International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 08 | Aug 2020 www.irjet.net p-ISSN: 2395-0072

Real-Time Text to Braille and Audio Convertor

Sharavana K[1], E Sivasubramanian[2], Arpitha N[3], Monish M Gowda[4], Gajanan Manjunath Naik[5], Vijayalakshmi[6]

[1,6] *Assisstant Professor, ISE Department, HKBK College of Engineering, Bangalore, KA*
[2,3,4,5] *B.E. Students, ISE Department, HKBK College of Engineering, Bangalore, KA*

ABSTRACT: *It is extremely hard for the outwardly debilitated and hard of hearing individuals to acquire fundamental and vital data required for their living. They are at a danger of being socially avoided because of poor access to data. Many attempts have been made to reduce the issues of deaf and blind. Techniques have been created which include material sensation, for example, finger Braille, manual Alphabets and a few other electronic gadgets. Yet, absence of security and absence of similarity to PC condition are issues that stay unaddressed.*

Our project aims to solve these issues. Our device which consists a camera captures an image of any page in a book, detects the text and converts this text into audio and also into refreshable Braille characters, in reality. The goal of text recognition is to convert the recognized text from hardcopy into desired format (audio or Braille). The process involves several steps like pre- processing, segmentation, feature extraction, post- processing. This device basically would be used by visually impaired people as well as students.

Keywords: Text Recognition, Feature Extraction, Refreshable Braille.

I. INTRODUCTION

Braille is a material framework to speak to content. A book in braille comprises of a few Braille cells where every cell speaks to a letter set or image. Six individual pins/dots together represent a single cell. Blend of raising high and raising low of these pins form a letter set. There is a standard Braille code for every letter set of the vast majority of the dialects. By and large, outwardly debilitated people move finger over the braille messages that are embellished on paper to peruse the content. We propose a novel text-based method for searching through noisy text.

The Braille system in English is mainly of 2 types: 6-dot Braille 8-dot Braille

8-dot Braille has the 2 extra pins/holes to represent capital letters. Our project is based only on the 6-dot Braille system.

Digits and numbers are preceeded by, the braille equivalent of the # symbol.

The braille part of our project focuses on representing:

i. The 26 English alphabets

ii. The 10 digits and any number resulting by combination of multiple digits

iii. 16 punctuation marks

This adds upto 52 combinations, which is the number of possible characters each Braille Unit/Cell of our device can represent.

India is home to 33% of the world's visually impaired populace. As per a report distributed by the National Program for Control of Blindness (NPCB),the country has about 12 million individuals with visual impairment as opposed to the global total of39 million.

A study by ET Health world says that India will have 2 million blind children by 2020. Another study states that, people who are both deaf and blind (i.e., those suffering from Deaf-blindness) cover around 2% of the worldwide populace and are bound to live in destitution and be jobless, with lower instructive results than others with no incapacities.

According to an overview led by the National Council of Educational Research and Training (NCERT),only 29.16 percent of the blind in India are part of the education system .Only 6.86 percent of the schools have access to braille books and audio content, as pointed out by the same survey.

Youngsters with hard of hearing visual impairment are up to multiple times more averse to be in school than kids without incapacities.

We aim to include 4 Braille Units/Cells.

Our device performs 2 functions - text to speech conversion to obtain the audio files for the text and creation of refreshable braille texts, in real-time.

II. LITERATURE SURVEY

[1] This paper has a proficient, inventive and ongoing cost helpful method. Rather than perusing from content pictures, this gadget permits clients to hear the substance of content picture. It blends the idea of Text to Speech Synthesizer (TTS) and Optical Character Recognition (OCR). This framework assists with connecting with PCs viably through vocal interface. The provoking errand to separate content from shading pictures.

[2] This paper designs a system which produces Braille characters on Braille display by taking input from Braille keyboard, this system also displays English characters corresponding to the Braille character on the LCD, it also displays on laptop if it is connected. Reading documents as text is also it's another capability. The challenging task here is to convert the Braille characters to English characters if the pins of Braille keyboard do not rise up and down properly.

