Section: Research Paper

EXTRACTING TEXT FROM IMAGE AND BRAILLE TO SPEECH CONVERSION

R. Sathish Kumar [1]*, K. Sujai [2], R. Ajay [2], S. Vishva [2]

[1]Assistant Professor, Department of Computer Science Engineering,
Manakula Vinayagar Institute of Technology, Puducherry

[2]UG Students, Manakula Vinayagar Institute of Technology, Puducherry.

ABSTRACT

Recent substantial technical developments have allowed us to effortlessly do unanticipated tasks. Our daily lives are facilitated by technology. Every industry, from eBooks to online education, is being choked off by digitalization. One of the fastest-growing areas of science and technology in the modern world is image processing. Software is used as a computer resource in image processing to create an effective user interface. It's crucial to consider the requirements of people who are blind while talking about eBooks. For this reason alone, podcasts and audio books were first offered to us. Not only the blind, but also the illiterate who can now read the text with the aid of the provided ways, would gain from this and them is also feasible to extract text from braille visuals and turn them into voice. Here, text from an image will be extracted using a variety of methods, and it will then be turned into audio so that it may be read aloud. The major goal is to use optical character recognition and text-to-speech conversion to achieve the conversion of text in photographs to voice.

1. Introduction

The capacity of a computer to replicate intelligent human behaviour is referred to as machine learning in the field of artificial intelligence. Artificial intelligence systems are used to complete complicated tasks, much to how people approach problems. We currently live in a digital age when mobile phones and other smart gadgets are the key means of engagement and communication since technology has grown more fast than ever in recent years [15]. An application software type known as a mobile application, more commonly referred to as "an app," was developed expressly to operate on mobile devices like smartphones and tablets. Consumers frequently have access to services similar to those obtained on PCs using mobile apps [16]. Small, standalone software modules with constrained functionality are a common component of apps. Verbal communication is viewed as the best approach for successfully transferring and expressing oneself [6]. One of the fastest-growing areas of science and technology in the modern world is image processing. The mechanical or electrical conversion of handwritten or printed text pictures into machine-editable text is known as optical character recognition (OCR) [4]. We'll look at image processing utilising several methods, such image filtering. image pre- processing, image segmentation, image compression, image editing and manipulation, feature extraction and object character recognition OCR is frequently used to take text out of photos and edit them[8]. It has been long time since visually challenged/disabled individuals are able to access printed documents, except for a few documents that have been translated into Braille by clear-sighted people [10]. sPeople with low eyesight or blindness can read Braille, a system of raised dots, with their fingertips [3]. It needs to combine several Braille dots with a Braille character for Braille recognition operation which may accumulate more errors, thereby reducing the performance of Braille recognition [7]. Speech synthesis is the production of human voice or speech by a machine. It is mostly used to convert written information into spoken information for convenience [1]. Here this need to be focused on extracting text from braille dots and transforming them to speech. Text extraction from image to speech is an effective system to help the visually impaired and the illiterate to help them read banners, text from pamphlets, books and their surroundings [9]. The system is faster and more precise than humans at a wide range of pre-defined tasks, with little opportunity for human error.

The tactile writing method known as braille is used by people with visual impairments, such as those who are

blind, deafblind, or have impaired vision [11]. It can be read on embossed paper or refreshable braille displays linked to computers and cell phones [13]. A slate and stylus, a braille writer, an electronic braille note taker, or a computer linked to a braille embosser can all be used to write in braille.

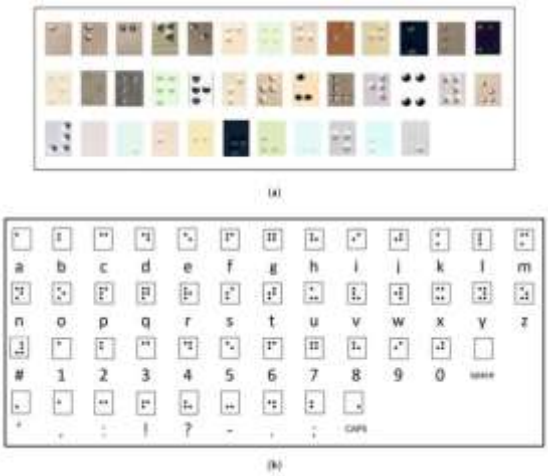

Figure 1: Braille dots

OCR, also known as optical character recognition, is the electronic or mechanical process of converting images of typed, handwritten, or printed text into machine-encoded text[5]. The scanned documents, photos of documents, scene photos (such as the text on signs and billboards in a landscape photo), or subtitle text superimposed on an image. One important approach to make any visually blind and illiterate person to analyze what is going on around him/her[18]. In order to make them more independent. And implementing this system in their Native languages is useful.

