Project/Thesis No. :

CSE 4000: Thesis/ Project

DiyaGlass: An Assistive Device to Read and Move

By

Sadman Sakib

Roll: 1807020

&

Md Shanjid Arefin

Roll: 18070104

Department of Computer Science and Engineering

Khulna University of Engineering & Technology

Khulna 9203, Bangladesh

February, 2024

DiyaGlass: An Assistive Device to Read and Move

By

Sadman Sakib
Roll: 1807020
&
Md Shanjid Arefin
Roll: 18070104

A thesis submitted in partial fulfillment of the requirements for the degree of
"Bachelor of Science in Computer Science & Engineering"

Supervisor:

Dr. Muhammad Sheikh Sadi
Professor
Department of Computer Science and Engineering
Khulna University of Engineering & Technology

Signature

Department of Computer Science and Engineering
Khulna University of Engineering & Technology
Khulna 9203, Bangladesh
February, 2024

Acknowledgement

All the praise to the almighty Allah, whose blessing and mercy succeeded us to start and make progress in this thesis work fairly. We gratefully acknowledge the valuable suggestions, advice, and sincere co-operation of Dr. Muhammad Sheikh Sadi, Professor, Department of Computer Science and Engineering, Khulna University of Engineering Technology, under whose supervision this work is going on. His intellectual advice, encouragement and guidance make us feel confident and motivate us to continue this journey to complete the thesis. We would also like to convey our heartiest gratitude to all other faculty members, officials and staff members of the Department of Computer Science and Engineering as they have always extended their cooperation to continue this work. Finally, we wish to thank our friends and family members for their constant support.

Authors

Abstract

Visual impairments have emerged as a prevalent issue over recent decades. The daily lives of visually impaired individuals are often fraught with challenges, leading to a significant dependence on others for assistance with everyday tasks. Though various systems have been developed to aid visually impaired on their day-to-day life, many of these are constrained by their specific domains. In this thesis, we have taken a comprehensive approach to develop a solution aimed at enhancing the mobility and reading capabilities for visually impaired. The solution is provided by developing an embedded spectacle glass with real-time obstacle avoidance, fall detection and textual information extraction modules. The system generates an audible alert upon detecting obstacles ahead. In the event of an accident, the fall detection module sends an emergency response to a designated caregiver. The optical character recognition (OCR) module includes various image preprocessing techniques, an OCR engine, and a text-to-speech converter. The overall accuracy of object detection module in indoor and outdoor environment is 99.46% and 99.36% respectively. In optical character recognition module, overall character error rate (CER) and word error rate (WER) is 6.97% and 9.46% respectively. The thesis provides a comprehensive solution for visually impaired by developing a portable, wearable, lightweight and low-cost spectacle by which they can get assistance to detect obstacles around them and get textual information easily.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Introduction

This chapter provides an overview of the research problem that this thesis addresses. It begins by defining the problem and then discusses the objectives of this research. The chapter then outlines the research scope and the overall planning of this thesis progression. Finally, the chapter concludes by summarizing the applications of the work and providing an overview of the remaining chapters.

1.2 Problem Statement

The eyes are a person's primary sense organ. It only takes a quick look around to realize that the majority of the information in our surroundings is visual. Blindness and visual defects lead to a variety of public health, social, and economic problems all over the world [1]. A recent survey by the World Health Organization (WHO) shows that approximately 2.2 billion people live with some form of vision impairment globally [2]. The leading causes of vision impairment and blindness at a global level are refractive errors and cataracts**.** The situation is more alarming in the context of Bangladesh as over 90% of individuals with blindness and low vision live in developing countries [3]. The three most common causes of blindness today are glaucoma [4], cataracts [5], and diabetic retinopathy [6]. Accidents can also cause some people to go blind. Performing daily tasks without the help of others is troublesome for visually impaired people. And the most noteworthy part is when they explore a new place, they should understand the barriers' position and other objects in their course for their safe navigation [7]. Their frequent mobility is compromised as they often fall victim to unexpected troubles due to their inability to navigate around obstacles, which can lead to emotional distress or unforeseen incidents [8]. As such, individuals need assistance from others or assistive technology in order to carry out their everyday activities, such as constant navigation and other duties, without difficulty. Ensuring secure and safe mobility for those with visual impairments is a challenging endeavor that demands constant data collection and precise real-time analysis.

Improving mobility and safe navigation is essential to enhance the quality of life for visually impaired individuals. Recent developments in sensor-based systems have created new opportunities to improve the mobility and independence of blind people. Any system to improve mobility and safe navigation needs to meet the hard deadlines and any delay can cause potentially hazardous situations. Various sensor-based systems have demonstrated promising results in responsiveness and real-time functionality, as they can instantly gather information from the environment [9], [10], [11], [12]. Additionally, IoT (Internet of Things) enhances the functionality of sensor-based devices by enabling seamless connectivity and real-time data exchange, thereby optimizing their effectiveness in aiding visually impaired individuals [13], [14].

People who are blind find it difficult to read written text since they cannot see the words. Using braille, audiobooks, human readers, etc. are a few of the conventional techniques. People who are blind frequently find themselves in situations where they need to access visual information that isn't easily accessible in standard formats. When reading books, documents, signs, or labels, for example, this barrier may severely hinder their ability to obtain information in a variety of circumstances. In order to solve this problem, optical character recognition (OCR) technology is essential. It transforms printed text into digital text that can be expressed in Braille on refreshable Braille displays [15] or spoken out by a speaker system [16]. Blind people can now freely access and understand printed information thanks to optical character recognition (OCR) devices, which use image processing algorithms to evaluate text characters from images taken by cameras or scanners [17]. Blind people can increase their access to written information and get past obstacles to reading by using OCR technology. This improves their chances in the workplace, in school, and in their personal lives.

