speak) the text on the screen in a variety of languages. Apps like Google Playbook for reading books aloud, Google Translate for reading aloud translations that offer invaluable insight into word pronunciation, Google Talkback and other spoken input accessibility-based software, as well as third-party tools, use text-to-speech. Each language's voice data should be mounted.

F. Speech synthesis

The artificial processing of human speech is known as speech synthesis. A speech machine, also known as a speech synthesizer, is a computer device that performs this function. It can be implemented as software or hardware. A text-to-speech device translates written text into spoken words. Linking bits of captured speech contained in a database may be used to construct synthesized speech. Systems vary in the size of the stored speech units. The highest display spectrum is provided by the system that stores mobile, but clarification can be lacking. The storage of whole words or sentences allows for high-quality production in unique user domains. A synthesizer, on the other hand, may use a representation of the vocal tract as well as other human speech features to produce a fully "synthetic" voice output.

4. SOFTWARE SPECIFICATION

Raspbian, which is derived from the Debian operating system, was used to run this project. The algorithm is written in the Python scripting language. The functions in the algorithm are based on open-source Computer Vision library functions. Open-source Computer Vision is a free and open-source computer vision library written in C and C++ that runs on Linux, Windows and Mac OS X. Open-source Computer Vision was created with a strong focus on ongoing applications and computational efficiency in mind. Computer Vision is an open-source project written in enhanced C that can take advantage of multi-core processors. The open-source Computer Vision library provides over 500 functions that cover a wide range of vision topics, such as manufacturing plant object investigation and medical imaging, security, UI, camera adjustment, stereo vision, and robotics. Since computer vision and machine learning are often combined, Open-Source Computer Vision also has a comprehensive, universally accessible Machine Learning Library (MLL). We must implement Optical Character Recognition and Text Speech engines with predefined libraries to assist Optical Character Recognition and Text to Speech activities.

5. APPLICATION

Picture text recognition technology is now being used in healthcare to handle the paperwork. Healthcare providers are constantly dealing with vast numbers of forms for particular patients, including insurance and general health forms. To keep track of all of this data, it's a good idea to enter relevant data into an electronic archive that can be retrieved as needed. They will retrieve information from forms and enter it into databases using image recognition technologies, ensuring that any patient's data is automatically registered. As a result, healthcare providers will concentrate on offering the best treatment available to each particular patient.

There has also been a major trend in the legal industry to digitize paper papers. Documents are copied and entered into electronic libraries to conserve space and remove the need to sift through boxes of paper papers. Image text recognition simplifies the process even further by making records text-searchable, making them easy to navigate and deal with after they've been stored in the archive. Law professionals today have convenient and simple access to a large library of electronic records that can be found by simply entering in a few keywords.

6. CONCLUSION

The results of the procedure mentioned above are seen in the, which displays the picture taken with the camera, as well as the pre-processed scan image that is sent to the tesseract OCR engine for material removal. The accuracy can be increased by using a high-resolution camera to view the performance of the tesseract OCR engine. The performance is saved in a dropbox, so the report won't be lost and you won't have to cart it around with you all the time.

ACKNOWLEDGEMENT

Without the help of our Mentor, Ashok Kumar, this paper and the study that went into it would not have been feasible. The kindness and experience of every one of us have changed this analysis in many ways and rescued us from making many mistakes.

REFERENCES

[1] Velumurugan D. et al, “Hardware Implementation of Smart Reader for Visually Impaired using Raspberry Pi”, IJAREEIE, March 2016.

[2] Faiz Alotaibi, Muhamad Taufik Abdullah, “Optical Character Recognition for Quranic image similarity matching,” IEEE 2017.

[3] Shalini Sonth, Jagadish S kallimai, “OCR based facilitator for the visually challenged,” IEEE 2017.

[4] Sanjana, Monisha, “OCR based automatic book reader for the visually impaired using raspberrypi,” IEEE,2016.

[5] Nirmala Kumari K. et al, “Image TTS Conversion using Optical Character Recognition Technique in Raspberry Pi”, IJAREEIE, May 2016.

[6] Velmurugan D.et al, “A Smart Reader for Visually Impaired People Using Raspberry PI”, IJSEC, vol. 6, pp. 2997-3001.

[7] Nagaraja L. et al, “Vision Basedon the Text Recognition using the Raspberry Pi, International Journal of Computer Applications”, NCPSIA 2015.

[8] Kwang in Kim, Keechul Jung, “Texture based approach for the text image detection,” IEEE 2013.