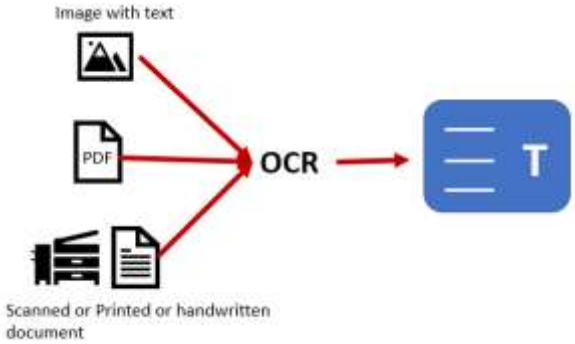

Figure 2: OCR

2. LITERATURE SURVEY

The Optical Braille Recognition (OBR) method was put out by Parmesh Kaur et al. [1] and enables the capture of Braille characters from documents, image conversion, and further processing to produce plain language counterparts. Niranjana Krupa, Sheetal Panchakshari, Sahana Ramu, and Parmesh Kaur suggested This process preserves records and makes copies of them as needed. OBR presents certain difficulties because no ink is used to produce the papers, making it difficult to distinguish between raised dots and flat surfaces. To make the pictures and dots, which are more complicated when the Braille is double-sided, clearer, several image enhancing techniques must be used. The authors of this study offer a novel approach for transforming the acquired Hindi Braille pictures into new types of images employing Utilising certain image processing methods for noise reduction and dots enhancement, Hindi text is then generated using deep learning. The text is then converted to voice using a text-to-speech (TTS) system and concatenative speech synthesis.

This project, which offers a web-based platform that turns any written fragment into an audible.mp3 format that the end user can listen to whenever they want, was proposed by Sanket Munot et al. Sanket Munot, Akshay Patil, Utkarsha Kandale, and Supriya Ambarkar suggest this work. The major goal is to process a user's input image via the web interface to extract text, which will subsequently be translated into speech. Additionally, our site offers a Mail Service through which created audio files can be delivered to the user-supplied email address. our platform provides an audio of the book to the user so that they can listen to book instead of reading. Visually impaired people require special type of books which tend to be costlier than the normal one. As we observe these students lack in receiving proper education due to their disability and we live in a society where spoon feeding is preferred more henceforth there is little effort made to assist these students. One of the goals is to give them free access to the book's audio version.

A Braille encoding scheme was proposed by G. Gayathri Devi et al. [3] to represent textual documents in a usable format for those who are blind or visually impaired. The lack of Braille-compatible reading materials makes it difficult for people who are blind or visually impaired to access basic services like employment and education. People with visual impairments often find it challenging to read text documents. Only Braille script may be used to read by those who are blind. The vast majority of printed books do not have speech or Braille adaptations. The idea that Braille symbols are created within units of space known as Braille cells was put out by G. Gayathri Devi and G. Satyanarayanan. Six raised dots, set in two parallel rows of three dots each, make up a complete Braille cell. The numerals one through six are used to identify the locations of dots. Twelve vowels, eighteen consonants, and one unique character—the ayudha ezhuthu—make up the Tamil alphabet. Thus, there are a total of 247 combinations of a consonant and a vowel, a mute consonant, or a vowel alone in the full alphabet, which consists of the 31 letters in their standalone form and an extra 216 composite characters. Sixty characters make up the Telugu script: sixteen vowels, three vowel modifiers, and forty-one consonants. Thirteen vowel letters, 36 consonant letters, and a few more symbols make up the Malayalam alphabet. The 49 phonemic letters in the Kannada language are broken down into three groups: vowels, consonants, and modifier glyphs (half-letters).There are 510 composite characters (34 * 15). This study suggests a method for successfully converting southern Indian languages to Braille.

The degree to which people who are visually handicapped can engage with their surroundings as that of a sighted person will be considerably improved by an OCR-based speech synthesis system, according to S. K. Singla et al. R.K. Yadav and S. K. Singla made a proposal The OCR-based system entails the subsequent procedural steps: Image acquisition, binarization, image pre-processing, segmentation, matching, and recognition are the first four steps. An OCR-based voice synthesis system that can be utilised as a reliable form of interpersonal communication has been discussed in this research. LabVIEW 7.1 has been used to implement the system. OCR and voice synthesis make up the developed system. In OCR, papers with printed or written characters are scanned, and a picture is acquired using IMAQ Vision for LabVIEW. Utilising LabVIEW-developed segmentation and correlation-based techniques, the various characters have been identified. Using Microsoft Speech Object Library (Version 5.1), recognised text has been turned into speech in the second portion. The created OCR-based voice synthesis system is simple to use, economical, and provides results instantly. Additionally, the programme has the necessary adaptability to be quickly changed as needed.