1.3 Objectives

This thesis aims to achieve specific objectives in order to develop an inexpensive and portable system to enhance the mobility and reading capability for visually impaired. These are to

- design real time obstacle avoidance with an audible warning system.

- design real-time fall detector with the option to notify the user's chosen caregiver in the case of an accident.
- develop a computer vision-based real-time text-from-image generator with the functionality of audio output for the user

1.4 Scope

The scope of this thesis encompasses the development and evaluation of a system to enhance the mobility and reading capability for visually impaired people. This thesis aims to enhance mobility by detecting any obstacle that lies before the user, notifying them promptly to ensure safety and confidence. It also targets the avoidance of mishaps like falls, providing caregivers with real-time alerts to save additional damage. The system's resilience is increased by its capacity to operate in both indoor and outdoor settings. Furthermore, the emphasis on real-time text extraction meets a critical demand that visually impaired people have on a regular basis.

1.5 Unfamiliarity of the Problem

The issue this thesis attempts to address is the lack of comprehensive and integrated approaches in the currently available assistive technologies for the blind and visually impaired. It specifically concentrates on creating a profitable, lightweight, and portable device with real-time smart reading, obstacle avoidance, and fall detection capabilities. Existing systems frequently suffer from being heavy-weight and bulky, which can result in a lack of user-friendliness. Moreover, these systems often struggle to effectively read smaller fonts on documents, leading to limitations in usability and functionality. This thesis aims to address these issues by developing a lightweight and compact solution that enhances accessibility for visually impaired individuals. Through improved design and functionality, we strive to create a system that offers ease of use while effectively reading documents of varying font sizes with precision and accuracy.

1.6 Applications of the Work

The proposed system provides crucial support for visually impaired individuals by addressing key challenges they face in navigation, safety, and information access. Through advanced technologies, including computer vision and sensor analysis, our system offers are the following.

- **Obstacle Detection and Avoidance:** Real-time detection of obstacles such as curbs, poles, and stairs, with immediate alerts to enable safe navigation.
- **Fall Detection**: Automatic recognition of fall events, triggering prompt emergency response to ensure timely assistance.
- **Emergency Response:** Quick distress signal transmission to pre-defined contacts or emergency services, including precise geolocation for swift intervention.
- **Optical Character Recognition (OCR):** Text recognition from various sources like signage and documents, converting it into speech or Braille for improved information access.

In summary, our system enhances mobility, safety, and information accessibility for visually impaired individuals, fostering independence and inclusivity in their daily lives.

1.7 Thesis Organization

The thesis is organized as below:

Chapter 1 briefly explains the introduction of the thesis and background of works. The problem statement, objectives of the thesis, and Project planning Gantt Chart discussed elaborately here.

Chapter 2 represents the existing works in the related field and focuses mostly on the advantages and drawbacks of existing works.

Chapter 3 explains the proposed methodology of the thesis.

Chapter 4 discusses about the dataset used, experimental setup and implementation and the result of the thesis.

Chapter 5 discusses about the societal, health, environment, safety, ethical and cultural issues of the thesis.

Chapter 6 addresses complex engineering problems and activities with the thesis.

Chapter 7 concludes this thesis along with recommendations and future works.

Chapter 2

Literature Review

2.1 Introduction

This chapter provides an overview of the current state-of-the-art advancements in both mobility and readability, outlining their methodologies and discussing their respective limitations. Some related terminologies are also discussed in this chapter.

2.2 Obstacle Avoidance and Fall Detection

Ahsan Habib et al. developed a hybrid system for staircase detection for visually impaired [18]. The technology uses an RGB-D camera to take pictures of staircases and compares them to pre-trained templates by combining an ultrasonic sensor and a pre-trained model. To precisely recognize stairs, the framework combines ultrasonic sensors with Faster R-CNN for object detection with an average accuracy of 98.73%.

Md. Mostafa Kamal et al. introduces a wearable, lightweight set of glasses that uses smartphone processed RGB data to help blind people find obstacles and move [19]. The obstacle detection module makes use of two sonar sensors. Surface smoothness is computed by the system in both light and dark conditions by using image processing techniques. The highest smoothness of the surface is achieved by the system is 96.341% and 98.683% for the day-night and dark respectively.

Nabila Shahnaz Khan et al. developed a portable, lightweight device meant to replace conventional walking sticks [20]. This innovative system incorporates ultrasonic sensors and a GPRS module, enabling it to detect obstacles, manholes, and potholes. Additionally, by just tapping a button, users can send their location via SMS to their trusted contacts

Emilia Sipos et al. presents a feature-rich and reasonably priced smart assistant system to help visually impaired people navigate different surroundings efficiently [21]. The assistant, which consists of a central unit and a smart cane, allows for easy voice message transmission and is usable by users with varying levels of IT proficiency. With features including an electronic compass, GPS, Wi-Fi, optical, ultrasonic, and RFID reader, the assistant helps users navigate safely even when they are not online. Preliminary tests provided encouraging results, indicating that the prototype has the potential to help visually impaired people to achieve a high level of independence in daily activities.

Wan-Jung Chang et al. proposed a system that integrates wearable smart glasses, an intelligent walking stick, a mobile app, and a cloud-based information management platform [22]. By wearing the smart glasses and utilizing the walking stick, visually impaired individuals can detect aerial obstacles and detect falls while navigating outdoor environments. The intelligent walking stick provides haptic feedback to guide users away from potential collisions with aerial obstacles. Additionally, in the event of a fall, the system promptly notifies family members or caregivers. Experimental results showed that the proposed system could detect aerial obstacles within 3 meters, and the average accuracy of fall detection reached up to 98.3%.

Tayla Froneman et al. constructed a system where Six 40 kHz ultrasonic sensors were installed on a polyester belt that was worn around the user's waist to form the system [23]. The sensors were set up to identify objects both on the ground and in the air, and they were accompanied with vibrating motors that were placed thoughtfully throughout the belt to provide feedback. The system's operation mode was signaled by a buzzer that could be changed via a toggle switch according to the user's surroundings. The proposed methodology has an accuracy of 98.57%.

