Visvesvaraya Technological University
BELAGAVI, KARNATAKA.

ವಿಶ್ವೇಶ್ವರಯ್ಯ ತಾಂತ್ರಿಕ ವಿಶ್ವವಿದ್ಯಾಲಯ
ಬೆಳಗಾವಿ, ಕರ್ನಾಟಕ

A PROJECT REPORT

ON

"Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning"

Submitted to Visvesvaraya Technological University in partial fulfillment of the requirement for the award of Bachelor of Engineering degree in Computer Science and Engineering.

Submitted by

Name	USN
Aishwarya S	4JN20CS004
Chandana C S	4JN20CS025
Nikhil M	4JN20CS064
Ranjan P N	4JN20CS082

Under the guidance of

Mr. Narendra Kumar S B.E,M.Tech,(Ph.D)
Assistant Professor, Dept. of CS&E,
JNNCE, Shivamogga.

Department of Computer Science & Engineering
Jawaharlal Nehru New College of Engineering
Shivamogga - 577 204

May 2024

National Education Society ®

Jawaharlal Nehru New College of Engineering

Department of Computer Science & Engineering

CERTIFICATE

This is to certify that the project entitled

"Smart Book Reader for Visual Impairment Person Using Iot Device and Deep Learning"

Submitted by

Name	USN
Aishwarya S	4JN20CS004
Chandana C S	4JN20CS025
Nikhil M	4JN20CS064
Ranjan P N	4JN20CS082

Students of 8th semester B.E. CS&E, in partial fulfillment of the requirement for the award of degree of Bachelor of Engineering in Computer Science and Engineering of Visvesvaraya Technological University, Belagavi during the year 2023-24.

Signature of Guide

Mr.Narendra Kumar S B.E., M.Tech.,(Ph.D).
Assistant Professor, Dept. of CS&E

Signature of HOD

Dr. Jalesh Kumar B.E., M.Tech., Ph.D.
Professor and Head, Dept. of CS&E

Signature of Principal

Dr. Y Vijay Kumar B.E., M.Tech., Ph.D.
Principal, JNNCE

Examiners: 1. ________________ 2. ________________


DOI: 10.55041/IJSREM34565

ISSN: 2582-3930
Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT
An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Aishwarya S

in recognization to the publication of paper titled

Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning

published in IJSREM Journal on Volume 08 Issue 05 May, 2024

www.ijsrem.com
Editor-in-Chief
IJSREM Journal
e-mail: editor@ijsrem.com


DOI: 10.55041/IJSREM34565

ISSN: 2582-3930
Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT
An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Chandana C S

in recognization to the publication of paper titled

Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning

published in IJSREM Journal on Volume 08 Issue 05 May, 2024

www.ijsrem.com
Editor-in-Chief
IJSREM Journal
e-mail: editor@ijsrem.com


DOI: 10.55041/IJSREM34565

ISSN: 2582-3930
Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT
An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Nikhil M

in recognization to the publication of paper titled

Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning

published in IJSREM Journal on Volume 08 Issue 05 May, 2024

www.ijsrem.com
Editor-in-Chief
IJSREM Journal
e-mail: editor@ijsrem.com


DOI: 10.55041/IJSREM34565

ISSN: 2582-3930
Impact Factor: 8.448

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT
An Open Access Scholarly Journal || Index in major Databases & Metadata

CERTIFICATE OF PUBLICATION

International Journal of Scientific Research in Engineering & Management is hereby awarding this certificate to

Ranjan P N

in recognization to the publication of paper titled

Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning

published in IJSREM Journal on Volume 08 Issue 05 May, 2024

www.ijsrem.com
Editor-in-Chief
IJSREM Journal
e-mail: editor@ijsrem.com

ABSTRACT

Individuals primarily have good vision where they can see the outer world easily and they can read and write normally. Braille script is created for visually impaired to read and write just like other people who have normal sight. Visually impaired (VI) according to WHO (2023) data 15 million people with significant vision loss.

The Braille system consists of cells with six raised dots, and each raised dot has a number from one to six organized in two columns. It is vital to allow visually impaired people to keep pace with the world around them. Providing Braille-assisted technology and integrating it in daily living are necessary to make the lives of visually impaired people more comfortable and efficient for communicating with others. After collection of Braille datasets, it is pre-processed where the noise or disturbances in the script is processed and removed. Later, the braille script is segmented and converted to individual characters and it is converted to other specified Indian language script and this is converted as voice output. By this we can focus on multiple languages and convert other language braille script to voice output. For this we can use IoT and deep learning, where support for multiple users can be provided.

ACKNOWLEDGEMENT

On presenting the Project report on "**Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning**'', we feel great to express our humble feeling of thanks to all those who have helped us directly or indirectly in the successful completion of phase 1.

We are grateful to our institution Jawaharlal Nehru New College of Engineering and Department of Computer Science Engineering for imparting us the knowledge with which we can do our best.

We would like to thank our respected guide **Mr. Narendra Kumar S** B. E, M.Tech(Ph.D), Assistant Professor, Department of CS&E, who have helped us a lot in making this project and for their continuous encouragement and guidance throughout the project work.

We would like to thank our respected project coordinators **Dr. Poornima K M**, **Mr. Hiriyanna G S, Dr. Ganavi M** and **Mrs. Pushpa R N**, who have extended their warm support with respect to all aspects of project.

We would like to thank **Dr. Jalesh Kumar**, HOD of CS&E Dept and **Dr. Y Vijay Kumar**, The Principal, JNNCE, Shivamogga for all their support and encouragement.

Finally, we also would like to thank the whole teaching and non-teaching staff of Computer Science and Engineering.

Thanking you all,

Project Associates

Aishwarya S **4JN20CS004**

Chandana C S **4JN20CS025**

Nikhil M **4JN20CS064**

Ranjan P N **4JN20CS082**

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The ability to see the world through vision is the biggest boon for a human being. However, some unfortunate people do not have this opportunity. The visual impaired who have no vision or low sighted vision use other techniques to read or write, and they can learn easily. But the problem is, not everyone can understand the Braille script and learn fast. As a result, to improve their efficiency in studies visually impaired can listen to the converted braille script. This can be achieved by IoT device, deep learning and corresponding output can be provided in the form speech.

In the realm of assistive technologies, the intersection of Internet of Things (IoT) devices and Deep Learning has given rise to a transformative innovation: the Smart Book Reader for individuals with visual impairment. As we embark on a comprehensive literature survey, our focus is directed towards understanding the landscape of existing research, technological advancements, and user experiences surrounding this groundbreaking convergence.

Braille is the system of reading and writing used by people who are blind where they feel raised dots on a Braille page with tips of their fingers. It is coming from the military system which used raised dots to send the message. The Braille code system has become widely used by several communities because of its simplicity, comfortably and usability in reading and writing for visually impaired. Braille was applied or translated into several languages including Arabic and English languages.

The fusion of IoT devices and Deep Learning algorithms has opened new possibilities in creating assistive tools that go beyond traditional solutions for individuals with visual impairments. Our literature survey aims to unravel the tapestry of research and development efforts, exploring how various scholars and technologists have approached the challenge of enhancing accessibility to literature for the visually impaired.

Deep Learning is a subfield of machine learning that focuses on artificial neural networks and algorithms inspired by the structure and function of the human brain. It has gained significant attention and popularity in recent years due to its remarkable success in solving complex problems across various domains.

Deep Learning excels at automatically learning hierarchical representations of data, allowing it to discover intricate features and patterns. It eliminates the need for manual feature engineering, enabling the algorithm to learn relevant features directly from the data.

Deep learning has significantly advanced braille script identification by automating the recognition process and continues to contribute to the development of accurate, user friendly braille script identification system.

Convolutional Neural Networks (CNNs) are a class of deep neural networks specifically designed for processing and analysing visual data, such as images and videos. They have been highly successful in various computer vision tasks and have become a fundamental architecture in the field of deep learning.

A DCNN typically consists of multiple convolutional layers followed by pooling layers, creating a deep architecture. The depth of the network allows it to learn intricate and abstract features from input data. Deep CNNs utilize convolutional layers to apply convolution operations on input data, extracting features using filters or kernels. Each convolutional layer captures different levels of abstraction.

Fig 1.1 Basic Overview

1.1 PROBLEM DESCRIPTION

The main problem addressed by this project is the difficulty faced by blind individuals or those with low vision when it comes to reading books. Braille is the language used to read books by visual impairment people by touch. This difficulty is overcome by "smart book reader for Visual Impairment People using IoT Device and Deep Learning" where the visual impaired people can easily be understood by audio.