[3] This paper operates the cell pins in a way that when the users rest their fingers over the cell it can make feel like moving fingers on cell. This paper also shows the development and design Braille book system that is low cost and portable and has one Braille cell. The disadvantage of this is paper is that when the cell moves over the fingers it's difficult to recognize the characters.

[4] This paper helps to display Braille characters by using permutations of slots that are pre-protruded. There are two eleven slots pre-protruded sliders. The display used in this paper does not need individual slot instead carry out linear sliding of the sliders and requires low power. It is not like those conventional piezo-electric displays that generate characters actuating slots dynamically and consumes large amount of power.

[5] This proposed paper is based on searching through noisy text. This system constitute words as vectors, candidates and queries into a common space acquired from the OCR, and using metric ranks the candidates to nearest- neighbor search.

III. PROPOSED METHODOLOGY

A. METHODOLOGY

The proposed text-to-braille conversion system overcomes certain limitations of the existing text-to-braille conversion system. It is based on detecting the texts from any book, document or magazine through the images captured using a camera. The text detection is achieved using Tesseract, an OCR tool from one of python's rich sets of libraries for Natural Language Processing. The OCR algorithm plays a very crucial part in our project. Python also helps in the conversion of detected texts into audio files by providing the GTTS (Google Text-to-Speech) library. The Arduino microcontroller can be programmed either using Embedded C or MicroPython. The Arduino Mega 2560 Rev3 microcontroller supports 8 bit parallel camera interface and more than 50 individually programmable gpio pins. We use 24 of the available gpio pins to activate the 24 braille pins electromechanically. The Camera which is the Miniature TTL Serial JPEG Camera, is used in our project due to its ability to transmit data with ease over a TTL serial link to the Arduino. The 4 Braille Units each consisting of 6 pins are raised electromechanically, i.e., like how a solenoid is raised or lowered due to the electromagnetic effect caused due to current flow.

B. ARCHITECTURE

The device captures images of books, then sends this image to the controller which detects the text in them and performs text to speech conversion to obtain the audio files for the book. Also, it creates writings in Braille in realtime.

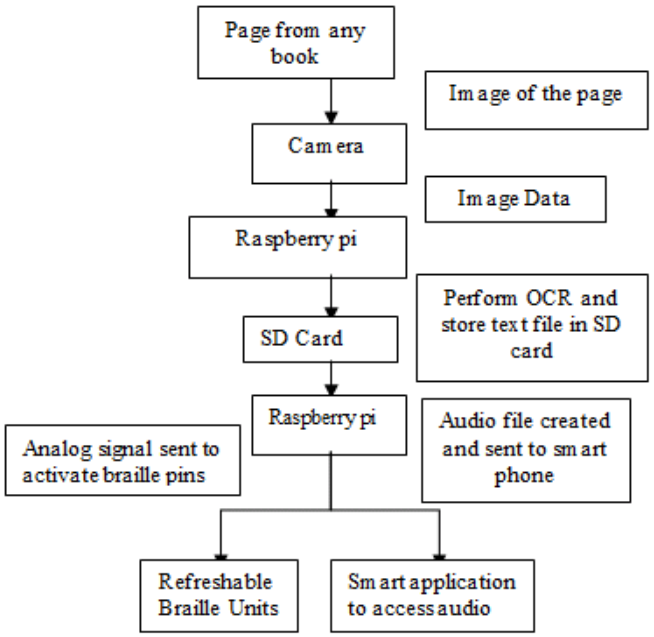


Fig 1: System Architecture

IV. REQUIREMENTS

This stage will talk about the different apparatuses that are utilized to accomplish the venture objectives and targets.