Md. Atikur Rahman et al. constructed a system that gives an auditory alert to the user and helps the blind identify a number of things [24]. The system uses four laser sensors (VL53L0X) to identify items facing the front, left, right, and ground. An accelerometer (ADXL345) is used by the suggested system to detect free fall. The suggested system recognizes items and currency in a real-time scenario in both indoor and outdoor settings using the Single Shot Detector (SSD) model with MobileNet and Tensorflow-lite. The experimental result showed an accuracy of 99.33% for obstacle distance at 210 cm.

2.3 Optical Character Recognition

Muhammad Farid Zamir et al. presented a real-time smart reader system that makes use of optical character recognition (OCR) in their study [25]. Using a camera, this system takes pictures, which are subsequently subjected to a number of image preprocessing techniques including binarization and adaptive thresholding. Finally, the Tesseract OCR engine is used to extract the required output from the photos that have been processed.

Anush Goel et al. proposed a raspberry pi based real time smart reader system [26]. Using a camera, this system takes pictures, which are subsequently subjected to a number of image preprocessing techniques including binarization, noise removal and skeletonization. Then segmented words are then fed to Tesseract OCR engine to get the output text.

Gauri Vaidya et al. introduced a camera-based text reading framework that makes it easier for those with visual impairments to understand text in everyday settings, product labels, etc. [27]. The work at hand is split into two parts: text detection and text recognition. Text localization includes binarization and removing non-text areas of a picture using geometric and stroke width filters. Finally, the image is given to OCR engine to get the text from image.

M. Rajesh et al. introduces a real-time smart reader and face recognition system developed on raspberry pi 3B [28]. In the proposed system, after converting the collected image to grayscale and applying Gaussian filtering to minimize noise, adaptive Gaussian thresholding and binary conversion are applied to improve clarity. Following the removal of non-character regions, the image is subjected to Tesseract OCR processing for text recognition, yielding an output text file. From the text file, e-Speak then creates an analog signal that is sent to headphones for audio output.

Hao Jiang et al. developed a raspberry pi based smart reader. The process entails pressing a button to take a picture, which is subsequently converted to grayscale image [29]. The grayscale image is cleaned up of redundant noise, and edges are detected using edge detection techniques. After applying contour finding, the largest contour serves as the region of interest (ROI). After extracting text from the image using Tesseract, the text is sent into the eSpeak text-to-speech program.

Vijayanarayanan A. et al. developed an android based OCR system [30]. In the proposed system, there are three primary processes in the character recognition process. The OCR algorithm first processes the image by identifying the region containing the sign with the text that has to be recognized. It then outputs the raw text. After that, the text is transmitted to another algorithm that is in charge of editing it and, if feasible, retrieving sentences. Lastly, the Text-to-Speech algorithm is used to translate the outcome into text.

A. G. Sareeka et al developed an optical character recognition with text to speech system for visually impaired [15]. The system involves preprocessing the input image through binarization, noise reduction, normalization, skew correction, and slant removal to prepare it for character analysis. Features extraction follows, where each character is represented as a feature vector. Finally, classification and recognition are performed using a classifier to compare extracted features with a data bank, followed by text generation using an OCR system, potentially incorporating deep learning architectures like LSTM and gTTS for text-to-speech conversion.

2.4 Research Gap Solutions

The limitations of existing systems and their potential solutions that our thesis provides are discussed in table 2.1.

Table 2.1: Addressing Similar Research Landscape Challenges: Comparative Analysis of Innovative Approaches

Authors	Limitations	Proposed Solution
Ahsan Habib et al. [18]	Used ultrasonic sensor that has lesser range, less accuracy and sensitive to noise. Used a comparatively low processor device which needs time to process data	Using laser sensor with long range and more robust to noise. Using high processing capable raspberry pi
Md. Mostafa Kamal et al. [19]	Used vibrator to notify user of the danger ahead	Using sound system to notify user of the danger ahead
Nabila Shahnaz Khan et al. [20]	Used arduino uno which is a slow processor with low storage	Using raspberry pi which is faster, has large storage and capable of multi-functionalities

Table 2.1: Addressing Similar Research Landscape Challenges: Comparative Analysis of Innovative Approaches

Authors	Limitations	Proposed Solution
Emilia Sipos et al. [21]	System is bulky, need additional space for central unit and not embedded in a single device	The whole system is embedded in one single device. The device is light-weight and portable
Wan-Jung Chang et al. [22]	Overall the system is large and costly as two separate devices are used. Alert user using vibrator only	Using sound system to notify user of the danger ahead
Tayla Froneman et al [23]	Used ultrasonic sonar sensor, alert user using vibrator and buzzer. Wearable device is not user-friendly, can be awkward when wearing it outside	Using laser sensor with long range and more robust to noise. Audible Bluetooth based sound system to notify user. The developed prototype is smart and lucrative
Muhammad Farid Zamir et al. [24]	Do not address the issue of skewed or curved lines in the captured document images, resulting degradation in accuracy for OCR systems.	Provides solutions for skewed or curvy lines.
Anush Goel et al. [25]	Doesn’t clarify how line segmentation is done	Proposed a clear methodology to show how various lines from the image are extracted
Gauri Vaidya et al. [26]	Used global thresholding for binarization, thus leading problems for the image with uneven illumination.	Used local thresholding to solve the problem
M. Rajesh et al. [27]	Provides no solution if the document lines are curvy and skewed	Provides a solution for de-skewing the skewed lines.