1.2 MOTIVATION FOR THE PROJECT

The motivation behind the creation of the Smart Book Reader for visually impaired individuals is deeply rooted in the aspiration to empower and inclusively integrate this community into the world of literature and knowledge. By allowing visually impaired individuals to read and comprehend books at an accelerated pace without the need for Braille learning, the Smart Book Reader seeks to eliminate barriers to education and information. The overarching goal is to provide equal opportunities, bridging the gap between visually impaired and sighted individuals in terms of accessing written content. Furthermore, by addressing societal challenges faced by the visually impaired, such as slow reading and the complexity of Braille, this technology-driven solution promotes a more accessible and user-friendly means of consuming information. The use of cutting-edge technologies, including image processing and deep learning, reflects a commitment to leveraging innovation to enhance the lives of visually impaired individuals and foster a more inclusive society.

1.3 OBJECTIVES

- Accessibility and User-Friendly Design.
- Text Recognition Accuracy.
- Real-Time Text-to-Speech Conversion.
- Environmental Adaptability.
- Continuous Learning and Improvement.
- Support for multiple languages, providing users with the flexibility to read content in their preferred language.
- Utilization of IoT Support for multiple users on a single device, making it suitable for shared environments such as classrooms or community centres.

1.4 ORGANIZATION OF THE REPORT

The further chapters in the report include the brief discussion of the project and its working. The next chapter has introduction to the areas and there is explanation about Deep learning, Colab, IoT, Raspberry Pi, and methods such as grayscale, gaussian blur, and adaptive threshold which are explored in this project. Further chapter has the system architecture which includes real-time image pre-processing, training image dataset, output matrix, SoftMax layer and final voice output. The next chapter includes results and analysis with respect to hyperparameter. Hyperparameter tuning is the process of finding optimal values for a learning algorithm.

CHAPTER 2

LITERATURE SURVEY

In this section the literature survey of various papers is explained briefly.

2.1 BRAILLE RECOGNITION FOR REDUCING ASYMMETRIC COMMUNICATION BETWEEN THE BLIND

- Bi-Min Hsu, Department of Industrial Engineering & Management, Cheng Shiu University, Kaohsiung 83347, Taiwan; 30 June 2020

Methodology

- This research presents a novel approach to convert images of braille into English text by employing a convolutional neural network (CNN) model.
- Optical Braille recognition (OBR) have been created to convert braille images into text and to account for the differences between normal characters and braille symbols.
- The CNN model with the best hyperparameters is saved then used in the OBR algorithm.

Advantages

- CNNs are adept at extracting hierarchical features from images, enabling them to capture intricate details in hand gestures. This allows for more accurate recognition and classification of different gestures.
- CNNs are inherently invariant to spatial transformations such as translation, rotation, and scaling. This makes them resilient to variations in hand placement, orientation, and size within the input images or video frames.

Disadvantages

- While the CNN model used in this study achieved a high prediction accuracy of 98.73% on the test set, there is still a possibility of errors in the recognition of braille characters. This could lead to miscommunication between the blind and non-blind individuals.
- The reliance on technology for braille recognition could be a disadvantage for individuals who prefer traditional methods of communication or do not have access to the necessary technology.
- The dataset used in this study contained braille images corresponding to 71 different English characters. This suggests that the technology may not be able to recognize

braille characters from other languages, limiting its usefulness for non-English speaking individuals.

Result

- The image is captured from the web camera. Following the capture of an image, the image is identified and classified using the CNN algorithms.
- The braille script is matched to the dataset, and the braille script with the highest matching rate is categorized.
- The performance of the CNN model yielded a prediction accuracy of 98.73% on the test set.

2.2 PRE-PROCESSING THE BRAILLE IMAGE FOR IMPROVING OPTICAL BRAILLE RECOGNITION PERFORMANCE

- Vishwanath Venkatesh Murthy *et al,* International Journal of Advanced Research in Computer Science, July-August 2020

Methodology

- Image Binarization: The initial step involves converting a colour image into a binary image. RGB values of each pixel are combined to calculate brightness.
- Histogram Equalization: Enhances the intensity contrast of pixels in an image.
- Linear Spatial Filtering: Linear smoothing filters include Mean and Gaussian filters. Linear edge-enhancing filters include Sobel, Prewitt, and Laplace filters.
- Noise Reduction: Mean filter reduces noise by blurring the image. Median filter reduces noise with less blurring compared to the mean filter.
- Gaussian Filtering for noise removal.
- Fuzzy Sets are used for transforming image intensity.
- Edge Detection algorithms use various filtering, amplification, and threshold operations.

Advantages

- Many of the mentioned filters and processes, such as median filtering, Gaussian filtering, and mean filtering, are effective in reducing different types of noise in images, leading to cleaner and more visually appealing results.

- Filters like Sobel, Prewitt, and Laplacian are widely used for edge detection. They highlight the boundaries and edges of objects in an image, aiding in object recognition and segmentation.
- Histogram equalization and other preprocessing techniques can help standardize the intensity distribution in images, making them more suitable for certain types of analyses or comparisons.
- Preprocessing can help align and register images more accurately.

Disadvantages

- It may result in a loss of fine details in the image. This blurring effect can be undesirable in scenarios where maintaining sharpness is critical.
- Converting colour images to grayscale using simple averaging (as described in the text) may lead to a loss of important colour information.
- Some preprocessing techniques may not be suitable for certain types of images or may require careful tuning for optimal results.

Result

Images are taken as sample images and filtered through median, mean, and Gaussian filters and noise free image is get.

2.3 AN IMPLEMENTATION OF K-MEANS CLUSTERING FOR EFFICIENT IMAGE SEGMENTATION

- Proposed by, Dr. K. Venkatachalam, Dr. V. Purandhar Reddy, M. Amudhan, A. Raguraman, E. Mohan in 2021 at 10th IEEE International Conference on Communication Systems and Network Technologies.

Methodology

- Initialize K cluster centres randomly or using a specific initialization technique such as K-means++.
- K-means clustering algorithm is performed on the extracted features. This involves assigning each feature to the nearest cluster centre and updating the cluster centres based on the assigned features. This process is repeated until convergence, where the cluster centres no longer change significantly.
- 33 Each pixel in the image is assigned to a cluster, based on the nearest cluster centre. This results in segmenting the image into K distinct regions or clusters.

Advantages

- K-means clustering is a relatively simple and easy-to-understand algorithm. Its simplicity makes it accessible to users with varying levels of expertise in image processing and machine learning.
- K-means clustering can handle images of varying sizes without significant performance degradation. This scalability is particularly beneficial when dealing with high-resolution images or large datasets.

Disadvantages

- k-means has trouble clustering data where clusters are of varying sizes and density.
- Centroids can be dragged by outliers, or outliers might get their own cluster instead of being ignored.

Result

Images are taken as sample images to compute the K Means clustering and the same are converted to the gray scale images as the first step and the pre-processing step of the image segmentation. In that, it is clearly showing that the difference of images under various luminance conditions.

2.4 A DEEP LEARNING-BASED RECOGNITION APPROACH FOR THE CONVERSION OF MULTILINGUAL BRAILLE IMAGES

- Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh, 11543, Saudi Arabia,17 January 2021

Methodology

- Braille Image Acquisition and Pre-processing: Original Braille images are acquired (OrgImgs). Image pre-processing (Image Preprocessing) is performed, resulting in pre-processed and segmented images (PreImgs, SegImgs).
- Braille Cell Cropping: Braille cells are cropped using the Character Positioning Cropping Algorithm (CPCA). The segmented image in the Braille cells is processed to extract individual Braille symbols.
- Braille Cell Recognition: A Deep Convolutional Neural Network (DCNN) model is employed to recognize the Braille cells obtained from the previous step. The DCNN model is trained using a training set (Trainset) and corresponding labels (Train Labels).

- Braille Multilingual Mapping: Recognized Braille cell labels (TstLabel) are used to search a lookup table for the corresponding characters and symbols of the selected language. Mapped characters and symbols are concatenated to form sentences in text or voice format.

Advantages

- The approach follows a well-defined and sequential process flow, making it easy to understand and implement.
- The inclusion of image acquisition and pre-processing steps ensures that the input Braille images are appropriately prepared for further analysis.
- The use of the Character Positioning Cropping Algorithm (CPCA) for Braille cell cropping is advantageous in isolating individual Braille symbols accurately.
- The ability to map Braille symbols into characters and symbols of different languages enhances the system's versatility and usability across various linguistic contexts.

