WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Handwritten Detection and Translation with Audio for Visually Impaired

R. Suchitra[1], K. Neelima[2], M. Sravanthi[3], K. Bhargav[4], B. Aditya Harshit[5]

[1]Associate Professor, [2,3,4,5]Student -8th Semester
Department of Computer Science & Engineering, Lendi Institute of Engineering & Technology, Vizianagaram, India.
[1]kuppilineelima9@gmail.com, [2]manyamsravanthi.1108@gmail.com, [3]bhargavkadali39*@gmail.com*, [4]Adithya.harshith179@gmail.com

ABSTRACT-

We show a camerabased system that can help visually impaired people read letters and symbols. This is a framework that helps visually impaired people read it and convert it to audio output. Understanding cell behaviour is an important area of research. It has many uses such as helping blind people and government agencies who need to know how to gather information. Based on this, we propose a system that can recognize characters and convert them to speech using text-to-speech. This work uses the method of searching for texts in different languages. All activities mentioned above related to this application are completed with publication instructions. The machine enables visually impaired people to read and understand Forensic Investigation Reports (FIR) business cards, doctor's notes and other documents and information. Our planning process is managed to produce good results and good presentations.6tr

KEY WORDS: - *Tesseract Engine, Adaptive recognition algorithm, NLTK, text blob.*

1. INTRODUCTION

In communication, this is the process of signing, editing and displaying on the internet, using mostly text format and again concluding many contracts for machines. (subtraction), intelligence, production of text and electronic data in large samples. OCR is a research topic in artificial intelligence, pattern recognition and computer vision. This work has the potential to improve the quality of life and independence of visually impaired people and opens up new opportunities for them to interact with the world around them. Textbook is a form of communication, but it can be problematic for visually impaired people who rely on touch or hearing to navigate their environment.

This project aims to solve this problem by developing an efficient text editor that can recognize and interpret text content from various sources such as letters, notes and signs. The system will use machine learning techniques and computer algorithms to identify certain features of writing, including punctuation, line spacing, and spelling. The ultimate goal is to create a simple and easy-to-use device that visually impaired people can use on a mobile phone or other portable device to capture, capture and interpret text in a timely manner. Completion of this project has the potential to improve access to information for the visually impaired and to integrate them into their independent and daily lives.

The idea of this project is to provide solutions to improve accessibility for the visually impaired. Literacy has come a long way and this project aims to use its potential to help visually impaired people interact with text in their daily lives. The project aims to build a bridge between the invisible and the written information world by creating a new system that can define and interpret the text together, and enable them to easily access important information, notes and messages. The original version of the OCR required learning for each character and image and could only work on one document at a time. Most fonts have clear recognition support.

With the support, it can create a variety of high-end digital image data formats. Text recognition is one of the most popular forms of computer vision used by many tech companies, including Apple and Google. Apple recently announced that the "Live Text" feature will be included in iOS15. This feature is similar to how Google Lens works on Android phones and how Google Search and Photos apps work on iOS. So, for this feature to work, the person needs to point the camera at the image or text.

The Live Text feature shows text in images, whether it's a document or other text. These features are based on Tesseract OCR service or technology. (Optical character recognition). OCR is the only method of converting paper documents into computerized data and is still the preferred method (other than EDI and invoices) for converting invoices to summary information that can be attributed to finance. Electronic data transmission now gives organizations a better way to manage areas such as invoicing and sales, reducing costs and allowing employees to focus on more productive activities.

TESSARECT

Tesseract is a Open Text Recognition (OCR) engine available under the Apache 2.0 license. It can be used directly or (for developers) with an API to extract text from images. Many languages are supported. Tesseract does not have a built-in GUI, but some are available on the 3rdParty websites. Tesseract is compatible with many programming languages and frameworks.

Tesseract 4.00 includes a new neural network subsystem configured as a text stream recognition engine. Derived from OCRopus' Python-based LSTM implementation, but redesigned for Tesseract in C++. Tesseract's neural network system predates TensorFlow but is compatible with it as there is a network description language called Variable Graph Specification Language (VGSL), which is also available for TensorFlow.