Table 2.1: Addressing Similar Research Landscape Challenges: Comparative Analysis of Innovative Approaches

Authors	Limitations	Proposed Solution
Hao Jiang et al. [28]	Used global thresholding for binarization, recognize few numbers of characters.	To resolve the uneven illumination problem, local thresholding is applied. The proposed system can recognize all the 26 letters.
Vijayanarayanan A. [29]	Some lines are not detected in real-time if image is distorted perspectively.	Methodology involves removing perspective distortions

2.5 Related Terminologies

This section includes all related terms including the theories of the technologies used for the system.

2.5.1 Raspberry Pi

The Raspberry Pi compact Compute Module 4 has been picked as the microcontroller in our framework plan because of its reduced size and convenient comfort. It has one USB port through which USB Hub can be associated. CM4 is a popular choice for embedded and IoT applications [31] and is used in numerous commercial devices. CM4-NANO-A is the mini board of Raspberry Pi Compute Module 4, which is a baseboard of Raspberry Pi Compute Module 4 with a 5V/2.5A USB Type C interface.

2.5.2 VL53L0X

The VL53L0X is a Time-of-Flight distance sensor with a tiny invisible laser source and a matching sensor. It precisely measures the time taken for light to bounce back, providing accurate distance readings. Its narrow light source makes it ideal for close-range surface

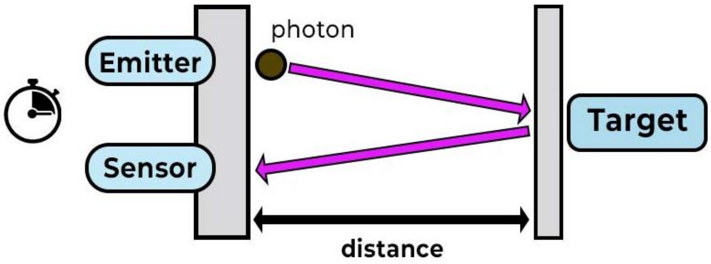


Figure 2.1: Working Principle of Time-Of-Flight Sensor

detection. Unlike sonars or IR sensors, it avoids linearity issues and double imaging. This small sensor is perfect for robotics or interactive projects, featuring easy integration with a breakout board that includes a regulator and level shifting for 3-5V microcontrollers. The working principle of VL53L0X is depicted in Figure 2.1. Equation (2.1) illustrates how the distance is computed using the light beam's triangulation. At an occurrence point, the sensor emits a laser light shaft that strikes an object, bounces off it, and can be identified [24]. The measured object, the detector, and the laser source form a triangle. Distance is calculated using elementary mathematics by determining the precise region where the laser hits the locator.

$$Distance = (speed\ of\ light \times elapsed\ time) \ldots\ldots\ldots\ldots\ldots \quad (2.1)$$

It operates in a range of approximately 50 millimeters to 1.2 meters in default mode and up to 1.5 to 2 meters in 'long-range' mode on reflective surfaces. Ranging accuracy varies from 3 to 12%, influenced by ambient lighting and surface properties, with better results in well-lit and shiny environments [32].

2.5.3 Accelerometer

The ADXL345 as depicted in Figure 3.3 is a compact 3-axis accelerometer offering high-resolution (13-bit) measurements up to ±16 g. It communicates through SPI or I2C interfaces and is well-suited for tilt-sensing applications, measuring static gravity acceleration, as well as dynamic acceleration from motion or shock. With a high resolution of 4 mg/LSB, it can detect inclination changes below 1.0° [33]. Special functions include activity and inactivity sensing, tap sensing for single and double taps, and free-fall detection. These functions can trigger interruptions on two output pins. The sensor features

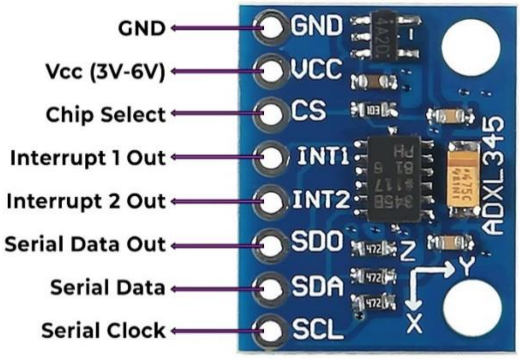


Figure 2.2: ADXL345 Accelerometer Pin Diagram

a 32-level FIFO buffer for data storage, minimizing processor intervention, and offers low-power modes for efficient motion-based power management with threshold sensing and low power consumption during active acceleration measurement.

2.5.4 Multiplexer

The TCA9548A multiplexer as shown in Figure 3.4 is a device with eight bidirectional translating switches controllable through the I2C bus. It allows selection of individual or combinations of downstream channels to resolve I2C slave address conflicts. For example, if eight identical digital temperature sensors are needed in the application, one sensor can be connected at each channel :0-7 [34]. The system master can reset the TCA9548A in the event of a time-out or other improper operation by asserting a low in the RESET input.

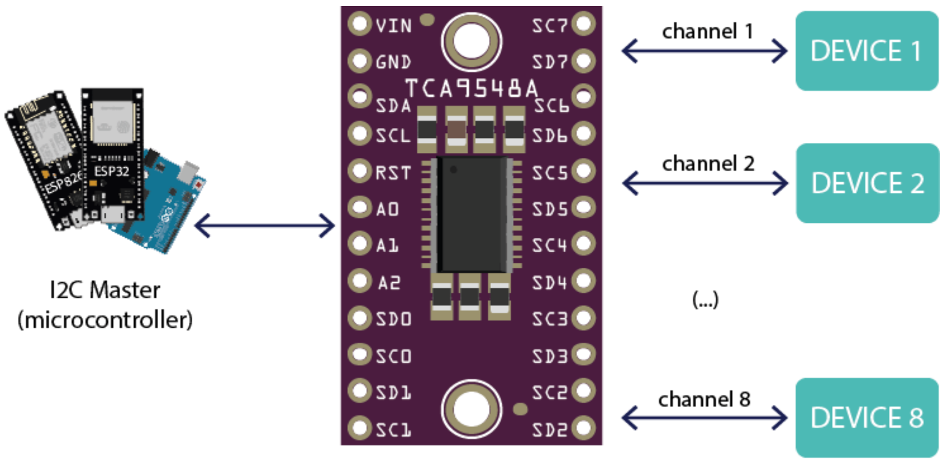


Figure 2.3: TCA9548A Multiplexer [35]