Disadvantages

- The quality of the input Braille images is crucial for accurate recognition. Poor image quality, noise, or distortions may lead to recognition errors, especially during the image acquisition and pre-processing steps.
- The accuracy of the Braille cell cropping algorithm (CPCA) may be affected by variations in Braille cell spacing, alignment, or non-standard Braille layouts.
- The performance of the proposed approach may be sensitive to variations in Braille cell characteristics, such as size, orientation, and quality. Small changes in these parameters may impact recognition accuracy.

Result

Original Braille images are acquired. Processes Braille images through a trained DCNN model. Maps the recognized labels to characters based on language and context and concatenates them to form readable text.

2.5 INTERNET OF THINGS (IoT DEVCES)

- Rajeev Kumar, Rajat2 Mani Raj, Vandana, Rohan and Ritik Raj Department of Bioscience and Bioengineering, Lovely Professional University Phagwara, Punjab. India.2Global Research Institute of Management and Technology. Radaur, Haryana, India, (May/June 2022)

Methodology

- As we know communication through internet networks can be easily facilitated. The transfer of data packets to a network saves time and money.
- The same information that can be transmitted can be done less than ever, just by internet of things.
- The more information you have, the easier it to make an appropriate decision. You have access to real-time data and information that is far away from your location.
- The authors likely established specific inclusion criteria to select relevant literature for the review, ensuring that the included sources were pertinent to the scope of the review and contributed valuable insights to the discussion of IoT devices.

Advantages

- Time Saving By programming the work, whenever needed or required will be completed and doing this will save human valuable time and energy. IoT has now proved that it has a more valuable existence in this world.
- Connectivity empowers internet of things by bringing together everyday objects.
- Processes in which machines have to work with each other become more effective and produce better, faster results.

Disadvantages

- Compatibility: As devices from different manufacturers will be interconnected in IoT, presently, there is no international standard of compatibility for tagging and monitoring equipment.
- Complexity: The IoT is a diverse and complex network. Any failure or bugs in the software will have serious consequences. even power failure can cause a lot of inconvenience. The designing, developing, and maintaining and enabling the large technology to IoT system is quite complicated.

2.6 A POSITION PAPER ON RESEARCHING BRAILLE IN THE COGNITIVE SCIENCES: DECENTRING THE SIGHTED NORM

- Robert Englebretson1, M. Cay Holbrook2 and Simon Fischer-Baum3,41 Department of Linguistics, Rice University, Houston, TX, USA, 2 Department of Educational and Counselling Psychology and Special Education, University of British Columbia, Vancouver, BC, Canada,3 Department of Psychological Sciences, Rice University, Houston, TX, USA and 4 Social, Behavioural, and Economic Sciences Directorate, National Science Foundation, Alexandria, VA, USA,15 February 2023.

Methodology

- Decentring Sight-Centric Views: The paper emphasizes the importance of decentring sight-centric views of reading and highlights the need for researchers to understand the contextual variables intrinsic to the heterogeneous population of braille users.
- Participant Recruitment: The authors stress the significance of being intentional about where and how to recruit participants, especially given the sparse and widely distributed population of braille users.
- Understanding Contextual Variables: The paper underscores the need for research methods to consider the heterogeneity of braille users from the outset.
- Responsible and Informed Research: The authors advocate for responsible and informed research about braille, particularly in terms of understanding the diverse nature of reading and writing, and the implications of print-centric positionings of braille on research work.

Advantages

- Advancing our understanding of braille as a writing system: Researching braille can help us better understand the unique characteristics of this writing system, including its tactile nature, its use of embossed dots, and its relationship to spoken language.
- Improving braille literacy: By studying braille, researchers can identify effective teaching methods and interventions that can improve braille literacy among individuals who are blind or visually impaired.

- Enhancing accessibility: Researching braille can help improve the accessibility of written materials for individuals who are blind or visually impaired, by identifying ways to make braille more widely available and easier to access.
- Promoting equity and inclusion: By centring the experiences and perspectives of braille users, research on braille can help promote equity and inclusion for individuals who are blind or visually impaired, and can help reduce the marginalization of this population.

Disadvantages

- Limited research and resources: The low incidence of blindness and the relatively small population of braille users may present challenges in recruiting participants and conducting large-scale studies, leading to limited research and resources in this area.
- Heterogeneity of braille users: The diverse nature of braille users, including differences in the onset of visual impairment, educational experiences, and exposure to braille, may present challenges in generalizing research findings and developing standardized interventions.
- Accessibility and technology: Researching braille in the cognitive sciences may also need to consider the impact of technological advancements and the availability of alternative formats for accessing written materials, which could influence the use and relevance of braille.

Result

The paper also highlights the challenges and complexities involved in researching braille, particularly in terms of participant recruitment, understanding contextual variables, and the urgent responsibility of ensuring solid and well-grounded research.

2.7 SMART READER FOR VISUALLY IMPAIRED PEOPLE USING RASPBERRY Pi

- Department of Electronics and Communication Engineering Lakireddy Balireddy College of Engineering Mylavaram, Andhra Pradesh, India. April 2019

Methodology:

- **Image Processing Module**: - Text Extraction: The system utilizes optical character recognition (OCR) to extract text from colour images. OCR is employed to recognize characters in the scanned or printed text images and convert them into editable text.

- Character Recognition Steps:

  - Image Capture: A webcam is focused on the text, capturing the image.

  - Pre-processing: The colour image is converted into grayscale and then into a binary image.

  - Character Extraction and Resizing: Characters are extracted from the image and resized.

  - Template Matching: Templates are loaded and matched for character recognition.

  - Background Removal: The background is removed from the image.

  - Edge Detection: The system performs edge detection and writes the output to a text file.

- **Voice Processing Module:** Text to Speech Conversion: The extracted text is converted into speech using a text to speech synthesizer (TTS) installed in the Raspberry Pi.

  -Audio Output: The output of the TTS is amplified using an audio amplifier and then played through a speaker.

- **System Components:** Raspberry Pi B+ Processor Board: The project is built around the Raspberry Pi B+ processor board, which controls the system.

  -Camera: A camera is used to capture the image of the text.

  -Speaker: The audio output is played through a speaker, allowing the user to listen to the converted text.

- **Working Components:**

  Power Supply: An electrical device supplies the required power, with a 5V supply used in the system.

Advantages

- The system provides an accessible way for the visually impaired to interact with computer systems. By converting text files into audio, the system enables users to access and consume information that would otherwise be inaccessible to them**.**
- The system is designed to be user-friendly, with a simple interface that allows users to capture images of text and convert them into audio files with ease. The system does not require any specialized knowledge or training to operate.

- The integration of optical character recognition and text to speech synthesizer in the Raspberry Pi platform makes this system an innovative solution for text to speech conversion. The system combines the capabilities of these two technologies to provide a powerful tool for the visually impaired community.
- The system can be customized to meet the specific needs of individual users. For example, users can adjust the speed and volume of the audio output to suit their preferences.

Disadvantages

- The accuracy of OCR (Optical Character Recognition) in extracting text from images may vary based on the quality of the captured images and the complexity of the text. Inaccuracies in OCR could lead to errors in the converted audio output.
- The system may face challenges in accurately processing complex text formats, such as handwritten text, stylized fonts, or non-standard layouts, which could impact the accuracy of the text to speech conversion.
- The effectiveness of the system is reliant on the quality of the captured images. Poor lighting conditions, blurriness, or other image quality issues may affect the accuracy of text extraction and subsequently impact the quality of the audio output.
- Users, particularly those unfamiliar with technology, may require some time to become proficient in using the system effectively, especially in capturing high-quality images and interpreting the audio output.

Result

Text to Speech Conversion using Raspberry Pi system was able to successfully detect text on the image and convert it into an audio file. The system was able to convert both capital and small letters, making it a versatile tool for text to speech conversion. The system was tested on the Raspberry Pi module, and the output was played through a speaker. Overall, the Text to Speech Conversion using Raspberry Pi system appears to be a promising solution for text to speech conversion, particularly for the visually impaired community

2.8 SMART BOOK READER FOR VISUAL IMPAIREMENT PERSON USING IoT DEVICE

- Norharyati binti Harum, Nurul Azma Zakaria, Nurul A. Emran, Zakiah Ayop, Syarulnaziah Anawar Centre for Advanced Computing (C-ACT), Faculty of Information and Communication Technology Universiti Teknikal Malaysia Melaka, Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia,2021

Methodology

- Requirements Planning Phase analyse problems that occurs among blind people when reading a book, and then determine adequate solutions that might solve the problems.
- In User Design Phase the problems that occur among visual impairment person is analysed to determine adequate solution/modules that might help them and their family in having a low cost, portable and easy to use product.
- In the cutover Phase the functionality of the system is improved based on testing in the previous stage.