The ability to quickly search for content is essential for offices that need to work with a lot of analytics or big data. You can now fix problems by copying and pasting instead of editing them. Tesseract OCR is fast and reliable. This means that the structure of the paper is preserved and time is saved [2].

Introduction to Tesseract OCR

A description of the Tesseract OCR Engine defines Tesseract as follows: "Tesseract is an open source Optical Character Recognition (OCR) engine. HP first started it as a project. It was modified, developed and managed by Google and then released as open source software. It is now available at "(Smith, 2007)" It is highly portable compared to other software and supports many platforms . Its purpose is to provide less rejection and higher accuracy. Currently only the command root version is available, but there are many projects whose UIs are built on top of it and can be forked. As of now, Tesseract 3.02 version has been released and opened for use. Today Tesseract is created and maintained by Google. Supports up to 139 languages. Broken letters are some of the facts.

Architecture

Tesseract OCR is a nice engine with several layers. It works step by step as shown in the block diagram in Figure 1. image. The second step is to establish a connection between the components of the image to complete the task of extracting the signals. This step is the most important part of this cycle because it OCRs images with white text and black text. The Tesseract was the first to use this circuit to process input images. After that, the contours extracted from the image are converted to Blobs (Binary Long Objects). It is then organized into lines and regions and defined in addition to some fixed areas. After extracting, the extracted items are split into words and separated by spaces. Then it started recognizing the text, two passes. As shown in Figure 1, the first part is the time to try to recognize each word. All interesting words are taken and the last second starts typing the rest of the words. This brings with it the role of adaptive classifiers. An adaptive classifier classifies text in more detail. Adaptive classifiers need training before they can work properly. When the classifier receives some data, it must solve the problem and assign the appropriate position of the text. For more information on each step, visit The maximum number of blobs can be combined in a match. The search collects new possible states from the values in the column and evaluates them by dividing them by distribution parameters [5].

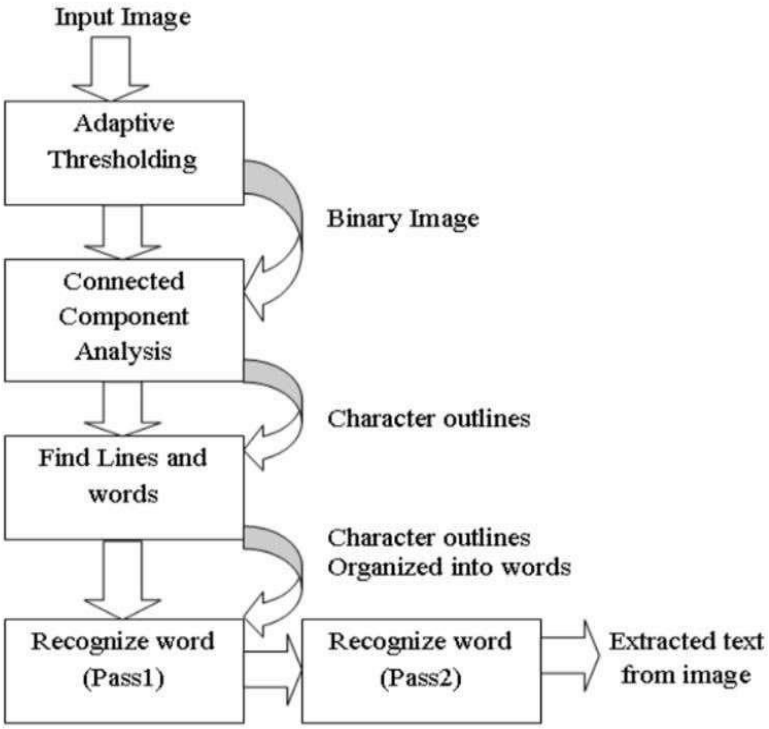

Fig 1: Tesseract flow

Working of Tesseract

Tesseract is used almost as a browser. Its interface is very simple because it uses simple commands entered from the command line. We need to write a picture along with the text. For example, the diagram shown in Figure 2.This is followed by the Tesseract. The Tesseract Basic command has only two arguments: the first is the input image containing the text, and the second is the output file, which is mostly text. Tesseract defaults to the output extension .TXT. There is no need to specify additional output data [1].