According to Pooja Bendale et al. [5], OCR is the basic theoretical foundation for Text to Speech synthesis.Optic character recognition is known as OCR. The method of optical character recognition (OCR) turns handwritten text and scanned or printed text images into editable text for further processing. A dependable approach for extracting text and converting it into speech dates back to the telegraph in the 20th century and was provided in the Pooja Bendale, Sanika Badhe, Sarthak Bhagat, and Pravin Rahate suggested T article. Unit selection synthesis, which started to be employed in the 1990s, was essentially a

diphone synthesis update that improved pitch. Neural networks, which can be trained to anticipate future words and sentences, became popular after machine learning techniques. The naturalness, precision, and understandability of TEXT to voice are its most crucial qualities. In the project, we used a straightforward Python technique to create TTS. In this project, we upload a text message that we wish to turn into a voice, and then we click the play button to hear that voice.

According to Teena Varma et al. [6], visual text in contemporary natural or artificial scenes may offer very significant and helpful information. In order to extract and understand the data from these photographs, the researchers have begun to digitise them. They will next perform text-to-speech synthesis (TTS). Teena Varma, Stephen S Madari, Lenita L Montheiro, Rachna S Pooojary propsed the sequence in which the information is read out for the user's convenience and benefit. TTS and text extraction can be used in conjunction to enable computer-assisted reading for those with visual impairment and reading difficulties. In this study, a novel text identification system based on connected component analysis and MSER methods is proposed. CCs, which are considered to be letter candidates, are extracted using CCs. Optical character recognition is used to identify and extract the words, and then text-to-speech synthesiser is used to turn the retrieved text into acceptable speech. On photos representing various scenes, ranging from documents to nature sceneries, the suggested algorithm is put to the test. Promising outcomes have been published, practical application in real-world circumstances.

3. EXISTING SYSTEM

Optical Character Recognition – OCR

The method of optical character recognition (OCR) turns printed texts into digital image files. It is a digital copier that automates the process of turning scanned documents into machine-readable, editable PDFs that can be shared. When you use your computer to scan a receipt, OCR will likely be evident. Your computer then stores the scan as an image. But you can use OCR to create a text document from the image with the content saved as text. The words in the image cannot be searched, changed, or tallied. Data can be extracted from scanned documents, camera images, and image-only PDFs using OCR software. It removes the need for manual data entry by making static content editable. It is a common practise to digitise printed texts so they can be electronically edited, searched, stored more compactly, displayed online, and used in machine processes like cognitive computing, machine translation, (extracted) text-to-speech. This method of digitising printed texts is used to enter data from printed paper data records such as passports, invoices, bank statements, computerised receipts, business cards, mail, printouts of static-data, or any suitable documentation. The invention of reading aids for the blind and early optical character recognition methods are also related. At the same time, Edmund Fournier d'Albe created the Opto-phone, a portable scanner that emitted tones that matched particular letters or characters when it was moved across a printed page. Emanuel Goldberg created what he called a "Statistical Machine" in the late 1920s and early 1930s for employing an optical code recognition method to search microfilm archives. He received USA Patent 1,838,389 for the innovation in 1931. Buying the patent was IBM. Pattern recognition, artificial intelligence, and computer vision are all areas of study in OCR. Early versions required training with pictures of each character and worked on one font at a time. Modern systems often handle a wide variety of digital picture file types and have a good level of accuracy for most font recognition. Some systems have the ability to reproduce formatted output that closely resembles the original page, including columns, graphics, and other non-textual elements. As a result, In the existing work will mainly have 3 steps.

- Firstly, we get an input through camera or pre saved image file or an handwritten text image which is processed by OCR.

- Secondly, the OCR will extract text from the image indeed of its background and save it as a (.txt)file.

- Lastly, the text file is processed and converted to audio as an output for the user.