Advantages

- IoT Integration: By utilizing IoT technology and a Raspberry Pi, the Smart Book Reader provides a modern and innovative solution for accessing and processing book content.
- The device supports both softcopy and hardcopy books, providing users with a wide range of reading materials.
- Through IoT connectivity protocols such as Wi-Fi and 4G services, the Smart Book Reader can access online text-to-voice converters, further expanding the range of available reading materials.
- The Smart Book Reader promotes independence by enabling visually impaired individuals to read books without relying on assistance from others.

Disadvantages

- The cost of the Raspberry Pi and other components required to build the Smart Book Reader may be prohibitive for some individuals or organizations.
- Building and operating the Smart Book Reader requires technical expertise, which may be a barrier for some users.

- The Smart Book Reader relies on technology, which may be subject to malfunctions or other technical issues that could impact its functionality.

Result

The embedded camera is used to capture the image of the book**.** The captured image is sent to Tesseract that recognizes the word in the image and converts it into text file. Then Flite will read the text file and the text file will be converted to voice and played by speaker. In this prototype,10 second is set to give time to blind people turning page.

2.9 CHARACTERIZATION OF ENGLISH BRAILLE PATTERNS USING AUTOMATED TOOLS AND RICA BASED FEATURE EXTRACTION METHODS

- Department of Computer Science and IT, University of Azad Jammu and Kashmir, Muzaffarabad 13100, Pakistan 2022

Methodology

- The dataset comprised of "x" and "y" coordinate values of the tapped dots against one alphabet stored in a ".txt file" and converted to a ".csv file" for further processing.
- Feature Extraction: RICA- and PCA-based feature extraction methods were used to extract features from the dataset.
- Support Vector Machine (SVM), Decision Trees (DT), and K-Nearest Neighbour (KNN) with RICA- and PCA-based feature extraction methods were used for Braille to English character recognition.
- A statistical test was performed to justify the significance of the results.

Advantages

- The study found that RICA-based feature extraction methods outperformed other methods in predicting the English alphabet for corresponding Braille characters, resulting in improved Braille to English character recognition.
- The study collected a new dataset of English Braille patterns directly from visually impaired students using a position-free touchscreen-based Braille input application, which is less tiring and less complicated for visually impaired individuals.
- The study's findings can be applied in the development of Braille technology, such as the creation of more user-friendly devices for visually impaired individuals.

- The study contributes to the advancement of Braille technology by exploring new methods for characterizing English Braille patterns, potentially leading to improved tools and applications for visually impaired individuals.

Disadvantages

- The study focused on Grade 1 English Braille patterns, and there was a mention of the need to increase the number of datasets not only for Grade 1 but also for Grade 2 and contracted Braille. This limitation may impact the generalizability of the findings to a broader range of Braille patterns.
- The study mentioned that the proposed model's results were not satisfactory when implemented on mobile devices with limited computation power. This limitation suggests the need for further optimization of the models for mobile device usage.
- The study primarily focused on Braille to English character recognition and did not extensively explore other potential applications or functionalities for visually impaired individuals.

Result

The dataset was split into training and testing sets, with a 70:30 ratio for training and 80:20 for testing. Using PCA, SVM, KNN, and DT achieved accuracies of 86.32%, 75.40%, and 70.02%, respectively. However, better results were obtained using RICA, with SVM achieving an accuracy of 99.85%, and KNN and DT achieving accuracies of 99.50% and 99.79%, respectively.

2.10 A NEW MODEL FOR IMAGE SEGMENTATION BASED ON DEEP LEARNING

- Rafeek Mamdouh BADR University, Cairo, Egypt Mansoura University, Mansoura, Egypt.

Methodology

- This paper describes a combination of two fields of solving segmentation problem to convert through the workflow of a hybrid algorithm structure Convolutional neural network (CNN)
- The segmentation model for images can be expanded to colour segmentation by using "RBG" or their conversion (linear/non-linear).

- We propose a process to implement DCNN image segmentation, and use active contour technique as a method for segmentation as input data where use CNN network of deep learning.

Advantages

- The use of deep learning techniques allows for efficient processing of large DICOM image datasets, enabling faster segmentation and analysis.
- The CNN component of the model can adapt and learn from the complex features present in medical images, making it suitable for a wide range of medical imaging applications.
- The automated nature of the model reduces the need for manual intervention in the segmentation process, saving time and effort for medical professionals.
- The accurate segmentation of medical images has potential applications in medical diagnosis, treatment planning, and surgical training tools.

Disadvantages

- Deep learning models require large amounts of data to train effectively. This can be a challenge in medical imaging, where obtaining large datasets can be difficult due to privacy concerns and limited availability.
- Training deep learning models can be computationally intensive, requiring high-performance computing resources. This can be a barrier for smaller medical institutions or those with limited resources.
- The proposed model involves a combination of ACM and CNN, which can make it more complex and difficult to implement compared to traditional segmentation methods.

Result

The proposed model achieved high accuracy in segmenting DICOM images, particularly for liver segmentation. The DSC values for liver segmentation ranged from 0.92 to 0.96, indicating a high degree of overlap between the ground truth and the segmented images. Proposed model also demonstrates its potential as a promising solution for accurate and efficient segmentation of medical images, with potential implications for improving medical diagnostics and treatments.

2.11 DEEP LEARNING STRATEGY FOR BRAILLE CHARACTER RECOGNITION

- Tasleem Kausar, Muhammad Sajjad, Adeeba Kausar, Yun Lu, Muhammad Wasif and M. Adnan Ashraf ,2021

Methodology

- Image Preprocessing is done for image alignment and enhancement are performed using various image preprocessing techniques. This involves a series of steps such as image filtering, histogram equalization, chromatic adaptation, and morphological operations.
- The CNN model is based on a lightweight architecture using an inverted residual block (IRB) module to reduce computational cost for the classification if the image. The proposed network is designed to effectively recognize Braille characters and outperforms existing methods in terms of prediction accuracy and recognition speed.
- They evaluated the performance of the proposed model using statistical measures such as sensitivity, precision, F score, predictive accuracy, entropy, and peak signal-to-noise ratio (PSNR).
- The proposed method was tested on two publicly available Braille datasets, namely the English Braille character dataset and the double-sided Braille image dataset.

Advantages

- The proposed approach demonstrates enhanced recognition accuracy for Braille characters, achieving a prediction accuracy of 95.2% and 98.3% on the English Braille and double-sided Braille image datasets, respectively.
- The methodology incorporates various image preprocessing techniques to enhance image clarity and alignment, thereby improving the overall recognition process.
- The authors conducted thorough experimental evaluations, including the impact of different preprocessing techniques and backbone networks on recognition performance. The use of statistical measures and evaluation on publicly available datasets adds rigor to the study.
- The strategy addresses the challenge of network computational cost, making it suitable for real-time applications, which is crucial for assistive technologies for individuals with visual impairments.

Disadvantages

- Deep learning models, including CNNs, are often considered as "black box" models, meaning that the decision-making process of the model may not be easily interpretable, which could be a concern in applications where interpretability is crucial.
- The model's robustness to noise and variability in real-world Braille images, such as smudging, varying lighting conditions, and different writing styles, may need further investigation to ensure reliable performance in practical applications.
- The strategy aims to reduce computational cost, deep learning models, especially when deployed in real-time systems, may still require significant computational resources, which could be a limitation in certain environments.

Result

Achieved a prediction accuracy of 95.2% and 98.3% on the English Braille and double-sided Braille image datasets, respectively. In terms of time complexity, the proposed model involves fewer time computations compared to state-of-the-art methods, performing Braille character recognition quickly. The paper suggests potential future extensions, including the application of the algorithm to larger Braille datasets and the exploration of the impact of Braille image qualities on recognition systems.

2.12 BRAILLIE TO TEXT/VOICE CONVERTER

- Bayan Halawani, Wisam Younes, Samer Isieed Hebron-Palestine, May 2020

Methodology

- To design and implement a system that will be able to convert the Braille language scripts into multilingual scripts, using effective algorithms and techniques.
- The image processing elements, including image acquisition, which is the first and most important step in any pattern recognition system.
- The taken images naturally has noise that will make the image unclear and will cause some errors in the subsequent used algorithms, there are some noise filtering algorithms including mean filter, median filter, Gauss filter.
- The RGB image can be represented in one-dimensional array, because it will make the processing easier, faster and will reduce the amount of information in an image.