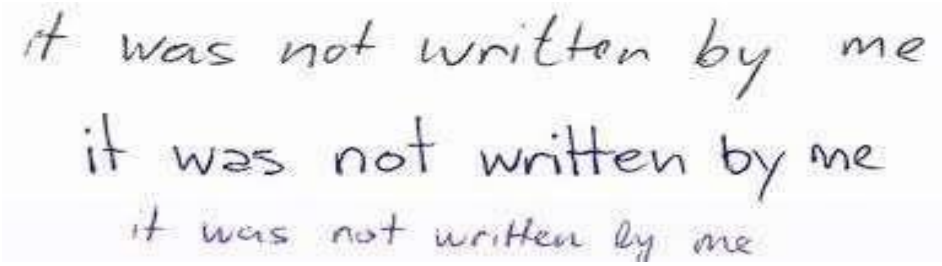

Fig 2: Image having text

All languages come with language learning materials. Language files should be stored in a place known to Tesseract. When using it in a project, it is recommended to save it in the project folder. This folder is the main Tesseract folder on your computer. Our aim in this research is to extract English characters from pictures, so we need to collect data in English. After the steps are completed, the output data shown in Figure 3 will be generated. On simple images with or without colour (grayscale). Testing shows that Tesseract can achieve good accuracy, but for some complex images with lots of background or beautiful text layers, Tesseract may provide better results if the image is in gray mode rather than colour mode. To prove this theory, we ran Tesseract on the same image in colour and grayscale mode and got different results on both. These results are then compared with the difference hypothesis processed after using the blob model in which the hypothesis is checked by editing the text, and then processed into the nltk . corpus model, which also includes the sample control text. print it out, put some real words or some random words in it. It also means that trust is used, which means how many words are correct in it, for example there are no words in the sentence entered by the previous text blob module.

2. LITERATURE SURVEY

A science study has been made from the optical behavior information and the current technology has achieved the following results. Many countries contribute to the structure and infrastructure of various projects to implement and support the digitalization process. They say the road to success is paved digitally. Digitization increases transparency and efficiency. OCR is a great way to bring simulated data to life in the world of cyberspace. Technology has long been used to create digital libraries, recognize natural phenomena, and understand written works. Using OCR technology, the data scanned or captured by the camera is an easily editable electronic product that can be easily edited, stored, copied and transmitted. Editing is very important in today's life. Until this process is implemented correctly, we rely on handwriting, which can be inaccurate. "Recording Text Using Tesseract OCR", V.Badrinarayan et al. (2019): In this article, the authors used Tesseract OCR to recognize the Tamil alphabet. They preprocessed the images using binarization and de-skew techniques and achieved 92% accuracy on the dataset [7]. "An Adaptive Recognition Algorithm for Handwritten Text Recognition" [10], S. Ahmadi et al.(2020): In this article, the authors present an adaptive recognition algorithm for text recognition. The algorithm is based on a combination of design and discrimination. The design model generates a set of candidates for each input image, while the discrimination model selects the best candidates based on the context. They achieved very high accuracy results using the IAM script. "Handwriting for Speech for the Blind" [9] , R.Zhang et al. (2020): In this paper, the authors present a real-time approach to writing and interpretation for the visually impaired. they use together tesseract OCR and neural machine translation to recognize and translate text. They achieved a recognition rate of 94.3% and correctly interpreted 87.6% of the data [16].