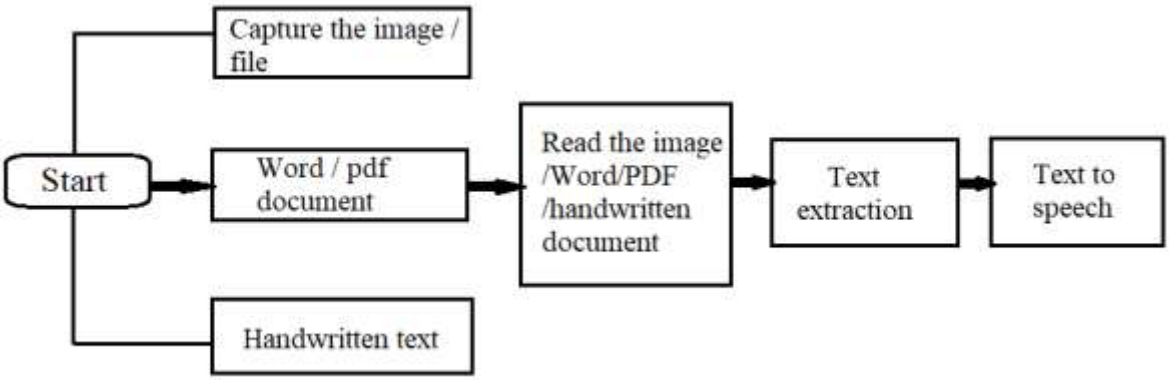

Figure:3 Existing System Architecture

4. PROPOSED WORK

Our approach overcomes most of the challenges faced by other algorithms, we are creating a clean UI which makes the user to feel more comfortable. The mobile application is created such a way where it supports in both android and ios. Creating a basic UI for our app and adding features like text extraction from images and converting the text into audio, translating the text into any desired language and generating voice note of it. All these functions are performed automatically in this application.

Creating a basic UI for our app and adding features like text extraction, converting the text into an audio , translating the text into any desired language and also generating a voice note of it. All these function are performed manually in this phase. Making all the manual function to work automatically. Enhancing the UI of the app that suits this work. The UI now will take snapshots of text by itself without manual control. This phase will support some extra important languages. The main objective of this module is to extract braille dots and convert them into the language of our choice. This will help the visually challenged people while learning braille for first time such that they can easily scan a braille paragraph with their mobile phone and understand the meaning of it. Multiple testing will be done in order to ensure proper working. Making the app to work in both platforms like Android & IOS. Additional features will be added as per the suggestion of mentors and project reviewing panel members.

5. IMPLEMENTATION

To implement the Braille to Text and Text to Braille conversion system, we followed few outlined below. We defined the requirements for the Braille to Text and Text to Braille conversion system, including the ability to convert Braille characters into their corresponding ASCII text equivalents and vice versa. We also defined the requirements for handling various formatting elements such as line breaks and paragraph breaks. We designed the Braille to Text and Text to Braille conversion system using Python programming language. Our system was designed to convert Braille characters to ASCII text and vice versa using a simple algorithm that maps Braille symbols to ASCII characters. We wrote the code for the Braille to Text and Text to Braille conversion system using Python programming language. Our system uses a combination of string manipulation and regular expressions to handle the conversion process. We tested the system to ensure that it worked as expected. We created several test cases to ensure that the system handled different input scenarios correctly. We also debugged any issues that arose during testing.

We optimized the system by making use of Python libraries to speed up the conversion process. We also added error handling to handle any unexpected input scenarios. We documented the Braille to Text and

Text to Braille conversion system by creating a user manual that explains how to use the system, how it works, and how it was implemented. We also provided details on the performance of the system and any limitations or areas for improvement. The system was able to accurately convert Braille characters to ASCII text and vice versa, and handled various formatting elements such as line breaks and paragraph breaks. Overall, our implementation of the Braille to Text and Text to Braille conversion system was successful, and the system met all of the defined requirements. The system was able to accurately convert Braille characters to ASCII text and vice versa. The system was also optimized for speed and included error handling to ensure that it functioned correctly even in unexpected input scenarios.