A. FUNCTIONAL REQUIREMENTS

Utilitarian prerequisites are highlights that the framework will require so as to convey or work. Right now, was imperative to accumulate a few prerequisites that will be expected to accomplish the goals set out already. The practical necessities have been accumulated are plot here.

Capture text images via camera, the miniature TTL serial JPEG camera is pointed to the book or magazines to capture image, these images are used to detect text by tesseract. The detected text is then used to convert into audio by Google text- to-speech convertor and then by Ardunio to convert into Braille. The images are stored in server or directly transmitted to devices for conversions. The visually impaired people can hear or run their hand on the Braille to hear and read the content respectively.

B. SOFTWAREREQUIREMENTS

- Python has rich set of libraries and tools for Natural Language Processing and one such tool is, the Tesseract which we use in our project to carry out fast and accurate Optical Character Recognition (OCR). The OCR algorithm plays a very crucial part in our project since it helps in detecting the text from images captured by our camera. This is the 1st step in the process of conversion of text to braille and audio.

- Python also helps in the conversion of detected texts into audio files by providing the GTTS (Google Text-to-Speech) library.

- To program the Arduino Mega 2560 Rev3 microcontrolller, we use C++.

- Arduino IDE, The Arduino Integrated Development Environment (IDE) is a cross-stage application (for Windows, macOS, Linux) that is written in capacities from C and C++.

Software Module

- Raspian OS
- Pycharm IDE
- Python 3.7
- Tessaract
- gTTS
- Xamp, PHP
- HTML,CSS
- Bootstrap, JavaScript

RASBIAN OS

- Raspbian is a Debian-based computer operating system for Raspberry Pi. Since 2015 till now it is officially provided by the Raspberry Pi Foundation as the primary operating system for the family of Raspberry Pi single-board computers.

- Raspbian was created by Mike Thompson and Peter Green as an independent project.

- The initial build was completed in June 2012.

- The operating system is still under active development. Raspbian is highly optimized for the Raspberry Pi line's low-performance ARM CPUs.

- Raspbian uses PIXEL, Pi Improved Xwindows Environment, Lightweight as its main desktop environment as of the latest update.

a. PyCharm IDE

PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company JetBrains. It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as Data Science with Anaconda.

PyCharm is cross-platform, with Windows, macOS and Linux versions. The Community Edition is released under the Apache License, and there is also Professional Edition with extra features – released under a proprietary license.

b. PYTHON 3.7

Python is a great and ground-breaking programming language that is anything but difficult to utilize (simple to peruse and compose) and with Raspberry Pi gives you a chance to associate your task to the genuine world. python punctuation is exceptionally perfect, with an accentuation on lucidness and utilizations standard English catchphrases. Begin by opening IDLE from the work area. The most effortless prologue to Python is through IDLE, a Python advancement condition. Open IDLE from the Desktop or applications menu.

c. Tesseract

Tesseract - an open-source OCR engine that has gained popularity among OCR developers. Even though it can be painful to implement and modify sometimes, there weren't too many free and powerful OCR alternatives on the market for the longest time. Tesseract began as a Ph.D. research project in HP Labs, Bristol. It gained popularity and was developed by HP between 1984 and 1994. In 2005 HP released Tesseract as an open-source software. Since 2006 it is developed by Google.

Tesseract is an open source text recognition (OCR) Engine, available under the Apache 2.0 license. It can be used directly, or (for programmers) using an API to extract printed text from images. It supports a wide variety of languages. Tesseract doesn't have a built-in GUI, but there are several available from the 3rdParty page. Tesseract is compatible with many programming languages and frameworks through wrappers that can be found here. It can be used with the existing layout analysis to recognize text within a large document, or it can be used in conjunction with an external text detector to recognize text from an image of a single textline.