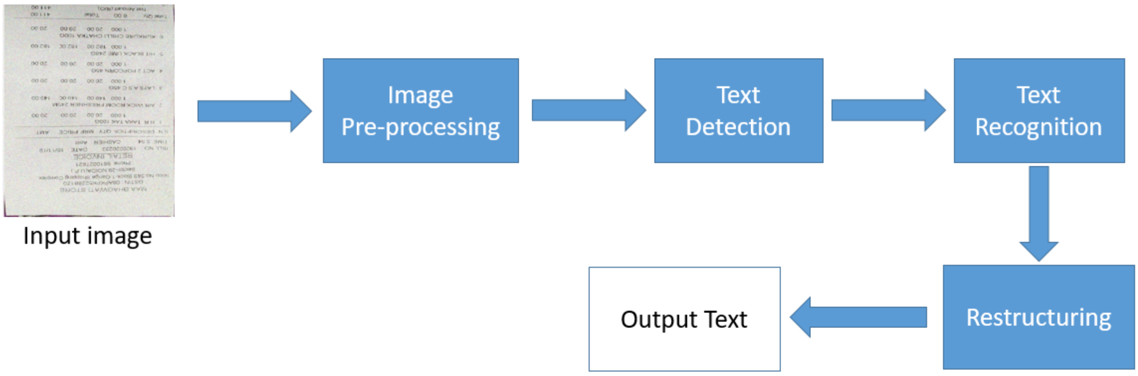


Figure 2.4: Pipeline of optical character recognition (OCR). [16]

Similarly, the power-on reset deselects all channels and initializes the I2C/SM Bus state machine. Pass gates of the switches are designed to limit the maximum high voltage (VCC) and enable communication between different bus voltages (e.g., 1.8V, 2.5V, 3.3V, and 5V) without additional protection. All I/O pins are 5V-tolerant, enhancing versatility in various applications.

2.5.5 OCR

Optical character recognition, usually abbreviated to OCR, is the mechanical or electronic conversion of scanned or photographed images of typewritten or printed text into machine-encoded/computer-readable text [36]. In the era of deep learning, optical character recognition has brought a revolution. The main pipeline of OCR engines involves image pre-processing, text segmentation from non-text part of image and text recognition.

2.5.6 Tesseract

Tesseract OCR functions step-by-step is shown in Figure 3.6. The first step in converting an image to a binary picture is called adaptive thresholding. The next step is linked component analysis, which extracts character outlines. Because it can perform OCR on an image with a black backdrop and white text, this approach is especially useful. Tesseract was perhaps the first to provide this kind of processing. Subsequently, the outlines are used to construct Blobs. Text lines are created by arranging blobs, and lines and regions are then checked for a predetermined area or text size equivalent. The text uses both definite and fuzzy spaces to separate words. Next, text recognition is started as a two-pass process, as seen in. Every word in the text is attempted to be recognized during the first pass. Every word that passes is fed into an adaptive classifier as training data. The adaptive classifier tries to recognize text more accurately. Since the adaptive classifier has gained new knowledge from the training set, several issues are resolved, and text is retrieved from images in the final stage [37].

2.5.7 Dilation

Dilation is one of the basic morphological operations. Morphology is a comprehensive set of image processing operations that process images based on shapes [38]. Dilation adds pixels to the edges of objects or add pixels in the image as shown in Figure 3.6. To begin, we execute a dilation operation over the picture object by traversing the structuring element, as illustrated in Figure 3.6. The output pixel values are calculated using the Eq. (3.2).

$$Pixel(output) = \begin{cases} 1(if\ hit) \\ 0(if\ miss) \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ (2.2)$$

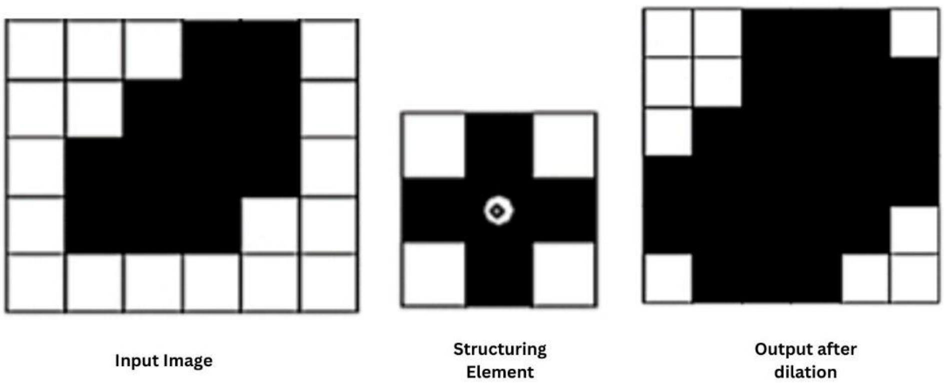


Figure 2.5: Dilation with structuring element

2.5.8 Morphological reconstruction

Three key components are used in the morphological reconstruction process: marker points, source image, and marker image.

Source Image: The reference image used in the morphological reconstruction is the source image, also known as the mask image in some research articles.

Marker Image: The source image is subjected to erosion or dilation to produce the marker image, which is used in the reconstruction process. It is also possible to use pre-existing photos for the marker image. The dimensions of the source image and the marker image must match.

Marker Points: These are user-specified locations inside the image that indicate the beginning of the reconstruction process.

Morphological reconstruction is of 2 types. They are as follow.

