Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.98549397

OPTICAL CHARACTER RECOGNITION FOR THE VISUALLY IMPAIRED USING RASPBERRY PI

Dr.S.Murugesan
Associate Professor, RMD Engineering college, Kavaraipettai, Tiruvallur District,
smn.cse@rmd.ac.in

Dr.N.Ram Shankar
Associate Professor, RMD Engineering college, Kavaraipettai, Tiruvallur District,
nrs.cse@rmd.ac.in

Sivapriya N
UG Student, RMD Engineering college, Kavaraipettai, Tiruvallur District,
ucs19408@rmd.ac.in

Subbalakshmi G
UG Student, RMD Engineering college, Kavaraipettai, Tiruvallur District,
ucs19416@rmd.ac.in

Vinitha R
UG Student, RMD Engineering college, Kavaraipettai, Tiruvallur District,
ucs19436@rmd.ac.in

Abstract— The use of a "third eye" for blind OCR (Optical Character Recognition) application has been proposed as a potential solution for individuals with visual impairments to access printed text and to detect objects. This article describes an OCR system for the visually impaired that makes use of a Raspberry Pi and earbuds for audio output. The device converts printed text into digital format using OCR technology, providing an accessible way for the visually impaired to access printed information. The device is designed to be portable and user-friendly, with tactile feedback and audio prompts to guide the user through the scanning process. Our tests showed that the device was highly effective, with accurate OCR conversion and clear audio output. The proposed device represents a significant advancement in assistive technology for the visually impaired, providing an innovative solution for accessing printed information.

Keywords— ocr application, portable device, audio output, blind, support, object detection, fastest scanning device, raspberry pi, text reader, ease and efficiency.

I. INTRODUCTION

Visual impairment affects millions of people around the world, making it difficult for them to access printed information. While there are various technologies available to help individuals with visual impairments, such as screen readers and braille displays, there is still a need for more innovative solutions. One such solution that has been proposed is the use of a "third eye"

for blind OCR (Optical Character Recognition) application. The concept of a third eye involves the use of a wearable device that captures images of printed text and converts them into a format that can be read by an OCR software to give an audio output to the user. This technology has the potential to provide individuals with visual impairments with greater independence and access to printed materials. However, the implementation of a third eye device for blind OCR application is not without its challenges. It requires high-quality imaging and advanced image processing algorithms to ensure accurate recognition of text. The device also needs to be comfortable and convenient for the user to wear and use. This journal aims to explore the concept of a third eye for blind OCR application in detail, including the current state of OCR technology, the challenges associated with developing a third eye device, and the potential benefits for individuals with visual impairments. We hope that this journal will provide valuable insights and inspire further research and development in this field.

II. EASE OF USE

Ease of use is a crucial factor to consider when designing a third eye device for blind OCR application. The device should be user-friendly and intuitive, allowing individuals with visual impairments to use it with ease. In this journal, we will focus on two key components of a third eye device: object detection and text reader. Object detection is a critical feature of a third eye device as it allows the user to navigate their surroundings safely. The device should be able to detect and recognize obstacles, such as walls, furniture, and other people, and alert the user accordingly. The object detection feature should be easy to activate and adjust, depending on the user's preferences and needs. The text reader is another essential component of a third eye device. The device should be able to capture images of printed text and convert them into audio output that the user can listen to. The text reader should be accurate and reliable, and it should be able to recognize various font types and sizes. Additionally, the text reader should be easy to use, allowing the user to capture and read text with a simple gesture or voice command. In summary, ease of use is a crucial factor when designing a third eye device for blind OCR application. The device should have intuitive interfaces, easy-to-activate features, and be lightweight and comfortable to wear. By prioritizing ease of use, we can ensure that the third eye device is accessible and empowering for individuals with visual impairments.