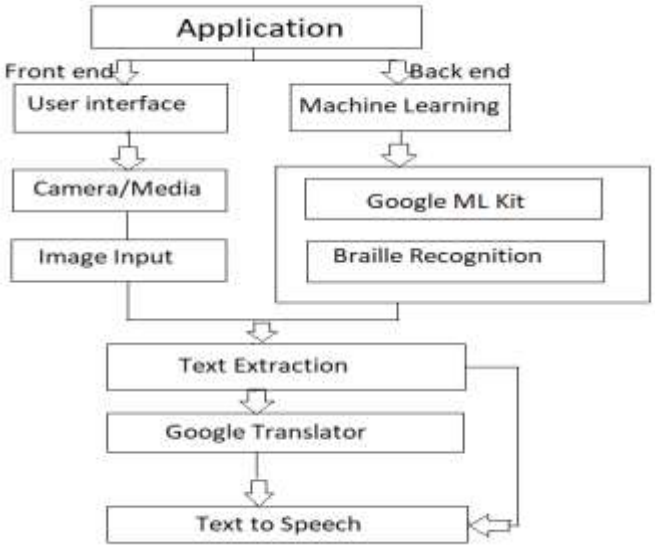

Figure: 4 System Architecture

6. RESULT AND DISCUSSION

I. APPLICATION UI:

User interface (UI) design is a technique used by designers to create user interfaces for software or electronic devices, with an emphasis on aesthetics or style.

Figure:5 App UI

II. Braille to Text:

Braille to text and text to braille are two important processes in the field of accessibility and assistive technology for individuals with visual impairments. In this discussion, we will explore the results and implications of these processes.

The process of converting Braille to text involves the translation of raised dots on a page into readable characters and words. This process can be achieved through the use of specialized software or hardware, such as a Braille display or scanner. The accuracy of this process depends on the quality of the Braille document and the technology used to perform the translation.

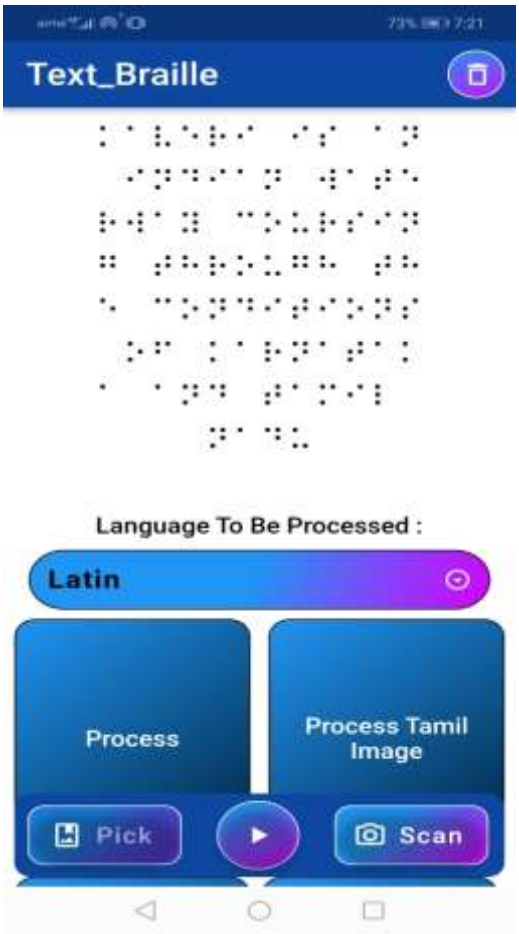

Figure:6 Picking Braille Image

Figure:7 Braille to Text Conversion

Implications:

The ability to translate between Braille and text is critical for individuals with visual impairments. Access to Braille documents allows for greater independence and accessibility in education, employment, and other areas of life. Additionally, the ability to translate text into Braille allows for the creation of Braille documents, making written materials accessible to those who rely on Braille.

Advancements in technology have greatly improved the accuracy and speed of these processes, making them more accessible and widely available. However, there is still room for improvement in terms of accuracy and efficiency, particularly with complex or technical documents.

III. Text to Braille:

Conversely, the process of converting text to Braille involves the translation of written language into raised dots on a page. This process is often achieved through specialized software or hardware, such as a Braille embosser. The accuracy of this process also depends on the quality of the technology used, as well as the input text.