Reconstruction by Dilation: In binary images, reconstruction by dilation reconstructs particles, while in grayscale images, it reconstructs bright patches. Neighboring pixels are reconstructed by distributing the brightness value, beginning at the marker points. Reconstruction by dilation reconstructs the surrounding pixels spanning from 0 to the maximal valued pixel, beginning with the maximal gray valued pixels of the marker [39].

Reconstruction by Erosion: Reconstruction by erosion restores holes in a binary image and dark areas in a grayscale image. Neighboring pixels are recreated by spreading the darkness value, beginning at the marker points. The least valued pixels of the marker serve as the starting point for reconstruction by erosion, which then reconstructs the nearby pixels up to the maximum value of the image [39].

2.5.9 Canny Edge

In a digital image, edges are notable local variations in brightness. A collection of connected pixels that creates a border between two discontinuous regions is referred to as an edge. Edge detection is a pre-requisite for various computer vision-based applications**.** Canny edge [40] is one of widely used edge detection algorithm. Canny edge detection is composed of 5 steps. They are as follows.

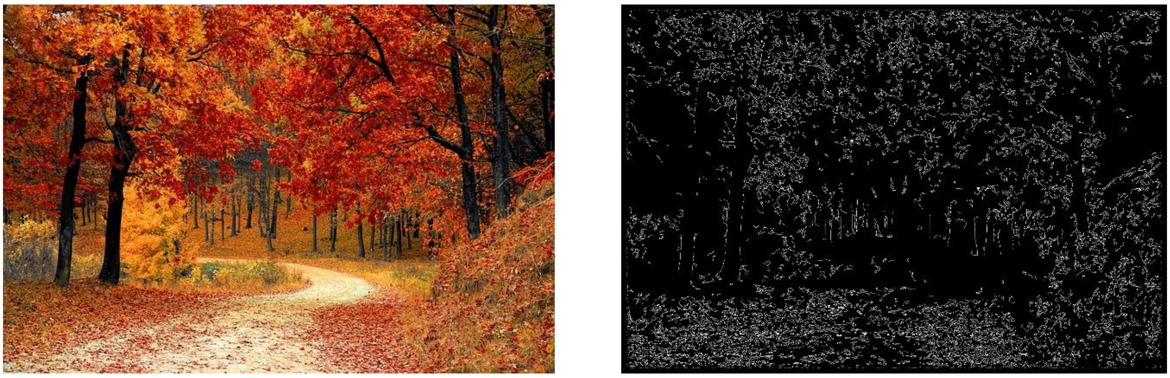
1. Gradient calculation: Applying sobel Filter that highlights intensity change in both the x and y-axis. Then, the magnitude G and the slope θ of the gradient are calculated.

2. Non-maximum suppression: Performs non-maximum suppression to thin out the edges.

3. Double threshold: Three types of pixels are intended to be identified by the double threshold step: strong, weak, and non-relevant. Strong pixels are those whose intensity is so great that we can be certain they add to the final edge. Pixels with an intensity value that is too high to be classified as strong, but too low to be regarded as irrelevant for edge detection, are referred to as weak pixels. For the edge, other pixels are regarded as irrelevant.

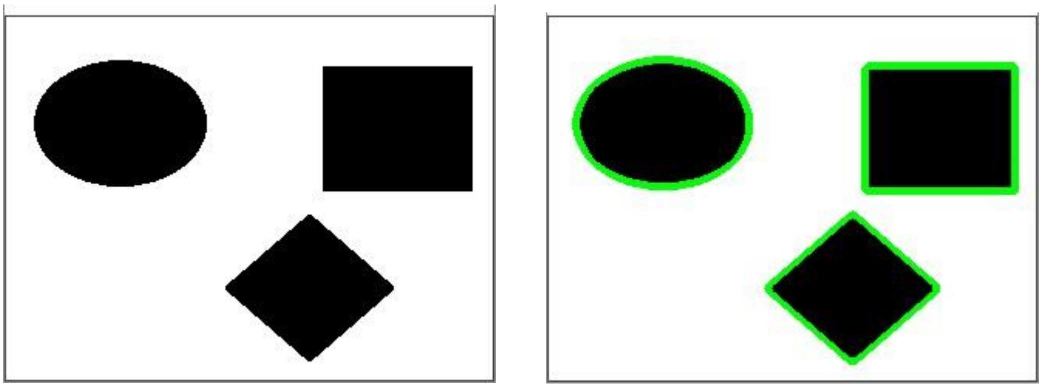
4. Edge Tracking by Hysteresis: The hysteresis is based on the threshold results and involves turning weak pixels into strong ones if and only if at least one of the surrounding pixels is a strong pixel.

An image and its corresponding canny edge detected output is shown Figure 2.6.



(a) (b)

Figure 2.6: (a) Input image, (b) Image after canny edge detection.



(a) (b)

Figure 2.7: (a)Input image, (b)Contours of the input image highlighted by green outline.

2.5.10 Contour finding

Any edge detection technique, such as the "canny edge," is used to identify the gradient of edges in order to find contours. Contour tracing algorithms can be used to identify continuous curves that represent object boundaries once edges have been detected. Contour approximation techniques can be used to simplify contours while maintaining their shapes once they have been recognized. An approach that is frequently used for contour approximation is the Douglas-Peucker algorithm. An image and its corresponding contour highlighted image is shown Figure 2.7.

2.5.11 Local and global threshold

Thresholding is one of the segmentation techniques that generates a binary image. A binary image is one whose pixels have only two values – 0 and 1 and thus requires only one bit to store pixel intensity. To generate a binary image from a given grayscale image it is required to separate it into two regions (background and foreground) based on a threshold value. Global threshold can be applied if the histogram of any image is bimodal. A bimodal histogram is one whose histogram distribution has only one valley as shown in Figure 2.8.