III. LITERTURE SURVEY

[1] Optical character recognition (OCR) is a significant branch of machine vision. Pattern recognition, image processing, digital signal processing, artificial intelligence, and numerous other fields are all involved. It is all-inclusive. It has significant practical and theoretical importance in high-tech domains such as word processing, office automation, machine translation, and real-time monitoring systems. The deep learning technique can automatically identify and learn the hidden feature rules in a large number of data sets through iterative training, resulting in superior standardization than classic image processing related algorithms. [2] Over the past few years, neural network-based approaches have gradually acquired performance and acceptability for a wide range of optical character identification tasks, from single digit detection to handprint recognition. We provide a NN classification technique based on enhanced multilayer perception and explore the National Institute of Standards and Technology Visual Image Processing Group's end-to-end solution for form-based handprint

OCR applications. The enhancements are based on neuron activation functions, which reduce the frequency of singular Jacobians; successive regularization, which limits the volume of the weight space; and Boltzmann pruning, which limits the dimension of the weight space. The NIST form-based handprint recognition system is also provided, as with performance characterization studies of systems assessed at the inaugural OCR systems conference.

LIPNet is a lateral inhibition pattern recognition pyramidal neural network inspired by the concept of receptive and inhibitory fields in the human visual system. Although this network can extract features implicitly and utilise them to accurately classify patterns in photographs, various parameters must be adjusted before the network can be trained and used.

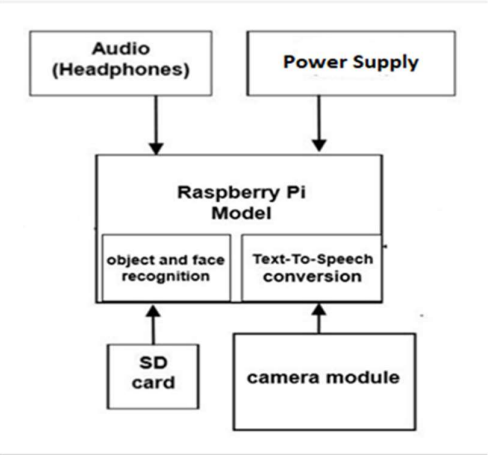
These features also have a substantial impact on recognition performance. This paper [3] describes an encoding approach that optimises the LIPNet structure using particle swarm optimisation. In preliminary results for a face identification task using a well-known benchmark set, our method exceeded the original LIPNet in classification rates.

The electronic translation of handwritten, typewritten, or printed text into machine translated pictures is known as optical character recognition (OCR). It is commonly used to recognise and search text in electronic documents, as well as to publish text on a website. The study [4] includes an overview of OCR applications in several domains, as well as experiments for three significant applications: Captcha, Institutional Repository, and Optical Music Character Recognition.
A comparison of commercial OCR software for the blind is presented in [5]. The researchers compared the accuracy, speed, and usability of three OCR software programmes: Kurzweil 1000, Abbyy FineReader, and Readiris. According to the study, Abbyy FineReader had the greatest accuracy rate and was the easiest to use OCR software for the blind.
A machine learning strategy for blind OCR is proposed in [6]. The study employed a deep learning technology known as convolutional neural networks (CNN) to classify text into several categories such as headlines, paragraphs, and lists. A recurrent neural network (RNN) is also used in the study to improve OCR accuracy by taking into consideration the context of the text. According to the study, the proposed OCR system performed well in recognizing various types of text.

IV. PROPOSED WORK

The purpose of this research is to develop an application that can detect linguistic inscriptions using optical character recognition (OCR). The application's OCR capability is the subject of our investigation. Artificial neural network (ANN) and convolutional neural network (CNN) technologies are used in the OCR module. Experiments were carried out to evaluate the recognition rates of two OCR systems using test data (both preprocessed and real images).
CNN was recognised as the best-performing OCR solution after a review of each OCR solution. The main shortcoming of this study is a lack of data, which will have a significant negative impact on OCR accuracy.