Protecting and accessing good physical data is hard work. Accurate data processing requires manual labor. This new technology

helps to store information in the system which makes it easy to store, manage and access information. It also provides very secure information. An example of handwriting software is Google Lens. Tesseract is an open source visual symbol engine that is considered one of the most accurate engines available today. To recognize text in an image, Tesseract first converts the input image to a binary image by matching it to the threshold. It analyzes the order of objects to store the structure of each object: lines of text are separated into words by letter spacing. Text recognition is done in two steps to improve accuracy. Follow the same level steps to recognize printed and handwritten text [14]. When reviewing the information on "Improving the Performance of the Tesseract OCR Engine", we often decided. The first thing to consider is current technology and limitations. There are many applications that are similar to normal tests. Second is the type of architecture used in the current application. A third assumption is that more pre-order types can be added without impacting access. The literature review attempts to solve these three problems and demonstrate the theoretical and empirical applicability of this approach.

The main goal here is to add the first step to improve accuracy and then build this system for the smartphone platform, in our case Android. Many publications have passed and finally started to address all these issues for further research and perhaps a successful experiment can lead to better performance than before for using algorithms or techniques [6]. Image Processing Techniques: Traditional image processing techniques such as edge detection, binarization, component analysis, and morphological operations are often used in early writing. Pre processors are often used in this scheme

to edit the input image and extract the text using a special process or law. However, these methods often require manual calibration and it can be difficult to account for written differences. Feature-based machine learning method: This method separates text and non-text by extracting features from input images and using machine learning algorithms. For text classification, properties such as texture, shape, and density are often used in conjunction with classifiers such as SVMs, decision trees, or random forests. These methods may require manual engineering, but may yield accurate results in some cases.

3. METHODOLOGY

Tesseract:

It recognizes text in digital images and outputs text in a variety of formats, including plain text and searchable PDF. Tesseract is mainly used in data scanning and data transformation applications, as well as machine learning and artificial intelligence projects. It is an open source OCR engine that can extract text or text from images. It was originally developed by HP and later taken over by Google.

That's why it's now called "Google Tesseract OCR" [2]. Regarding handwriting and voice translation for the visually impaired, Tesseract can be used as part of a larger system to recognize text and translate it into audio. This can be done by combining Tesseract with other technologies such as natural language processing and text-to-speech synthesis. Searching for text using Tesseract involves first manipulating the shape of the text to improve its quality and improve the visibility of characters.

This may include techniques such as initialization, binarization and noise reduction. After the image is processed, Tesseract can

be used to extract text from the image [6]. After the text is extracted, it can be translated into audio output using text-to-speech synthesis technology. This audio output can be played for visually impaired people who can hear the translations. can be a useful tool for creating text and machine translation for the visually impaired, but it's important to note that the accuracy of the OCR engine can be affected by images and good writing, including words and fonts. [7] can be used.

Adaptive Recognition Algorithm:

Adaptive recognition algorithms allow the system to improve its accuracy in recognizing the person who wrote the time, improving the overall user experience for applications that require text typing such as digitizing medical records, digitizing medical records, and more. This approach can help overcome handwriting recognition problems, which can vary in style and form even by the same author. By tracking these changes, the system can provide more accurate authentication and thus handle data collection more efficiently [13].

The Algorithm usually includes the following steps:

1) Image Preprocessing: Preprocessing the input image with handwritten text to remove noise, increase contrast, and improve image Quality.
2) Handwritten Text Detection: An OCR engine specially designed for handwritten text detection is used to identify handwritten text in the preview image.
3) Specific extraction: Recognized handwriting is analyzed to extract author-specific features such as line size, shape, and direction.
4) Model training: Extracted features are used to train learning models such as neural networks to get to know the author better.
5) Handwritten text recognition: The training model is used to recognize handwriting detected at level 2 and convert it to machine readable text.
6) Recommendations: The validation text is compared to the actual text and is used to update each bug test model so that the model learns from bugs and improves its performance just in time.
7) Output: The final output is text that can be used for further analysis or processing.

4. EXPERIMENT

Tesseract:

In 2006 Google took over the development of Tesseract and it has gone through many significant improvements since then.

Tesseract uses machine learning algorithms, specifically deep neural networks, to recognize text in images. It first takes steps to clear the image and make it coherent, such as removing noise, adjusting focus, and improving contrast. It then divides the image into smaller segments and performs a classification of each segment using a neural network-based classifier. Finally, it provides recognized symbols to form words and sentences [4]. Tesseract is available as both a command line tool and a library in programming languages such as C++, Java, and Python. It is widely used in many industries such as finance, healthcare and transportation to automate data processing, data extraction and data recording requests.