Advantages

- Image processing techniques are used in image compression to reduce the storage space required for images while maintaining an acceptable level of visual quality. This is important for efficient storage and transmission of images over networks.
- The user rolls a sensor over the Braille script, which will convert the Braille cells into signals, then the sensor sends these signals to the build-in microcontroller to analyse them, and find the corresponding alphabet character and then output this result as voice using speaker
- The most important stage in this system, which includes all operations needed to analyse the image and convert each cell to its equivalent binary code then using hash table to convert it from binary to Character.
- Braille Text/Voice Converter is a good algorithm, because there is no need to go through all pixel in the image, which means it is fast algorithm, and we were unable to compare our project with the other due to non-existence of such application that convert a Braille script into another script using the mobile.

Result

The existence of such a system to convert a Braille paper into another language will help those sighted people they are dealing with blind people in their daily life to be able to read a Braille paper without have any knowledge in Braille, because it is not easy to learn a Braille language. The existence of such a system to convert a Braille paper into another language will help those sighted people they are dealing with blind people in their daily life to be able to read a Braille paper without have any knowledge in Braille, because it is not easy to learn a Braille language.

CHAPTER 3

SYSTEM DESIGN AND IMPLEMENTATION

The chapter contains detailed description about the function being carried out in each module. This also includes some information about OBR.

3.1 System Architecture

This chapter includes the information about the system architecture.

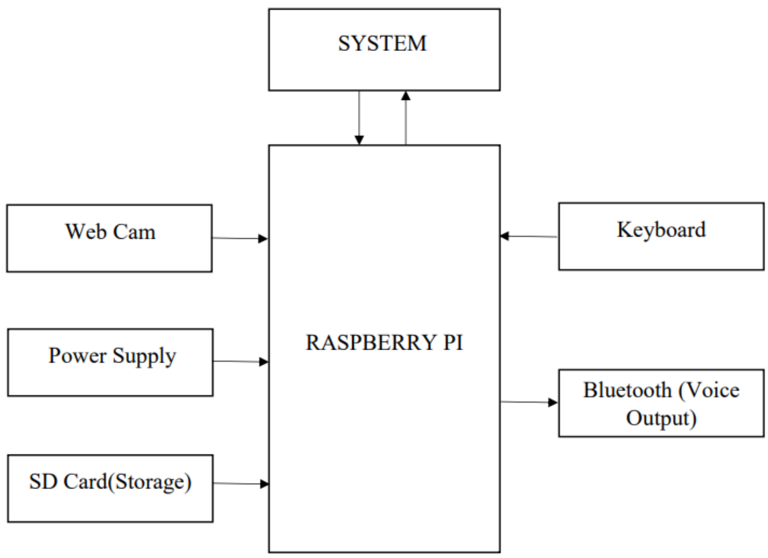

Fig.3.1 System Architecture

The fig 3.1 shows the different components used in our system architecture.

Power Supply: This block provides the essential electrical current to the Raspberry Pi, typically at 5 volts and a minimum of 2.5 amps.

Raspberry Pi: The heart of the system, this block is a single-board computer (SBC) containing a Central Processing Unit (CPU) – usually an ARM processor – and other integrated circuits like a Video Core Graphics Processing Unit (GPU) for handling visuals. The CPU executes instructions and performs calculations.

SD Card (Storage): This removable storage device acts as the Raspberry Pi's main memory. It stores the operating system (OS) like Raspbian, applications, and your personal data. The SD card is crucial for booting up the Raspberry Pi and running programs.

Webcam: This optional block adds camera functionality to your Raspberry Pi. You can connect it via a USB port to capture images or videos for various applications, like video conferencing or creating time-lapse videos.

Keyboard: An essential input device, the keyboard allows you to enter text commands and interact with the Raspberry Pi's operating system. It connects via a USB port and facilitates communication with the computer.

Bluetooth (Voice Output): While the Raspberry Pi may have a built-in headphone jack for audio output, Bluetooth provides a wireless alternative. By connecting a Bluetooth speaker or headset, you can enable voice output for functions like listening to music or receiving audio notifications.

3.2 Workflow

This chapter includes the information about the workflow of the model

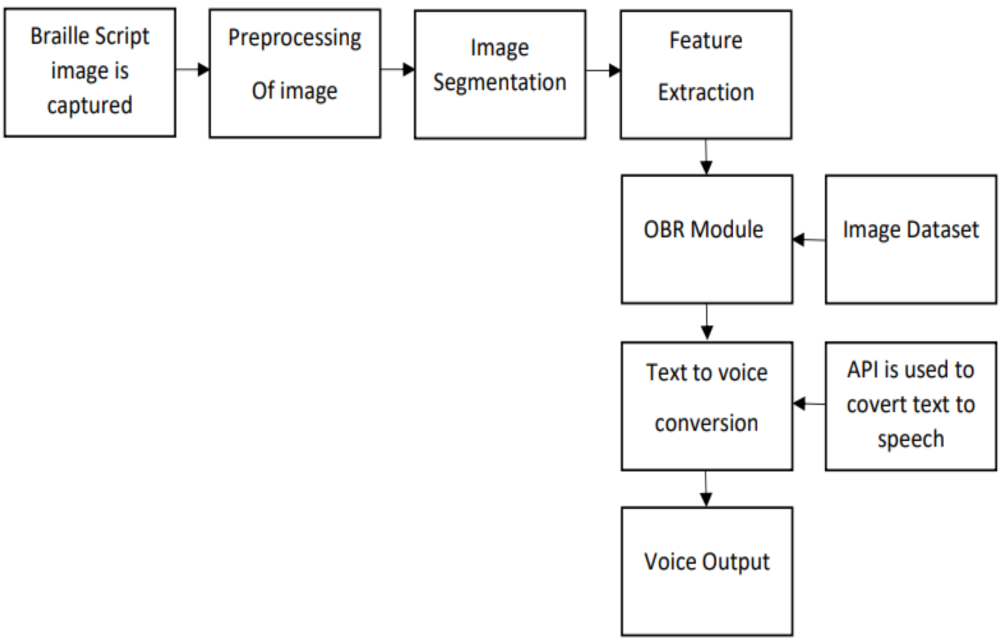

Fig.3.2 System Workflow

The fig 3.2 shows the different model in our system workflow.

At first the featured extracted image dataset is sent to the OBR module to recognize the braille text then image that must be classified is taken from the camera module.

Pre-Processing of Images: Extracted images form camera contains out of interest substance which should be cropped from the point of interest that is the blind script image is extracted here and the image is converted to the required resolution**.**

Image Segmentation: Here the image taken from previous module is segmented using K-Means clustering technique. This step helps to distinguish the blind script with other substances in the image.

Feature extraction: It is the process of selecting and transforming relevant information from raw data to create a simplified representation, highlighting essential patterns for further analysis or machine learning tasks

OBR module: Optical Braille Recognition (OBR) is a Windows software program that allows you to 'read' single- and double-sided Braille documents on a standard flatbed scanner. It scans the Braille document, analyses the dot pattern, and translates it into normal text that it presents on the computer screen.

Text to Voice Conversion: The classified text is converted to voice using an API.

The workflow provided outlines the process flow of a project aimed at converting braille to voice output flow using IoT (Internet of Things) and machine learning technologies. At its core, the project seeks to leverage data from a camera feed to convert braille script and then utilize this information to convert into voice output. Here is a detailed description of the workflow:

At the outset, the system declares its input and output ports, setting the stage for data exchange with external devices and systems. This initialization phase ensures that the system is ready to send and receive data as required.

The image of the objects in front of the screen can be captured through an internet connected via USB to the microcontroller when the capture button is pressed. The captured image is subject to OCR technology once the compression process has been determined. OCR technique allows the conversion to text or information that can be incorrectly perceived or modified in scanned images of text or symbols. However, OCR technology is part of a wider structure that requires identification capabilities such as quantitative plate recognition systems or many other applications such as tools that create opportunities for SALT development from print-based texts. We abuse the TESSERACT library on the OCR technology chart. Information from the abuse flight library is re-created in the form of audio files. After that the picture or painting of the goods logo is internally processed and the mark separated from the image by abusing the open CV library and eventually identifying the object. The image received is now recognised and converted with the Tesseract library into text. After the famous La-Bell name is changed to an icon, the changed text appears on the console display. Currently, when using the flight library, text- to-speech can be used to hear the naming name as audio from earphones connected to the audio jack port.

3.3 Tools Used

3.3.1 Hardware Requirements:

- Web camera
- Raspberry Pi
- Bluetooth

3.3.2 Software Requirements:

- Python
- Open CV
- Flask

WEBCAM

A digital camera is a video camera that sends images on a laptop or a network in real time, generally via USB, LAN, or Wi-Fi, as shown in fig 3.3. Creating video connections, allowing computers to function as video cable or video conferencing stations. This is the normal use of land as a video camera. The digital camera got its name from the internet. Other common applications include security police investigation and laptop visibility. Webcams are known for their low production costs and compatibility, making them a low-cost video calling option. Some government webcams can be controlled remotely by spyware, which can be a source of security and privacy issues.