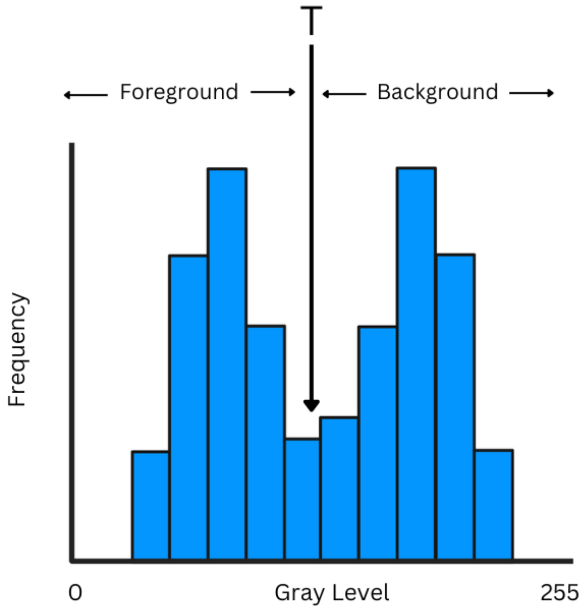


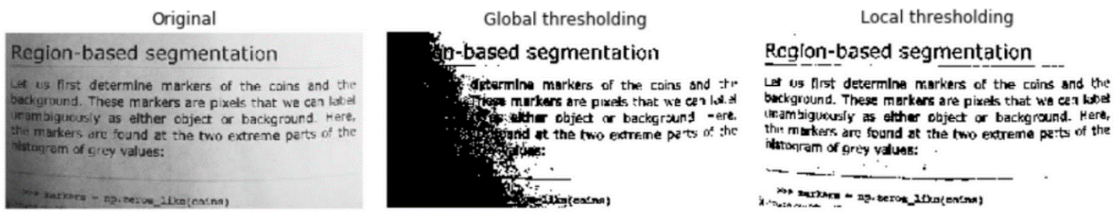
Figure 2.8: Bimodal Histogram.

For any image with bimodal histogram distribution, the foreground and background pixels have gray levels grouped into two dominant modes. One obvious way to extract the object from the background is to select a threshold *T* that separates these modes.

The thresholded image *g(x, y)* is defined as Eq. (3.3).

$$g(x, y) = \begin{cases} 1 \text{ if } (x, y) > \text{T} \\ 0 \text{ if } (x, y) \leq \text{T} \end{cases} \quad \text{.........................} \ (2.3)$$

There are various factors that can cause the illumination of the captured document to be uneven. The problem with global thresholding is that changes in illumination across the



(a) (b) (c)

Figure 2.9: (a) Input image, (b) Image after applying global thresholding (c) Image after applying local thresholding [41]

scene may cause some parts to be brighter (in the light) and some parts darker (in shadow) in ways that have nothing to do with the objects in the image. By setting thresholds locally, we may address such uneven illumination, at least partially. That is, we let the threshold itself gently change throughout the image rather than having a single global threshold [42]. One of the widely used adaptive thresholding techniques is Gaussian adaptive thresholding. In Gaussian Adaptive Thresholding, Gaussian Blur is performed using a small Gaussian Filter, and wherever there are pixel intensities that are more than the pixel intensities of the blurred image set those regions to 255 else 0.

Chapter 3

Proposed Methodology

3.1 Introduction

The proposed system consists of two modules that work together as a smart reader and navigation aid for people with visual impairments. An array of sensors is used in conjunction with the Raspberry Pi, which serves as the central processing unit, to collect input from the environment and make appropriate responses. The character recognition part is divided into three phases. The collected image is first preprocessed using a variety of methods. After that, an OCR engine processes the resulting image in order to extract text. To help the user understand, the captured text is finally turned into auditory sound.

3.2 Detailed Methodology

The system's overall architecture is depicted in Figure 3.1. A detailed discussion of the proposed methodology is described in the following subsections.

3.2.1 Obstacle Avoidance Module

Total three VL53L0X time-of-flight distance sensor is used in the proposed system for measuring the distance of obstacles. VL53L0X time-of-flight is a powerful sensor for measuring distance upto 4 meters with high accuracy. VL53L0X has a fixed i2c address

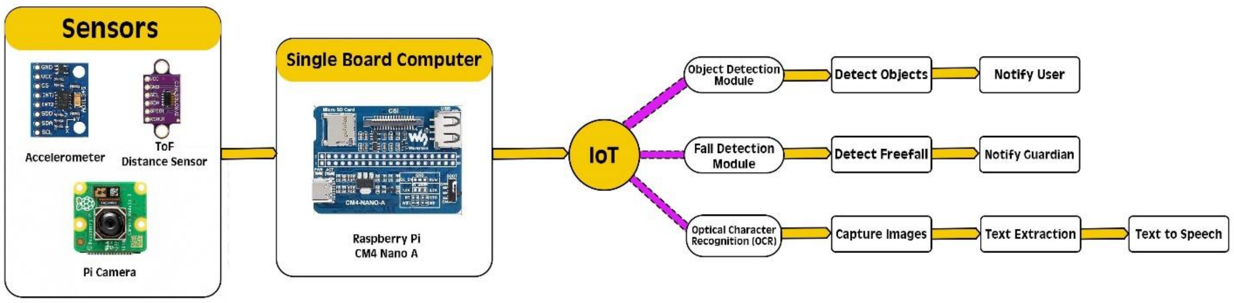


Figure 3.1: System's overall architecture

0X29. There are some problems to operate sensors with multiple Same address I2C concurrently. As SDA (data pin) of raspberry pi transfers data in serial mode, we cannot run sensors concurrently using common SDA pin if their i2c address is fixed. To resolve this issue, a TCA9548A multiplexer is used. Multiplexer expands the i2c bus to expand up to 8 channels theoretically. These time-of-flight distance sensors is continuously reading data from surroundings on system boot. Whenever there is any object within the specified range, a text-to-speech module is invoked to alert the user. Pseudocode of the procedure is given in table 3.1.