Textblob:

TextBlob is a Python library that provides a simple and intuitive interface for performing advanced word processing tasks such as sentiment analysis, part of speech tagging, and text translation. In the context of text detection and voice translation for the visually impaired, TextBlob can be used for

many tasks: • Language search: TextBlob can capture text input words, which is especially useful when translating text into different languages. . • Sensitivity Analysis: TextBlob can analyze the sentiment of text, which is useful in applications where the tone or sentiment of the text is important. • Translation: TextBlob interfaces various translation APIs, including Google Translate and Microsoft Translate, allowing text to be easily translated into different languages.

Python:

Python is an interpretive, high-level, general-purpose programming language. Invented by Guido van Rossum and first released in 1991, Python's design focuses on reading code and using empty field values. The language structure and goal orientation are designed to help programmers write clear, coded code for both small and large programs.

Tesseract Module:

PyTesseract is a popular Python library for optical character recognition (OCR) tasks. It's wrapped around the Google-developed Tesseract OCR engine, which recognizes text in images and converts it to machine-readable text. PyTesseract provides an easy-to-use interface for text on images or scanned documents and supports multiple languages and text formats. It can be integrated into different applications such as extracting text from images, data processing and text analysis, making it a powerful tool for knowing text in Python. PyTesseract offers many image preset options such as image rotation, binaryization and noise removal that can help improve the quality of input images and increase the accuracy of OCR results. tesseract: This is the main OCR engine module that does real text processing. Takes the input image and outputs the recognized text.

tessdata: This model contains the language data files required for text recognition. It contains special language symbols, dictionaries, and other information required for authentication. tesstils: This module provides tools for working with Tesseract, such as image processing and processing to work with OCR results.tessedit: This template contains configuration information for Tesseract that allows you to customize the behavior of the OCR engine. For example, OCR engine type, language to be used, etc. You can specify. ccmain: This module provides low-level functions for image processing such as initialization, deskew, and layout analysis. Used by the tesseract module for text recognition.

System Design :-

Design is the process of defining concepts such as architecture, modules and components, the differences between these components, and the information passing through the system. It aims to meet the specific needs and requirements of the business or organization by creating a consistent and efficient process. System Design mainly focuses on defining the architecture, components, modules, interfaces and information of the system to meet specific requirements. Design can be seen as the application of technology to production. System Design means a systematic approach to system design.

It can be a bottom-up or top-down approach, but in both cases, the process is systematic, taking into account all the changes that need to be made in the system, from the architecture to the required hardware and software to information. and how it is transmitted and exchanged system-wide. It then overlaps with systems design, systems analysis, systems engineering, and systems architecture.

The design process first appeared before World War II, when engineers were trying to solve control and communication problems. They should be able to carry out their work in a disciplined way, especially with the methods necessary for new jobs such as theory, research work, computer science.