LENSES: Mirror, photo machine, auxiliary physics and audio microphone are all common features of webcams. A variety of lenses are available, and the 29 most popular among consumer webcams are plastic lenses, which are stabilized to focus the image. Open are fixed focus lenses that have no adjustment. Since the depth of the camera device's field is great for large focal length (small aperture) narrow image size and lenses, the devices used in webcams have relatively wide depths. The lens has not much effect on the sharpness of the image by using a strong and quick focus. Image sensors, formerly most popular than low-cost cameras, are CMOS or CCD. In this price range, CCD cameras are often higher than CMOS. VGA quality footage is available at most customer webcams at a frame rate of thirty FPS. Most modern cameras generate multi-megapixel resolution video, while some play at fast frame rates such as the PlayStation Eye, capable of generating 320 x 240 video with 120 frames a second. We avoid Due to space and money, the computer and the physics that support it are built on a single silico chip. Convenience or convenience the video and video conferencing are made much easier for most webcams. Users are permitted to connect to computers without using device drivers by means of USB Video Device Class (Uvc) standards. UVC support is integrated into Microsoft Windows XP SP2, UNIX and Mac OS X (since October 2005) and does not require external system drivers, but is usually installed to include more options. Features (Logitech C100 Digital Camera) Plug and Play (UVC) Configuration Video Capture: up to 640 x 480 pixels Photo: up to one. (With the proposed system) High Speed Fixed Focus USB 2.0 Certified Notebooks, alphanumeric displays, and cathode ray tubes are all compatible with this universal clip-ray tube monitor

Fig.3.3 Webcam

RASPBERRY Pi

Raspberry Pi is a series of small, affordable single-board computers developed by the Raspberry Pi Foundation, primarily aimed at promoting computer science education and facilitating DIY projects. Its compact size, low cost, and versatility make it popular for various applications, including IoT (Internet of Things), robotics, home automation, and educational purposes. Raspberry Pi can handle tasks such as natural language processing (NLP) for converting text to speech or Braille output and managing user interactions through tactile interfaces or voice commands. Raspberry Pi provides an ideal platform for developing assistive technologies like the smart book reader. Its capabilities can be extended with various peripherals and accessories, allowing for customization and scalability according to specific user needs. Overall, Raspberry Pi plays a crucial role in enabling innovative solutions for enhancing accessibility and inclusivity for individuals with visual impairments.

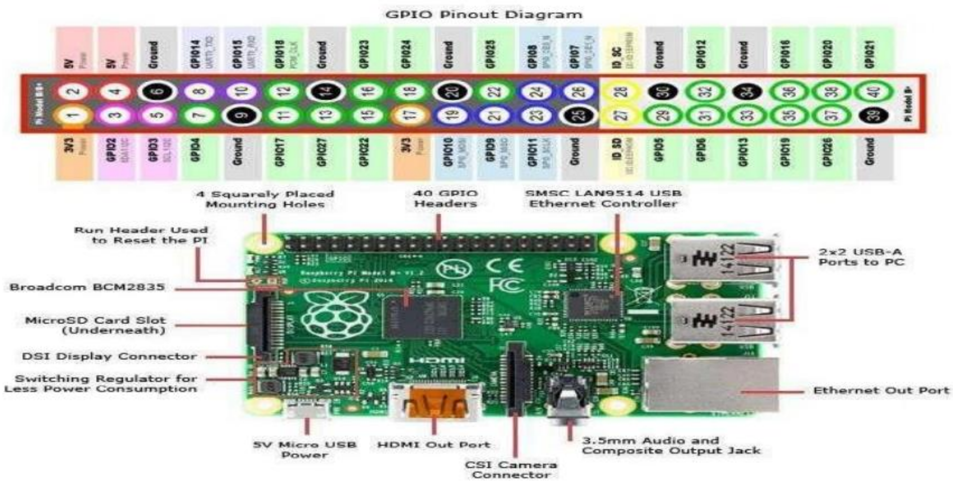

Fig.3.4 Raspberry Pi Model

Exploring the board on raspberry Pi

Model A and Model B are the two available versions of the Raspberry Pi. The Model B has a few more advanced options than the Model A. The Model B has 2GB RAM and two USB ports, while the Model A has only 256 MB RAM and one USB port. Also, Model B has a LAN port and Model A does not have a LAN port. The different elements of the upper area unit are classified in Figure 3.5 and each of the following sections contains a variable definition of the fraction.

SD CARD SLOT: Raspberry Pi does not have a phone-like drive, the SD card is used for the installation of software while storing it. This card must be inserted into the wound site of the raspberry pie. A 2GB, 4GB or 16GB SD card is also available.

USB POWER MICRO: Since it does not have an onboard voltage regulator, the power port 5V must be a micro- USB input and the supply may be specifically 5V. As a result, the power supply does not reach 5 volts.

HDMI OUTPUT: This HDMI output port is used to connect to the display (high-definition transmission interface) on the raspberry. As a result, every monitor or TV can be connected to an HDMI port.

Ethernet and USB port: The LAN9512 chip is equipped with an Ethernet and a USB port in Model B. It has a 10/100 LAN interface with high-speed USB 2.0 hub (DONAT, 2014). USB ports for connecting input devices (keyboard, mouse). Nearly all connections via USB to the computer can also be connected to the top of the raspberry.

RCA AUDIO OUT: Recording Corporation of America stands for RCA. Audio in foreign countries RCA audio jacks is available on the ground for audio output. The Raspberry Pi can support the audio from its HDMI output, although the USM microphones do not support or support it is also a standard 3.5mm audio jack to connect headphones and Bluetooth.

SUBTITLES GPIO (PINS): 24 Input / Output pins for a variety of purposes are the term "GPIO pins." These pins are used to attach the tops of the frame to different physical extensions. Comes with pre-installed libraries on Raspberry, allowing North American users to access PINs using programming languages such as C, C ++ or Python.

PYTHON

Python is widely used in the field of Object Recognition (OBR) due to its versatility, ease of use, and robust ecosystem of libraries. OBR, a crucial component of computer vision, relies on algorithms to identify and classify objects within digital images or video streams. Python's libraries such as OpenCV, TensorFlow, and Tesseract offer powerful tools for image processing, feature extraction, and neural network implementation, making it ideal for developing OBR systems. With Python, developers can quickly prototype and deploy OBR solutions, whether for autonomous vehicles, surveillance systems, medical imaging, or industrial automation. Its flexibility also enables integration with other technologies, facilitating the creation of sophisticated OBR applications tailored to specific needs.

Python libraries:

The code utilized several libraries for various tasks. Here is a breakdown of the libraries used:

1. PyTesseract is a Python library for integrating Tesseract OCR Engine, enabling text extraction from images or documents. It simplifies OCR tasks with its straightforward interface, making it popular for automated data extraction and document processing.
2. PTTS, or Python Text-to-Speech, is a library that converts text into spoken audio using various synthesis techniques. It offers easy integration of text-to-speech functionality into Python applications, facilitating accessibility features, voice interfaces, and audio content generation.

OpenCV

OpenCV (Open-Source Computer Vision Library) is a versatile tool for image and video processing, offering functionalities like object detection, facial recognition, and machine learning support. Its extensive set of libraries and algorithms make it suitable for various applications, from augmented reality to robotics. OpenCV's Python and C++ interfaces provide flexibility and efficiency for developers working on computer vision projects. Its active community and documentation resources further enhance its usability for both beginners and experienced practitioners.

Flask

Flask is a popular Python web framework that allows developers to build web applications quickly and easily. It is classified as a micro web framework because it paksa minimal set of tools and features needed to build a web application. Flask is known for its simplicity and ease of use. It is used for connecting python and frontend.

Some key features and concepts associated with Flask include:

• **Routing:** Flask provides a mechanism for defining URL, routes and mapping them 10 Python functions or methods. This allows developers to create custom endpoints for handling requests and returning responses.

• **Templates:** Flask uses Jinja2 as its default templating engine, which provides a flexible and powerful way to generate HTML and other markup languages. Templates can be easily customized and extended to create dynamic web pages. ENes In Flask, a view is a Python function or method that handles a specific URL route. Views can be decorated with various decorators to add functionality such as authentication, caching, or error handling multiple sessions. Flask allows developers to store user-specific data across requests using sessions. Sessions are stored as encrypted cookies by default and can be customized to use other storage options.