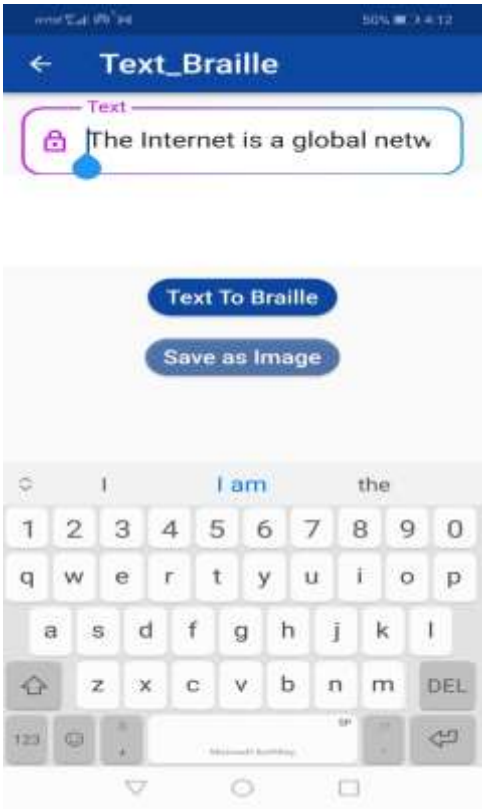

Figure:8 Importing Text

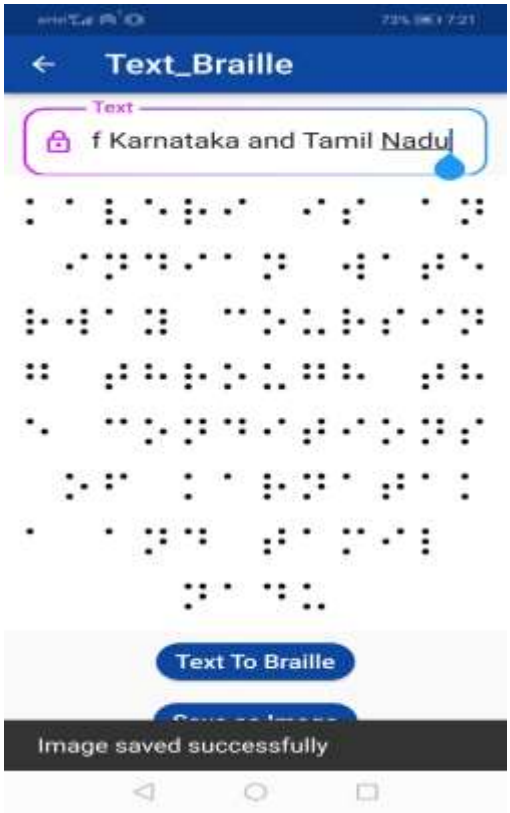

Figure:9 Text to Braille Conversion

In conclusion, the processes of converting Braille to text and text to Braille are essential for promoting accessibility and inclusivity for individuals with visual impairments. Continued advancements in technology will further improve the accuracy and accessibility of these processes, leading to a more inclusive and equitable society.

IV. Image To Text Conversion:

A visual depiction of something is an image. It can provide information to the visual system in two dimensions, three dimensions, or in another way. An artefact that resembles a topic, such a photograph or other two-dimensional representation, might be considered an image. An image in the context of signal processing is a distributed colour amplitude.

Text translation is the process of translating spoken or written words from one language to another while maintaining the original content's style, intent, and meaning. This can be accomplished using a variety of techniques, including automated translation using computer algorithms or manual translation by a human translator. A human translator who is fluent in the source and destination languages does manual translations. The information is then rewritten in the target language by the translator after reading the source text and comprehending its meaning. This approach calls for a

thorough comprehension of the subtleties of both languages as well as expertise in and sensitivity to other cultures.

Figure:10 Image to Text Conversion

Figure:11 Text Translation

On the other hand, automated translation translates text from one language to another using computer algorithms. Statistical models, rule-based systems, and artificial intelligence (AI) are used in this procedure to examine the source text and produce an equivalent text in the target language. Statistical machine translation, neural machine translation, and rule-based machine translation are a few of the different methods used in automated translation. In order to find patterns and rules that guide the translation process, statistical machine translation analyses enormous quantities of parallel texts in both the source and target languages. Deep learning models are used in neural machine translation to figure out how to map the source and target languages. On the other side, rule-based machine translation uses grammatical structures and pre-established rules to translate text.

The Translator service's text translation feature is a cloud-based REST API that makes use of neural machine translation technology to offer rapid and accurate source-to-target text translation in real time for all languages that are supported.

The rules of the target language, which the user is translating to, are used to create the translation once an automatic translation software system has analysed the sentence structure in the source language, which the user is translating from.