Use case Diagram : -

Fig 3: Usecase Diagram

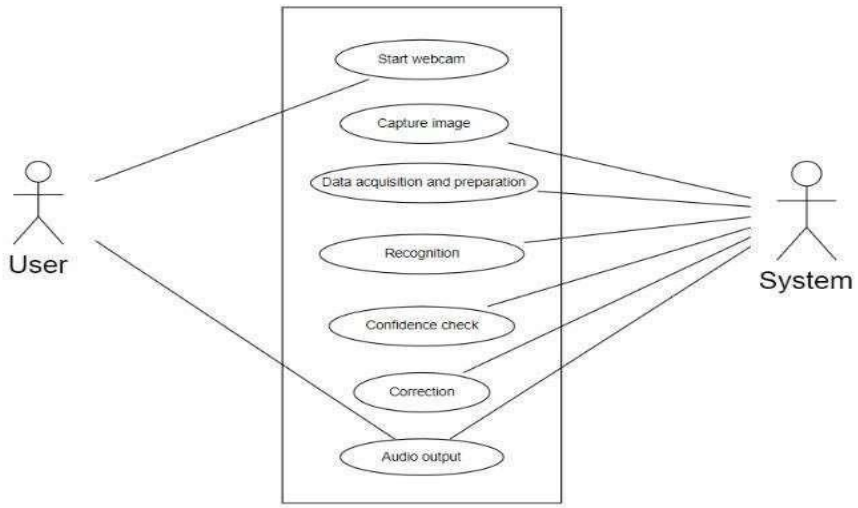

Figure 3 represents the usecase diagram for the methodology followed for generating the audio as output. The Input has been taken as a handwritten text and the text will be converted into audio. Now we will obtain final output. The handwritten text will convert into audio using tesseract (adaptive algorithm) we will get audio as output..

EXPERIMENT RESULTS :-

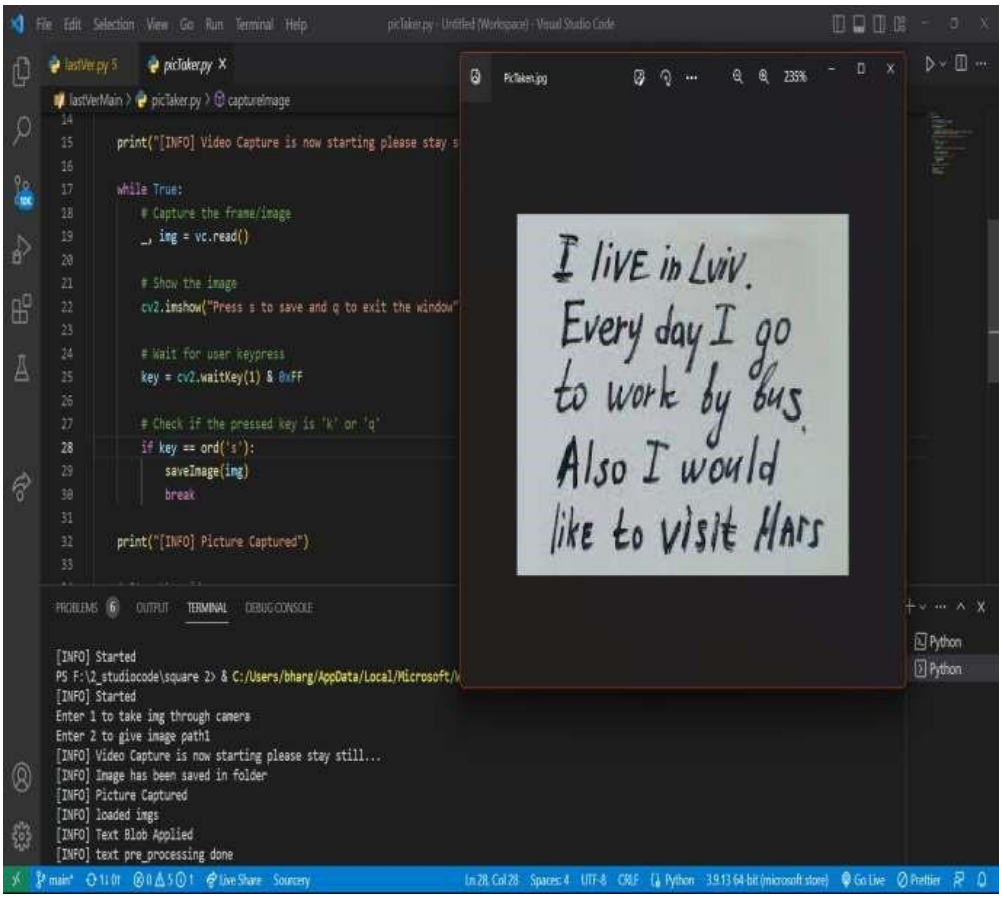

Fig 5 Input Handwritten Text

```
PS F:\2_studiocode\square 2> & C:/Users/bharg/AppData/Local/Microsoft/WindowsApps/python3.9.exe f:/2_studio
code/lastVerMain/lastVer.py
[INFO] Started
Enter 1 to take img through camera
Enter 2 to give image path2
Enter image path and make sure it's in jpgF:/2_studiocode/lastVerMain/myImages/pic2.jpg
[INFO] loaded imgs
[INFO] Text Blob Applied
[INFO] text pre_processing done
[INFO] H0
# was not written by me it was not written by me it was not writen Oy me
[INFO] H1
# was not written by me it was not written by me it was not written By me
[INFO] H2
# was not written by me
it was not walten by me
it was not writen Oy me

[INFO] Checking Confidence
[INFO] Playing the audio
```