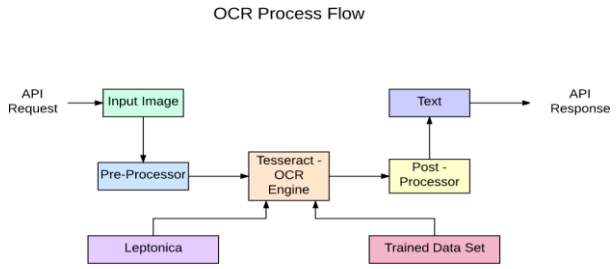


Fig2:OCR Process Flow

d. gTTS

There are several APIs available to convert text to speech in python. One of such APIs is the Google Text to Speech API commonly known as the gTTS API. gTTS is a very easy to use tool which converts the text entered, into audio which can be saved as a mp3 file.

The gTTS API supports several languages including English, Hindi, Tamil, French, German and many more. The speech can be delivered in any one of the two available audio speeds, fast or slow. However, as of the latest update, it is not possible to change the voice of the generated audio.

C. HARDWARE REQUIREMENTS

- The microcontroller which is the Arduino Mega 2560 Rev3 supports camera interfacing, wifi and bletooth connectivity and over 50 individually programmable gpio pins. We use 24 of the available gpio pins to activate the 24 braille pins electromechanically.
- The Camera which is the Miniature TTL Serial JPEG Camera, is used in our project due to its ability to transmit data with ease over a TTL serial link to the Arduino Mega 2560 Rev3.
- The 4 Braille Units each consisting of 6 pins are raised electromechanically, i.e., like how a solenoid is raised or lowered due to the electromagnetic effect caused due to current flow.
- An SD Card shield, wifi shield and bluetooth shield to provide storage network & connectivity features to the Arduino board.

Hardware Module

- Rasberry Pi 4b
- Rasberry Pi noir Camera
- 74CH595 Shift Registers
- Bread boards
- Jumper Cables
- Solenoids

Fig3.Rasberry Pi 4b

The Raspberry Pi 4 Model B is the latest version of the low-cost

Raspberry Pi computer. The Pi isn't like your typical device; in its cheapest form it doesn't have a case, and is simply a credit- card sized electronic board -- of the type you might find inside a PC or laptop, but much smaller One thing to bear in mind is that in its cheapest form, the Pi is just a bare board. You'll also need a power supply, a monitor or TV, leads to connect to the monitor -- typically a micro HDMI cable -- and a mouse and keyboard.

Once you've hooked up all the cables, the easiest way for new users to get up and running on the Pi is to download the NOOBS (New Out-Of-Box Software) installer. After the download finishes, follow the instructions here and it will walk you through how to install an OS on the Pi. The installer allows you to install various operating systems, although a good choice for first time users is the official OS, which is called Raspbian.

There are 40 Pins on a Model B altogether.

- There are three power supply pins [3.3v, 5.0v and 0v].
- 26 GPIO pins.

Rasbarian Operating system is one of the authority frameworks available for nothing to download and utilize. The framework depends on Debian Linux and is streamlined to work effectively with the Raspberry Pi PC. As we definitely know an OS is an arrangement of essential projects and utilities that keeps running on a predefined equipment, for this situation the Pi. Debian is extremely lightweight and settles on an incredible decision for the Pi.

GPIO pins can be configured to be I/O.

- GPIO pins can be enabled or disabled.
- Input values are intelligible. (typically high=1, low=0) \
- Output values can be read and are writable.
- The Raspberry Pi board have 40-pins
- Marked as P1 2.54 mm (100 mil) expansion header, arranged in a 2x13 strip.
- They provide 8 GPIO pins plus access to I^2C, SPI, UART).
- And as well as +3.3 V, +5 V and GND supply lines.
- Pin one is the pin in the first column and on the bottom row.
- Pins 8 and 10 (GPIO 14 and 15) are UART pins, designed for communicating with the Pi using the serial port.
- Pin 12 (GPIO 18) and pin 35 (GPIO 35) are hardware PWM capable, though the Pi is also able to provide software PWM through libraries such as pigpio.