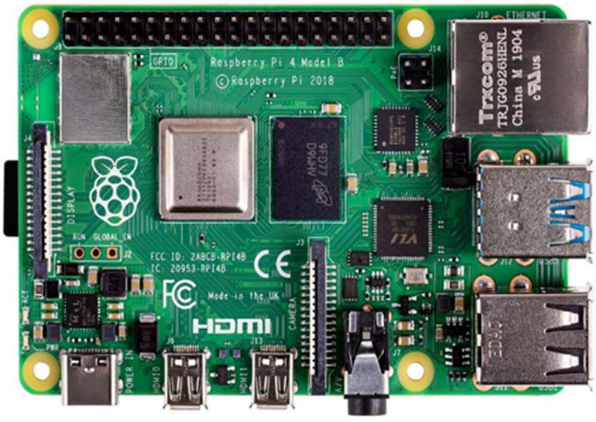
V. BLOCK DIAGRAM



VI. MODULES

1. OPENCV:

OpenCV provides a comprehensive set of image and video processing operations, including image and video capture, processing, filtering, feature identification, object recognition, and face recognition. Support vector machines, neural networks, and decision trees are among the machine learning approaches featured. OpenCV provides a comprehensive set of image and video processing operations, including image and video capture, processing, filtering, feature identification, object recognition, and face recognition. Support vector machines, neural networks, and decision trees are among the machine learning approaches featured. OpenCV provides a comprehensive set of image and video processing operations, including image and video capture, processing, filtering, feature identification, object recognition, and face recognition. Machine learning methods such as neural networks, decision trees, and support vector machines are also included.



2. PYTESSTRACT:

A Python Library for Optical Character Recognition. Pytesstract allows users to perform OCR on images in various formats, including PNG, JPEG, and BMP. The library provides a set of functions that enable users to preprocess images, such as resizing, cropping, and rotating. Pytesstract also includes functionality for improving OCR accuracy, such as deskewing and denoising. Pytesstract is a useful tool for developers looking to build OCR functionality into

their Python applications. With its ease of use and compatibility with other Python libraries, Pytesstract makes OCR accessible to developers of all skill levels. Pytesstract's reliance on the powerful Tesseract OCR engine ensures high OCR accuracy, while its customizable features provide advanced users with the flexibility, they need to develop sophisticated OCR systems.

3.YOLO:

YOLO (You Only Look Once) is an advanced object recognition technique that has sparked considerable interest in the computer vision world. YOLO acts as a real-time object identification system that can precisely identify and localize several entities in a picture. The algorithm performs both object identification and classification tasks using a single neural network, making it quicker and more efficient than previous approaches that employ several networks. YOLO can process images in real-time on a GPU, allowing it to be used in a variety of applications, including surveillance systems, robotics and autonomous vehicles. Furthermore, YOLO has produced cutting-edge performance on benchmark datasets like as PASCAL VOC and COCO.YOLOv2 introduced a number of improvements, including the use of anchor boxes to improve localization accuracy, batch normalization to improve training speed, and a new architecture that allowed for multi-scale feature extraction. YOLOv3 further improved these features, and introduced several new ones, including feature pyramid networks and improved training techniques.

VII. HARDWARE REQUIREMENTS

RASPBERRY PI 4:

The Raspberry Pi 4 Model B is a computer with a single board that was launched as the fourth edition of the Raspberry Pi range in June 2019. It is a debit card-sized computer with a 1.5 GHz 64-bit ARM Cortex-A72 quad-core processor and 2GB, 4GB, or 8GB of LPDDR4 RAM, depending on the model. Dual-band 802.11ac WiFi, Bluetooth 5.0, Gigabit Ethernet, two USB 2.0 ports, two USB 3.0 ports, two micro-HDMI connections (up to 4Kp60 enabled), and a 3.5mm audio socket are also included with the Raspberry Pi 4 Model B. It also has a DSI display port and CSI camera port for connecting a display and camera. The enhanced performance of the Raspberry Pi 4 iteration B is one of the significant enhancements over the previous iteration, the Raspberry Pi 3. It may run a variety of operating systems, including several versions of Windows 10 and Linux IoT Core. The Raspberry Pi 4 Model B is popular for a variety of projects, including media centers, gaming consoles, home automation, and educational purposes, due to its affordability, flexibility,
and ease of use.