Fig 6 : Output As Audio

5. CONCLUSION:

Tesseract OCR is a powerful and accurate text recognition tool that supports multiple languages and makes it accessible to

visually impaired people who speak and read multiple languages. The accuracy of Tesseract OCR can be improved using the information transfer system, making it easier to adapt to transfer to different scripts .

In addition, TextBlob has a good speech-making ability that translates the detected text into audio, making the invisible

accessible by voice. The use of audio screens allows visually impaired people to access written content independently without relying on vision, thus increasing their independence and ability to participate in daily life.

Overall, text search and voice translation using adaptive recognition algorithms and Tesseract OCR with TextBlobs has the

potential to improve accessibility and quality of life for the visually impaired. These technologies can improve their well-being by improving their access to written content, improving their education, work and social lives.

In addition, the cost-effectiveness and user-friendliness of this technology makes it accessible to many users, including the

visually impaired, and can be used as a test. The potential impact of this technology on the lives of the visually impaired is therefore significant and far-reaching.

In summary, the combination of Tesseract OCR with transformational recognition algorithms and TextBlob provides an

effective and powerful solution for handwriting and speech translation for the visually impaired. Accessibility is a great way to connect and empower the blind, giving them more freedom and improving their quality of life.

6. References :

Smith, R. (2007). An overview of the Tesseract OCR engine. In Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) (Vol. 2, pp. 629-633). IEEE.

Prasad, V. K., & Thirunavukkarasu, G. S. (2017). A review on Tesseract OCR engine. International Journal of Applied Engineering Research, 12(23), 13343-13347.

Singh, R. K., & Pujari, A. K. (2017). An adaptive OCR for handwritten character recognition using feature selection and SVM. Journal of Intelligent Systems, 26(4), 503-516.

Shukla, A., & Shukla, A. (2020). Text extraction and language translation for visually impaired people. International Journal of Advanced Computer Science and Applications, 11(5), 351-357.

Singh, S. (2020). Handwritten text recognition: a comprehensive review. Journal of Ambient Intelligence and Humanized Computing, 11(6), 2459-2479.

Liu, C. L., Yin, F., Wang, D. H., & Ding, X. Q. (2019). A survey of handwriting recognition. Frontiers of Computer Science, 13(1), 130-151.

Dehghani, M., Abolhassani, H., & Marvasti, F. (2018). Handwritten character recognition using deep learning methods: A comprehensive review. Journal of Computational Science, 28, 124-149.

Garg, R., Kaushik, A., & Kumar, S. (2021). Offline handwritten text recognition: A review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, e1485.

Gupta, V., Kaushik, A., & Bhatia, V. (2021). Offline handwritten text recognition using convolutional neural networks: a comprehensive survey. Soft Computing, 25(5), 3355-3380.

Saikia, H., & Bora, P. K. (2019). Handwritten text recognition: A review. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 130-135). IEEE.

Teng, F., Li, Y., Li, X., & Lin, Z. (2021). A survey on handwriting recognition: offline, online, and deep learning. Neural Computing and Applications, 33(18), 13073-13104.

Alsharidah, M., & Alshammari, F. (2021). Handwritten Arabic text recognition: A comprehensive survey. Journal of Ambient Intelligence and Humanized Computing, 12(11), 11645-11665.

Alharbi, R., & Omar, M. (2019). Arabic handwritten recognition: a comprehensive review. Journal of Ambient Intelligence and Humanized Computing, 10(6), 2257-2281a