RASBERRY PI 4B NOIR CAMERA

Fig4: Raspberry Pi 4b Noir Camera

The Raspberry Pi Camera Modules are official products from the Raspberry Pi Foundation. The original 5-megapixel model was realeased in 2013,and an 8-megapixel Camera Module v2 was realeased in 2016.For both iterations, there are visible light and infrared versions. A 12- megapixel High Quality Camera was released in 2020.There is no infrared version of the HQ Camera, however the IR Filter can be removed if required.

The infrared Camera Module v2 (Pi NoIR) replaced the original PiNoIR Camera Module in April 2016. The v2 Pi NoIR has a Sony IMX219 8-megapixel sensor (compared to the 5- megapixel OmniVision OV5647 sensor of the original camera).

We bundle a little square of blue gel with the Pi NoIR, which you can use with the camera to monitor the health of green plants The Pi NoIR is very popular among wildlife hobbyists: with a few infrared LEDs, you can monitor what nocturnal animals are doing in your garden without disturbing them.

74HC595 SHIFT REGISTER

74HC595 is a shift register which works on Serial IN Parallel OUT protocol. It receives data serially from the microcontroller and then sends out this data through parallel pins. We can increase our output pins by 8 using the single chip. We can also connect more than 1 shift register in parallel. So, let's say I have connected three shift registers with our microcontroller then our output pins are increased by 8 x 3 = 24.I hope I have cleared the idea, so now let's have a look at its pinout.

- As you can see from the figure it has below pinouts:
  - Pin # 1 to Pin # 7 are Output Pins Q1 - Q7.
  - Pin # 15 is also Output Pin Q0.
  - Pin # 8 is Ground.
  - Pin # 9 is Q7' (OutPut Serial Data).
  - Pin # 10 is Master Reset.
  - Pin # 11 is SHCP which is short for Shift Register Clock Input.
  - Pin # 12 is STCP which is short for Storage Register Clock Input.
  - Pin # 13 is OE which is Output Enable.
  - Pin # 14 is DS which is Serial Data input.
  - Pin # 16 is Vcc where we have to supply the power +5V.
- You can quite easily interface this shift register with different microcontrollers such as Arduino, PIC Microcontroller, Atmel etc.

Fig:5 74HC595 Shift Register

SOLENOIDS

A solenoid is a type of electromagnet, the purpose of which is to generate a controlled magnetic field through a coil wound into a tightly packed helix. The coil can be arranged to produce a uniform magnetic field in a volume of space when an electric current is passed through it. The term *solenoid* was coined in 1823 by André-Marie Ampère to designate a helical coil.

A solenoid is a coil of wire in a corkscrew shape wrapped around a piston, often made of iron. As in all electromagnets, a magnetic field is created when an electric current passes through the wire. Electromagnets have an advantage over permanent magnets in that they can be switched on and off by the application or removal of the electric current, which is what makes them useful as switches and valves and allows them to be entirely automated.

Like all magnets, the magnetic field of an activated solenoid has positive and negative poles that will attract or repel material sensitive to magnets. In a solenoid, the electromagnetic field causes the piston to either move backward or forward, which is how motion is created by a solenoid coil.