BROADCOM BCM 2711 PROCESSOR

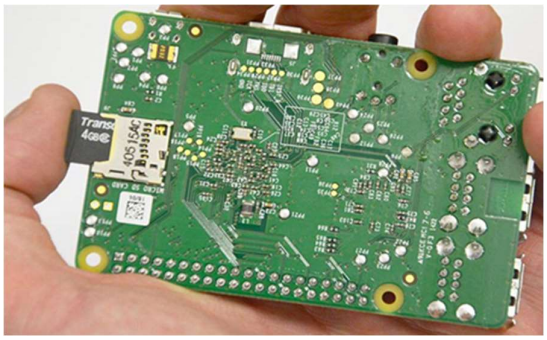
The Broadcom BCM2711 is an ARM Cortex-A72 quad-core CPU featured in the Raspberry Pi 4 Model B. It was first introduced in 2019 and is manufactured using a 28nm process. The processor has a clock speed of 1.5GHz and is capable of running at up to 1.8GHz with appropriate cooling. The BCM2711 has a number of features that make it well-suited for use in single-board computers like the Raspberry Pi 4 Model B. These include:

Quad-core processing: The processor is capable of running four independent processing cores, which can run multiple tasks simultaneously.
64-bit architecture: The processor is capable of addressing more memory than previous Raspberry Pi models, with up to 8GB of LPDDR4 RAM supported.
Improved graphics: The BCM2711 includes a Video Core VI graphics processor, which is capable of rendering 4K video at 60 frames per second.
Connectivity: The processor includes a range of interfaces, including USB 2.0 and 3.0 ports, dual-band 802.11ac wireless, Bluetooth 5.0, and gigabit Ethernet.

SD CARD PORT:

The microSD card slot on a Raspberry Pi is located on the bottom of the board, near the HDMI port. To insert a microSD card, gently push it into the slot with the gold contacts facing downwards until it clicks into place. To remove the card, gently push it in again, and it should pop out. It's important to use a high-quality microSD card that is compatible with the Raspberry Pi to ensure reliable performance. Higher capacity cards may also be used, but some older Raspberry Pi models may have limitations on the maximum card size. In addition to storing the operating system, microSD cards can also be used to store programs and data. For example, you could store your Python code or media files on the microSD card instead of the Raspberry Pi's limited internal storage.

SD CARD PORT IN RASPBERRY PI



XI. REFERENCES

[1] Yuming He “Research on Text Detection and Recognition Based on OCR Recognition Technology”, 2020
[2] Siva Ganesh, Giri “Neural network-based systems for OCR applications”, 2012
[3] Alessandra M. Soares, Bruno J.T. Fernandes “Lateral Inhibition Pyramidal Neural Networks Designed by Particle Swarm Optimization”, 2014
[4] Akshay Bhasin, Amarjot Singh “A Survey of OCR Applications”, 2019

[5] Das, A. K., & Mishra, A. (2018). OCR for the Blind: A Comparative Study of Commercial OCR Software. International Journal of Computer Applications, 181(24), 12-18.

[6] Bhattacharjee, S., & Dey, S. (2017). OCR for the Blind: A Machine Learning Approach. International Journal of Computer Applications, 169(9), 17-22.

[7] Ren S, He K , Girshick R , et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. 2015.

[8] Tian, Zhi, et al. "Detecting Text in Natural Image with Connectionist Text Proposal Network." (2016)

[9] Shi, Baoguang, X. Bai, and C. Yao. "An End-to-End Trainable Neural Network for Image-based Sequence Recognition and Its Application to Scene Text Recognition." IEEE Transactions on Pattern Analysis & Machine Intelligence 39.11(2015):2298-2304.

[10] Platt, J. "Fast Training of Support Vector Machines using Sequential Minimal Optimization. " 2000

[11] N. Ramshankar and P.M. Joe Prathap “Reviewer reliability and XGboost whale optimized sentiment analysis for online product recommendation” 2023.

[12] V Alappatt, and PMJ Prathap “Hybrid cryptographic algorithm based key management scheme in MANET” 2020.

[13] VG Menon and PM Prathap “Survey on latest energy based routing protocols for underwater wireless sensor networks”2016

[14] M Selvi, and PM Joe Prathap “Analysis & classification of secure data aggregation in wireless sensor networks” 2019.