• **Extensions:** Flask provides a range of extensions that add additional functionality to the framework, such as database integration, form handling, and authentication. These extensions are designed to work seamlessly with Flask and can be easily integrated into a Flask application

CHAPTER 4

RESULTS AND SNAPSHOTS

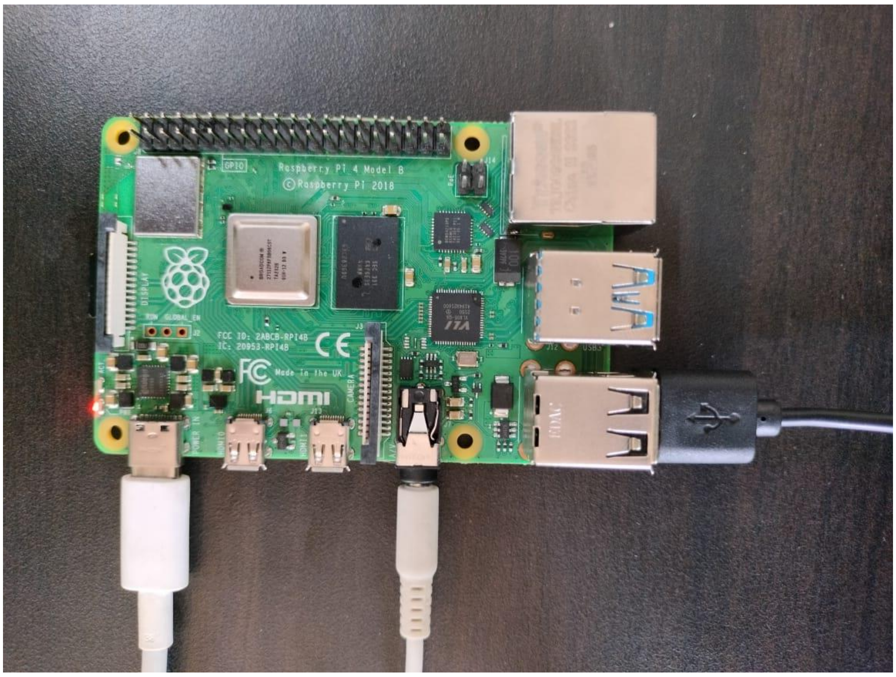

Fig.4.1 Top view of model

Fig.4.1 displays the top view of the model, where we used Raspberry Pi Model B.

Fig.4.2 Hardware setup for smart book reader

Fig.4.2 displays the hardware setup of the smart book reader.

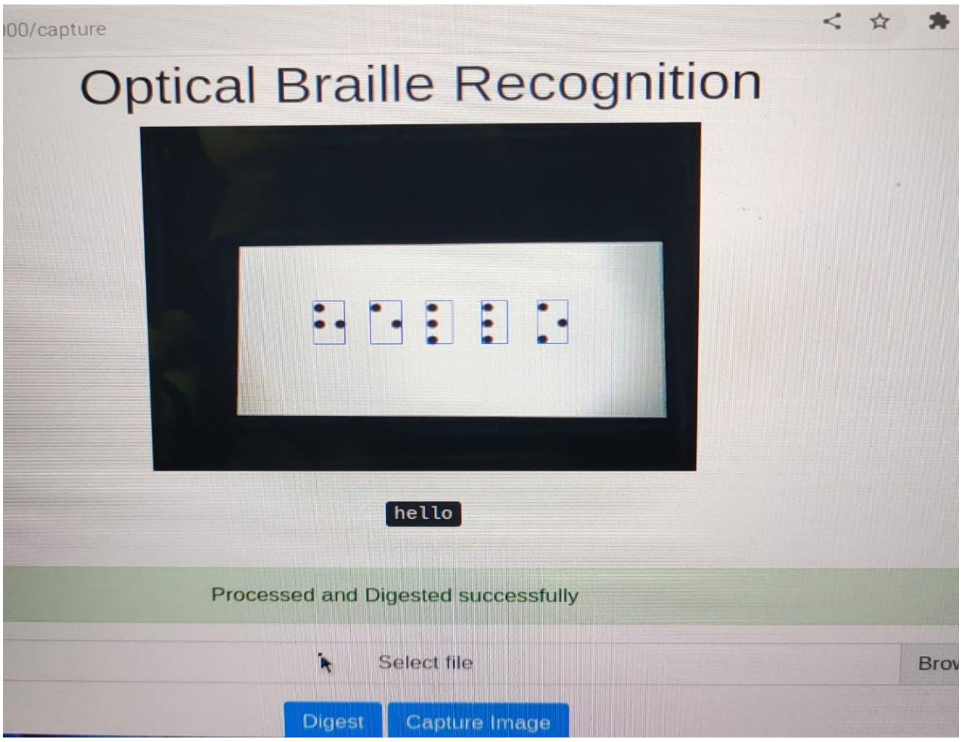

Fig.4.3 Image captured in real time and output is generated

Fig.4.3 displays the output of the captured image and the braille text is converted to voice output.

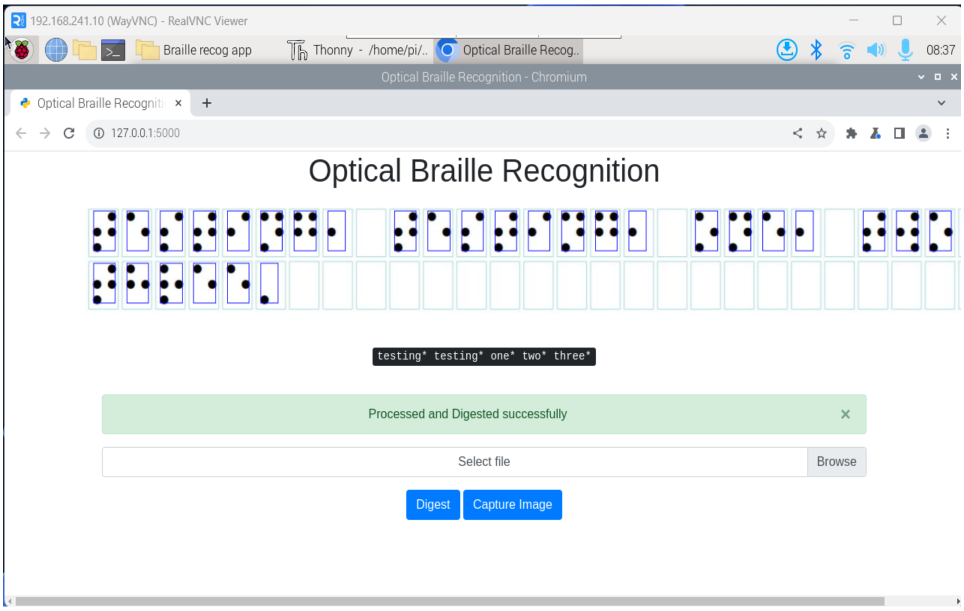

Fig.4.4 The generated output for the uploaded image

Fig.4.4 displays the output of the uploaded image and the braille text is converted to voice output.

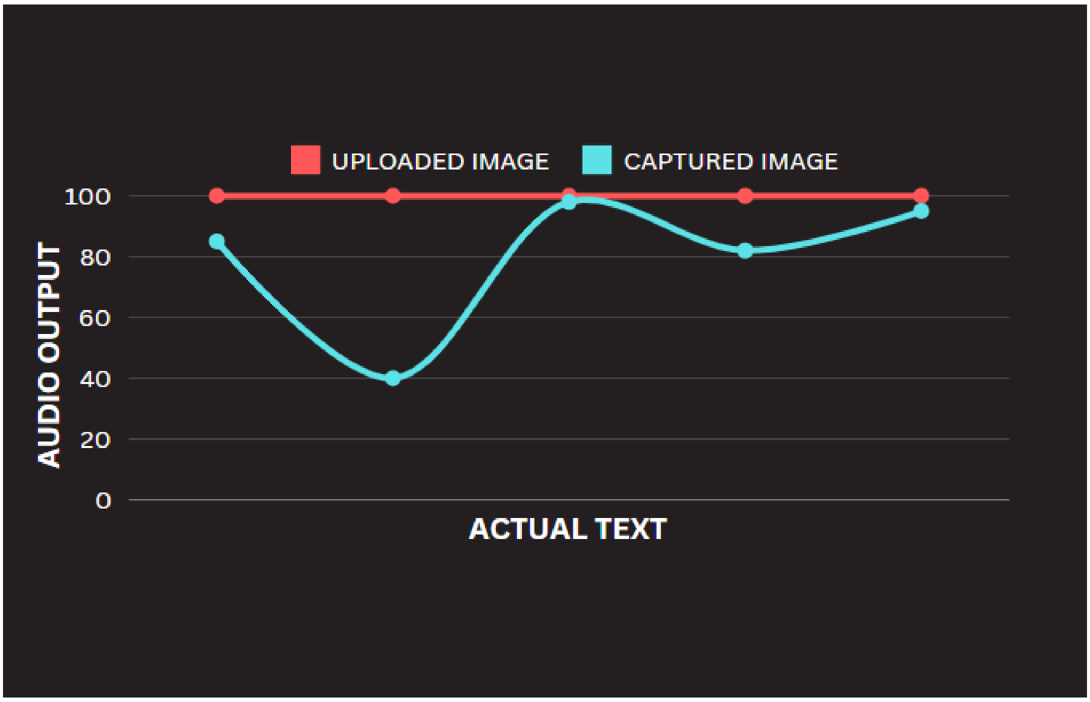

Fig.4.5 Accuracy graph

Fig.4.5 displays the accuracy graph, where it is plotted between actual text vs audio output. The graph shows the accuracy for both uploaded and captured braille text.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 Conclusion