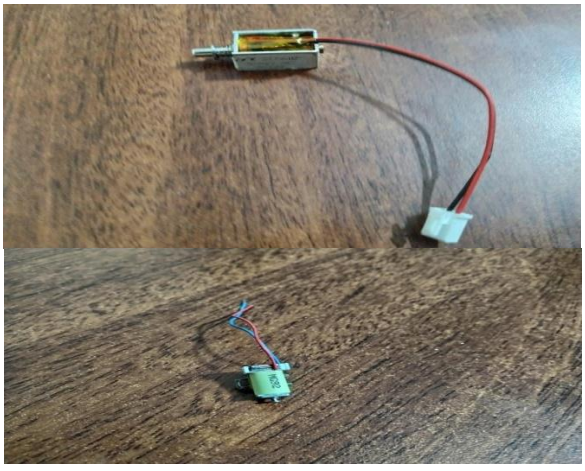


Fig7: Solonoids

RASBIAN OS

- Raspbian is a Debian-based computer operating system for Raspberry Pi. Since 2015 till now it is officially provided by the Raspberry Pi Foundation as the primary operating system for the family of Raspberry Pi single-board computers.
- Raspbian was created by Mike Thompson and Peter Green as an independent project.
- The initial build was completed in June 2012.
- The operating system is still under active development. Raspbian is highly optimized for the Raspberry Pi line's low-performance ARM CPUs.
- Raspbian uses PIXEL, Pi Improved Xwindows Environment, Lightweight as its main desktop environment as of the latest update.

V. IMPLEMENTATION

Implementation stage is where we convert our design into a working real-world system. We need to put together all the details we collected in the requirements and design stage and device a plan to give shape to our system. By incorporating all the designs and requirements we can start our implementation of the system by coding the entire system according to the architecture and the various functions and system properties we devised in the sequence diagrams as well as ensure all the use cases can be incorporated in the systems implementation.

Step-by-step implementation:

1. Give a Voice command snap to the system it captures a image from the Raspberry Pi noir camera or use the website to upload the image to Raspberry Pi 4b.
2. Raspberry Pi 4b would process this image to text using Tesseract module and performs OCR.
3. The text is saved in the Raspberry pi server with the book name and page number, so that it can be searched ,when its required again.
4. To search the already existing file give a command search to the system, then it asks for the name of the book, after entering that Raspberry Pi processor searches for the book, if it exists in memory prompts user exits and ask for read or listen(i.e, read is Braille conversion and listen is audio conversion, or it prompts it doesn't exists and to try again after loading.
5. Then after a new text is gained you can give a voice command read to convert it into Braille and listen to convert it into audio respectively. Raspberry Pi activates the Braille pins for the read command and generates the mp3 files for the listen command respectively.

VI. SYSTEM OUTCOMES

The fundamental target is to improve the understanding experience and ease of use of existing braille tablets and books, for the outwardly and hearing disabled populace. Proficiency in the use of this technology is an important part of braille literacy today.

The aim is to make a very portable and flexible product with refreshable braille technology.

A. Optical Character Recognition, OCR:

Our system represents words as vectors, projects queries and candidates obtained from the OCR into a common space, and rank the candidates using a metric suited to nearest- neighbour search.

Tesseract is an optical character acknowledgment motor for different working frameworks. Tesseract was viewed as one of the most precise open-source OCR motors.

B. Refreshable Braille:

Refreshable Braille show is an electro-mechanical gadget for showing Braille characters, for the most part by methods for round-tipped pins brought through gaps up in a level surface. The braille pins are raised and brought down utilizing a solenoid's standard.

C. Text-to-Speech Conversion:

Text to speech (TTS) convertor is the fake generation of human discourse. It changes over typical language content into GTTS (Google Text-to-Speech), a Python library and CLI apparatus to interface with Google, makes an interpretation of content to-discourse. It composes spoken mp3 information to a record, a document like article (byte string) for additional sound control, or stdout.

SNAPSHOTS

In this section we will be discussing the results of our implementation and display the snapshots of the application that has been developed. How each module that we discussed in the implementation will be represented and how the expected results are obtained. The app that has been developed can be shown with a screenshot and how the interactions happen. But the working of the model cannot be displayed in this report.