Figure:12 Image to Text Conversion

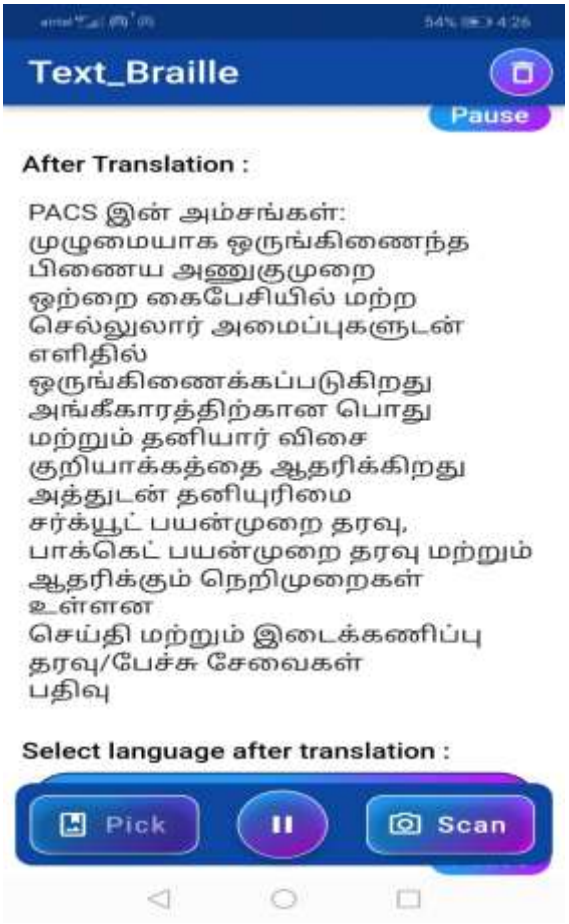

Figure:13 Text Translation

V. Text Scanner:

Text detection using a camera involves using software or algorithms to recognize and extract text from images or videos captured by a camera in real-time. This technology is often used in mobile devices and applications, allowing users to quickly and easily capture text information from their surroundings. The success of text detection using a camera depends on various factors, such as the quality of the camera and lighting conditions, the complexity of the text and its background, and the accuracy of the text detection and recognition algorithms being used. Despite these challenges, text detection using a camera has become an increasingly popular and useful technology for a variety of applications.

One common example of text detection using a camera is within document scanning apps. These apps use a camera to capture an image of a document, and then use text detection algorithms to locate and extract text from the image. The extracted text can then be saved or edited as needed. The success of text detection using a camera depends on various factors, such as the quality of the camera and lighting conditions, the complexity of the text and its background, and the accuracy of the text detection and recognition algorithms being used. Despite these challenges, text detection

using a camera has become an increasingly popular and useful technology for a variety of applications.