The Smart Book Reader stands as a revolutionary advancement, adeptly addressing the multifaceted challenges faced by visually impaired individuals. By obviating the requirement to learn braille, it democratizes access to literature, offering a fast and efficient means of reading. This device not only removes the barriers to traditional braille learning but also promotes equal opportunities for visually impaired readers in comparison to their sighted counterparts. Furthermore, the Smart Book Reader's capability to convert English into audio through intricate image processing and deep learning represents a pivotal breakthrough. The precision achieved in text-to-audio conversion through OBR module ensures accurate and reliable results. In essence, the Smart Book Reader emerges as a beacon of inclusivity, breaking down longstanding barriers and empowering the visually impaired to navigate the world of knowledge with unprecedented ease and independence.

5.2 Future Scope

- The proposed system can be further improved to efficiently work with multiple languages.
- Advancements in hardware equipment’s are expected to significantly increase the accuracy of smart book readers in real-time, enhancing their usability for visually impaired individuals.

PUBLICATION

[1] Narendra Kumar S, Aishwarya S, Chandana C S, Nikhil M, Ranjan P N, “Smart Book Reader for Visual Impairment People Using IoT Device and Deep Learning”, International Journal of Scientific Research in Engineering and Management (IJSREM), Volume: 08 Issue: 05, May – 2024, DOI: 10.55041/IJSREM34565

BIBLIOGRAPHY

[1] Bi-Min Hsu "Braille Recognition for reducing asymmetric communication between the blind", 30 June 2020.

[2] Vishwanath Venkatesh Murthy "Pre-Processing the braille image for improving optical Braille recognition performance", July-August 2020.

[3] Dr. K. Venkatachalam, Dr. V. Purandhar Reddy, M. Amudhan, A. Raguraman, E. Mohan "An Implementation of K-means clustering for efficient image segmentation ,2021.

[4] Abdulmalik AlSalman, Abdu Gumaei, Amani AlSalman and Suheer Al-Hadhrami "A deep learning-based recognition approach for the conversion of multilingual braille images",17 January 2021.

[5] Rajeev Kumar, Rajat2 Mani Raj, Vandana, Rohan and Ritik Raj "Internet of Things" (IoT DEVCES), May/June 2022

[6] Robert Englebretson1, M. Cay Holbrook2 and Simon Fischer-Baum3,41 "A Position paper on researching braille in the cognitive sciences: Decentering the Sighted norm", February 2023.

[7] Vinaya Phutak, Richa Kamble, Sharmila Gore, Minal Alave, R.R. Kulkarni "Smart Reader for Visually Impaired people using raspberry Pi", April 2019

[8] Norharyati binti Harum, Nurul Azma Zakaria, Nurul A. Emran, Zakiah Ayop, Syarulnaziah Anawar "Smart Book Reader for Visually Impaired people using IoT device",2022.

[9] Sana Shokat, Rabia Riaz, Sanam Shahla Rizvi, Inayat Khan, and Anand Paul

"Characterization of English Braille patterns using automated tools and RICA based feature extraction methods,"2022.

[10] Rafeek Mamdouh, Khaled Amer, Nashaat El-Khamisy, "A New Model for Image Segmentation based on Deep Learning", July 2021.

[11] Tasleem Kausar, Muhammad Sajjad, Adeeba Kausar, Yun Lu, Muhammad Wasif and M. Adnan Ashraf "Deep Learning Strategy for braille character recognition,"2021.

[12] Bayan Halawani, Wisam Younes, Samer Isieed, "Braille to text/voice Converter", May 2020.

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 05 | May - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

Smart Book Reader for Visual Impairment Person Using IoT Device and Deep Learning

Narendra Kumar S[1], Aishwarya S[2], Chandana C S[2], Nikhil M[2], Ranjan P N[2]

** 1Assistant Professor, Dept. of CS&E, JNNCE, Shivamogga, India.*

** 2UG Students, Dept. of CS&E, JNNCE, Shivamogga, India.*

---***---

Abstract - People with normal vision can easily see the world around them and can read and write without difficulty. For those who are visually impaired, the Braille script enables them to read and write just like sighted individuals. According to WHO data from 2023, about 15 million people worldwide have significant vision loss. The Braille system uses cells with six raised dots, each dot numbered from one to six, arranged in two columns. This system is crucial for visually impaired individuals to keep up with the world around them. Providing Braille-assisted technology and incorporating it into daily life is essential to make life more comfortable and efficient for visually impaired people, enabling better communication with others.

Key Words: Braille script, OBR.

1.INTRODUCTION

Vision is a profound sense, but not everyone enjoys its full benefits. Visually impaired individuals, whether blind or with limited sight, face unique challenges in accessing written information. Braille, a tactile system of raised dots representing letters and numbers, has been revolutionary in enabling blind individuals to read and write independently by feeling these dots with their fingertips. However, mastering Braille can be daunting due to its complexity and the time required for proficiency.

To overcome these challenges, modern technology has greatly enhanced accessibility and educational opportunities for the visually impaired. The integration of IoT devices and deep learning algorithms has led to significant advancements in assistive technologies. For instance, the Smart Book Reader utilizes IoT sensors and deep learning to convert Braille text into spoken language in real-time, facilitating easier comprehension and promoting greater independence in learning.

Deep learning, inspired by neural networks, has revolutionized Braille recognition by automating the interpretation of tactile patterns, thereby improving the accuracy and speed of converting Braille into audible content. This technological convergence holds promise in enhancing the lives of visually impaired individuals by breaking down barriers to information access and fostering greater participation in society.

2. RELATED WORK:

In this section, various authors have presented various IoT and Deep Learning techniques.

This section discusses various IoT and Deep Learning techniques used by different authors. One innovative method converts Braille images to English text using a convolutional neural network (CNN) model. Optical Braille recognition (OBR) systems have been developed to translate Braille images into text, addressing the unique characteristics of Braille symbols compared to standard characters. The process begins with image binarization, converting a color image into a binary format by calculating the brightness from the RGB values of each pixel [1].

A K-means clustering algorithm is then applied to the extracted features. This involves assigning features to the nearest cluster center and updating the centers based on these assignments until they stabilize. Each pixel is classified into a cluster, segmenting the image into distinct regions [3].

The workflow includes Braille image acquisition and pre-processing, where original Braille images are collected (OrgImgs) and processed into preprocessed and segmented images (PreImgs, SegImgs). Braille cells are then cropped using the Character Positioning Cropping Algorithm (CPCA) to isolate individual Braille symbols. These symbols are recognized by a Deep Convolutional Neural Network (DCNN) model trained with a specific dataset (Trainset) and labels (Train Labels). The recognized symbols (TstLabel) are mapped to corresponding characters and symbols of the selected language using a lookup table, forming sentences in text or voice format [7].

The paper emphasizes moving beyond sight-centric views of reading, highlighting the need to consider the diverse contextual factors affecting Braille users. It stresses the importance of intentional participant recruitment, considering the sparse and distributed population of Braille users. Researchers are encouraged to account for the heterogeneity of Braille users and advocate for responsible, informed research that recognizes the diversity of reading and writing experiences, avoiding print-centric biases [10].

In the image processing module, text extraction is achieved using optical character recognition (OCR) to convert text in color images into editable text. The process includes image capture via webcam, pre-processing to convert images to grayscale and binary formats, character extraction and resizing, template matching for recognition, background removal, and edge detection, with results written to a text file. The voice processing module converts extracted text to speech