Software interfaces

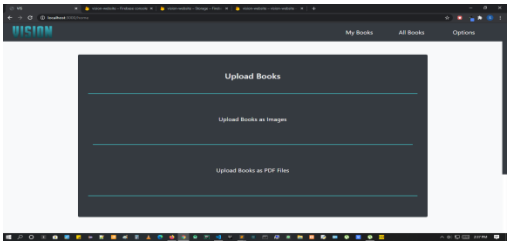


Fig8: Website Entry

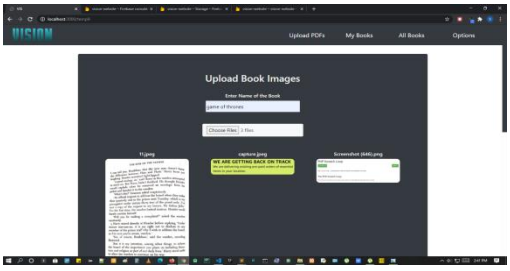


Fig 9: Upload files

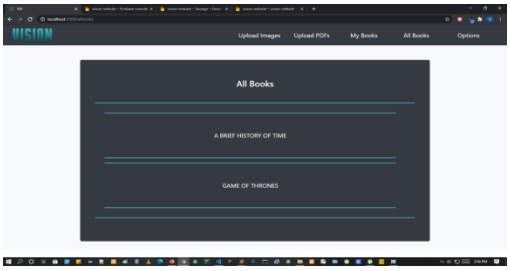


Fig10: List of files

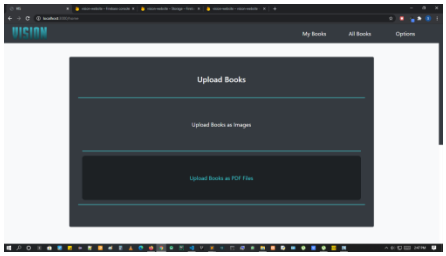


Fig11: Options in the website for file

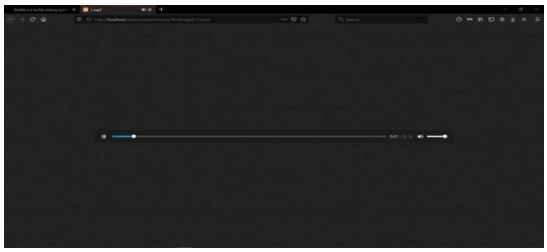


Fig12: Playing Converted text to audio

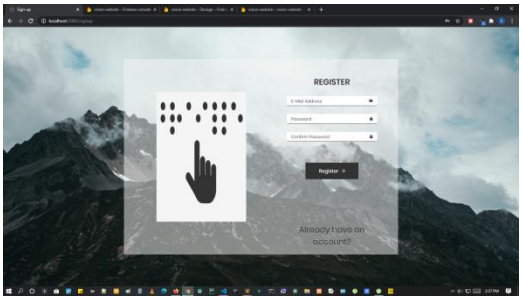


Fig13: Option to quick view the image

Fig14: Option to quick view the converted

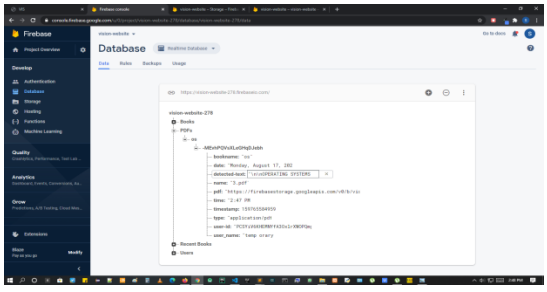
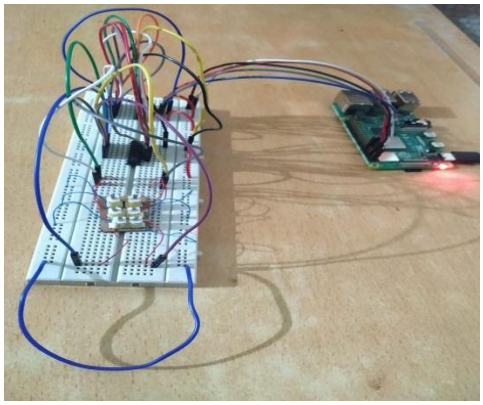


Fig15: Option to delete files

Hardware Interfaces:

Fig17: Raspberry Pi 4b model connected to Raspberry Pi Noir camera

Final Model



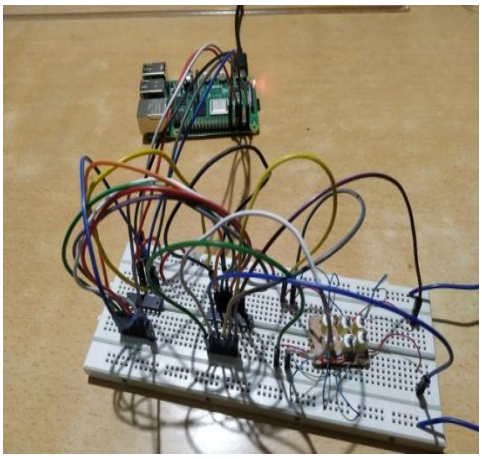


Fig18: Final Connection

VII.CONCLUSION

The present paper proposes a flexible and real-time text to Braille and audio conversion system. In its present form, the framework can be utilized in a paper, archive, book and so forth. For reading and easy understanding for the impaired audience. We've additionally added another usefulness of Text to Audio to help the outwardly hindered in understanding books, or tuning in to them straightforwardly as book recordings/content. In any case, there are a few upgrades which will be consolidated in future adaptations of the equipment interpreter. For instance, the present framework is an independent segment. We look forward to add the facial recognition feature as part of the device. Likewise for additional improvement, a multi-language-Braille interpreter will be considered. Look-into tables for changed dialects could be put away in streak memory with the goal that when interpretation of content in a specific language is required, the microcontroller loads the relating look-into table. Thereafter, we hope that the future hi-tech improvement will enlighten the life of the blind or deaf population. Our device allows social, communicational and educational barriers to be removed for people with disabilities. Thus, with some modifications in previous conventional communicating devices, we can accommodate large number of visually impaired people in an updated communication system.

References

[1] Archana A, Shinde D. Text pre-processing and text segmentation for OCR. International Journal of Computer Science Engineering and Technology. 2012:810–12

[2] Mithe R, Indalkar S, Divekar N. Optical character recognition. International Journal of Recent Technology and Engineering. 2013 Mar; 2(1).

[3] G. Devi Priya, N. Indumathi, N. Kalaimagal, A. Suriya, J.T. Vasuki, Hardware-Based Braille Pad on Mobile Phone, vol. 3, 2015.

[4] Paul H. Schimpf, "A Detailed Explanation of Solenoid Force", EasternWashington University Cheney WA USA, 2013.

[5] P. Smithmaitrie, J. Kanjantoe, P. Tandayya, "Touching force response of the piezoelectric braille cell", Disability and Rehabilitation: Assistive Technology, vol. 3, no. 6, pp. 360-365, 2008.

[6] T. Volkel, G. Weber, U. Baumann, "Tactile graphics revised:¨ the novel brailledis 9000 pinmatrix device with multitouch input", International Conference on Computers for Handicapped Persons, pp. 835-842, 2008.

[7] S. Agarwal, S. Godbole, D. Punjani, and S. Roy, "How Much Noise Is Too Much: A Study in Automatic Text Classification," Proc. Seventh IEEE International Conference on Data Mining (ICDM 2007), 2007, pp. 3–12. [2]

[8] A. Andoni and K. Onak, "Approximating edit distance in near-linear time," Proc. Forty-First Annual ACM Symposium on Theory of Computing, Bethesda, MD, pp. 199–204, June 2009. DOI: 10.1145/1536414.1536444.

[9] E. Rex, A. Koenig, R. Baker, "Foundations of braille literacy", American Foundation forthe Blind, 1994.

[10] A. Ray, S. Ghosh, B. Neogi, "An Overview on Tactile Display Haptic Investigation towards Beneficial for Blind Person", International Journal of Scientific & Engineering Research, vol. 6, no. 5, pp. 88-95, 2015.