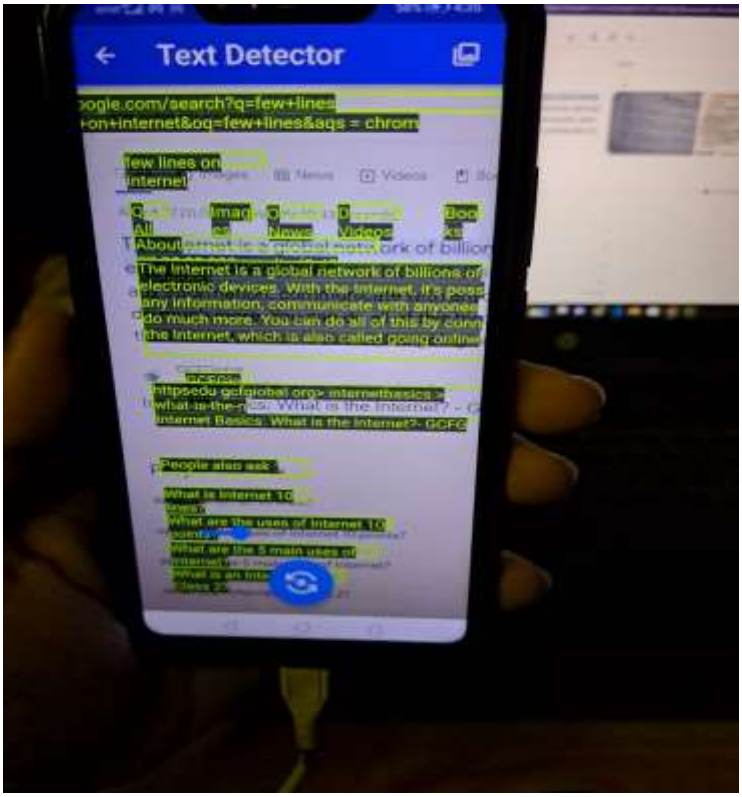

Figure:14 Text Detection from Browser

7. CONCLUSION

In conclusion, Braille to text and text to Braille conversion processes are crucial for fostering accessibility, inclusivity, and literacy for people with visual impairments. Technology advancements have significantly increased the precision, efficiency, and adaptability of these procedures, enabling wider access to education, work, and other facets of life. To fulfil the changing demands of people with visual impairments, it is crucial to keep promoting and supporting Braille reading while also upgrading Braille technology. By doing this, we can build a more just and inclusive society where everyone has the chance to realise their full potential. It's also crucial to remember that text-to-speech and Braille technology can work well together. While text-to-speech technology can be helpful for rapidly and effectively accessing information, Braille gives a distinctive tactile experience and fosters the development of reading. According to a person's needs and preferences, it is crucial to promote both Braille and text-to-speech technology. The availability and affordability of Braille and text-to-speech technology for those with visual impairments must also be prioritised. The Braille or text-to-speech output's quality must be taken into account. This entails making sure that the translation of the original text for Braille is accurate. This refers to making sure the voice output for text-to-speech is understandable and clear. To guarantee that people with visual impairments receive the best possible output, quality assurance measures should be in place. In conclusion, advancing Braille and text-to-speech technologies, boosting literacy, assuring affordability and quality, and raising awareness and education are all necessary to enhance accessibility and inclusivity for people with visual impairments. By doing this, we can make sure that people with visual impairments may participate fully in society and have equitable access to school, employment, and other facets of life.

REFERENCE

[1]. Parmesh Kaur, Sahana Ramu, Sheetal Panchakshari, Niranjana Krupa "Conversion of Hindi Braille to Speech using Image and Speech Processing" IEEE, 2020.

[2]. Sanket Munot, Akshay Patil, Utkarsha Kandale, Prof. Supriya S Ambarkar "Image To Speech Conversion Website" IJSART - Volume 5 Issue 9 – SEPTEMBER 2019.

[3]. Dr. G.Gayathri Devi, G.Sathyanarayanan "Conversion Of Text Image Document in Southern Indian Languages Into Braille For Visually Challenged People" International Journal of Computer Engineering & Technology (IJCET) 2018.

[4]. S. K. Singla and R.K.Yadav"Optical Character Recognition Based Speech Synthesis System Using LabVIEW" Journal of Applied Research and Technology 2021.

[5]. Pooja Bendale, Sanika Badhe, Sarthak Bhagat, Pravin Rahate "Text Extraction from Image, Word, PDF and Text-to-Speech Conversion" International Research Journal of Engineering and Technology (IRJET) 2022.

[6]. Prof. Teena Varma, Stephen S Madari, Lenita L Montheiro, Rachna S Pooojary "Text Extraction From Image and Text to Speech Conversion" International Journal of Engineering Research & Technology (IJERT) 2021.

[7]. Liqiong Lu , Dong Wu, Jianfang Xiong, Zhou Liang, Faliang Huang"Anchor-Free Braille Character Detection Based on Edge Feature in Natural Scene Images" Hindawi 2022.

[8]. Nivetha, Kameshwari " Image to Text and Speech Converter" International Research Journal of Engineering and Technology (IRJET) 2020.

[9]. Swami Datta Subramaniam, Vinayak Vasanta Shirsat "Smat Image To Text To Speech Using Deep Learning" Research Gate Publication Journal 2021.

[10]. Priya Bhardwaj, Ritesh Kumar Sharma "Image Text Reader for Visually Impaired in Desired Language" IEEE 2020 ,Conversion Of Text Image Document.

[11] L. M. Michelson, J. Kathleen, and D. Morgan, "Using a new electronic Brailler to improve Braille learning at the Florida school for the deaf and blind," Journal of Visual Impairment & Blindness, 2015.

[12] U. Beg, K. Parvathi, and V. Jha, "Text translation of scanned Hindi document to braille via image processing," Indian Journal of Science and Technology, 2017.

[13] M. Z. Iqbal, S. Shahid, and M. Naseem, "Interactive Urdu braille learning system for parents of visually impaired students," in Proceedings of the 19th International ACM SIG ACCESS 2017.

[14] S. Iqbal, A. Ali, M. Younus, M. Huzaifa, and Z. Abbas, Braille Instant Translator, pp. 327-328, National University of Computer and Emerging Sciences2017.

[15] J. Siqueira, F. A. Soares, and C. R. Silva, "Braille Ecran: a braille approach to text entry on smart phones," in Proceedings of the IEEE 40th Annual Computer Software and Applications 2016.

[16] M. Alnfiai and S. Sampalli, "SingleTap Braille: developing a text entry method based on braille Patterns using a single tap," in Proceedings of the 11th International Conference on Future Networks and Communications 2016.

[17] M. Alnfiai and S. Sampalli, "BrailleEnter: a touch screen braille text entry method for the blind," Procedia Computer Science 2017.

[18] M. Li, M. Fan, and K. N. Truong, "Braille Sketch: a gesturebased text input method for people with visual impairments," in Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, 2017.