Table 3.1: Pseudocode for Laser Sensor

Algorithm 1 Laser Sensor Control Algorithm

```
Define the MUX_ADDRESS as 0X70
Set all shutdown pins to the LOW state
Initialize VL53L0X objects: tof1, tof2, and tof3, with I2C addresses=0X29
Initialize count to 0
while power is on do
  if count is divisible by 50 then
    if flag equals 1 then
      Open tof1 and start ranging.
    else if flag is equal to 2 then
      Open tof2 and start ranging.
    else if flag is equal to 3 then
      Open tof3 and start ranging.
    end if
  end if
  if flag is equal to 1 then
    Obtain the distance measurement from tof3.
  else if flag is equal to 2 then
    Obtain the distance measurement from tof1.
  else if flag is equal to 3 then
    Obtain the distance measurement from tof2.
  end if
  if (distance is lower than threshold) then
    Text-to-speech ("Alert! Obstacle ahead").
  end if
end while
```

3.2.2 Fall Detection Module

Falls have been a common and major problem for the visually challenged and elderly people. This may cause serious injury to them. The proposed system can detect the falls and send notification to the guardian when the falls occur. The proposed system uses an accelerometer (ADXL345) to detect the freefall. A threshold value of acceleration is set and by meeting this threshold value we can determine whether the freefall has occurred. Acceleration data of all 3-axis is read from the accelerometer continuously and is also checked whether the values surpass the threshold value for a certain period of time. If the conditions are fulfilled, then it sets the FREE_FALL bit TRUE and the free fall is detected. Immediately a notification is sent to the guardian's device as "Fall Detected" so that the guardian can take necessary actions and the user may get immediate support. Pseudocode of the procedure is given in table 3.2.

Table 3.2: Pseudocode for Fall Detection

Algorithm 2 Accelerometer Freefall Detection Algorithm

```
 1: Initialize accelerometer.
 2: Define threshold values to trigger freefall detection and acceleration thresh- old.
 3: Read accelerometer data and return a tuple representing the acceleration in each
    axis.
 4: if total acceleration is below threshold then
 5:     Increase count.
 6: else if total acceleration is above threshold then
 7:     Reset count.
 8: end if
 9: if freefall count exceeds threshold then
10:     Set freefall flag to TRUE.
11: end if
12: Reset freefall detected flag when freefall conditions are no longer met.
13: Go to step 3.
```

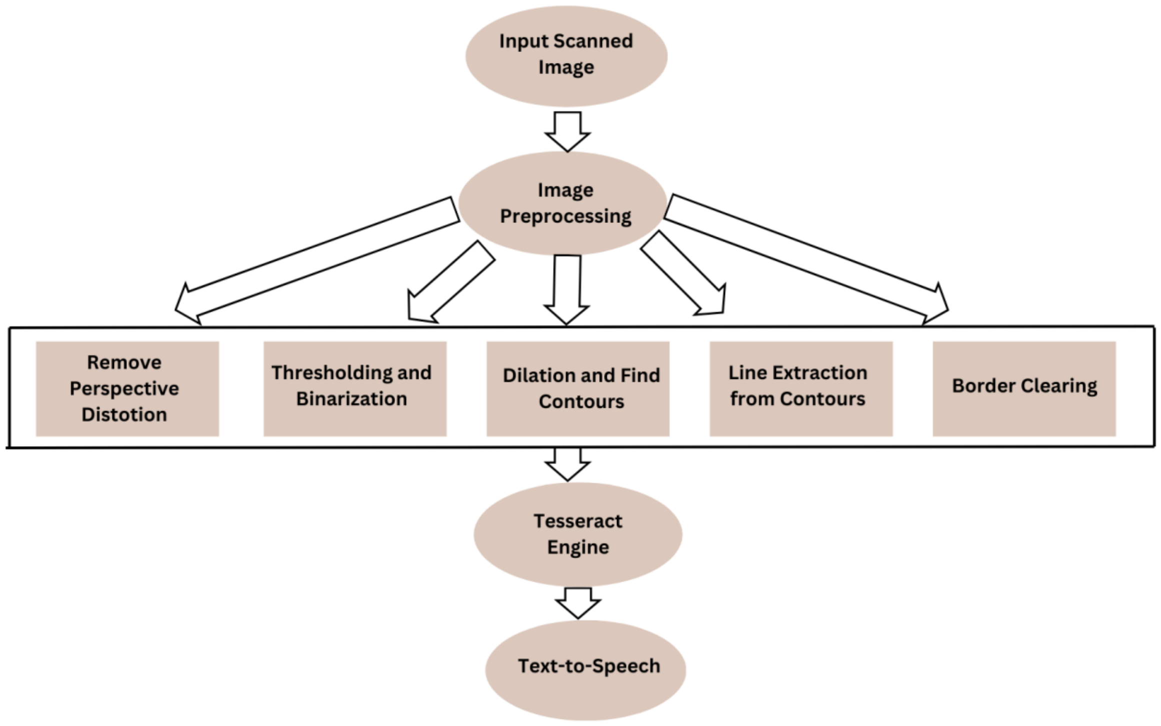


Figure 3.2: Overall procedure for optical character recognition (OCR)

3.2.3 Optical Character Recognition Module

The image is preprocessed in order to improve machine readability before the OCR process begins. The preprocessed image is then feed into the OCR engine so that text may be extracted. The retrieved text is then converted into an audio format for the user via a text-to-speech converter. Figure 3.2 provides an illustration of the whole procedure.

Remove Perspective distortion

Perspective distortion can occur when the camera angle is not perpendicular to the text source when being photographed. It is very difficult to quantify the tilting angles directly, though. OCR accuracy drops for perspective distortion. The following steps are performed to remove the perspective distortion.

1. To find edges in the input image, the Canny edge detection technique is first applied. To fill up any tiny spaces between edges, morphological closing is applied to the edges that have been identified.

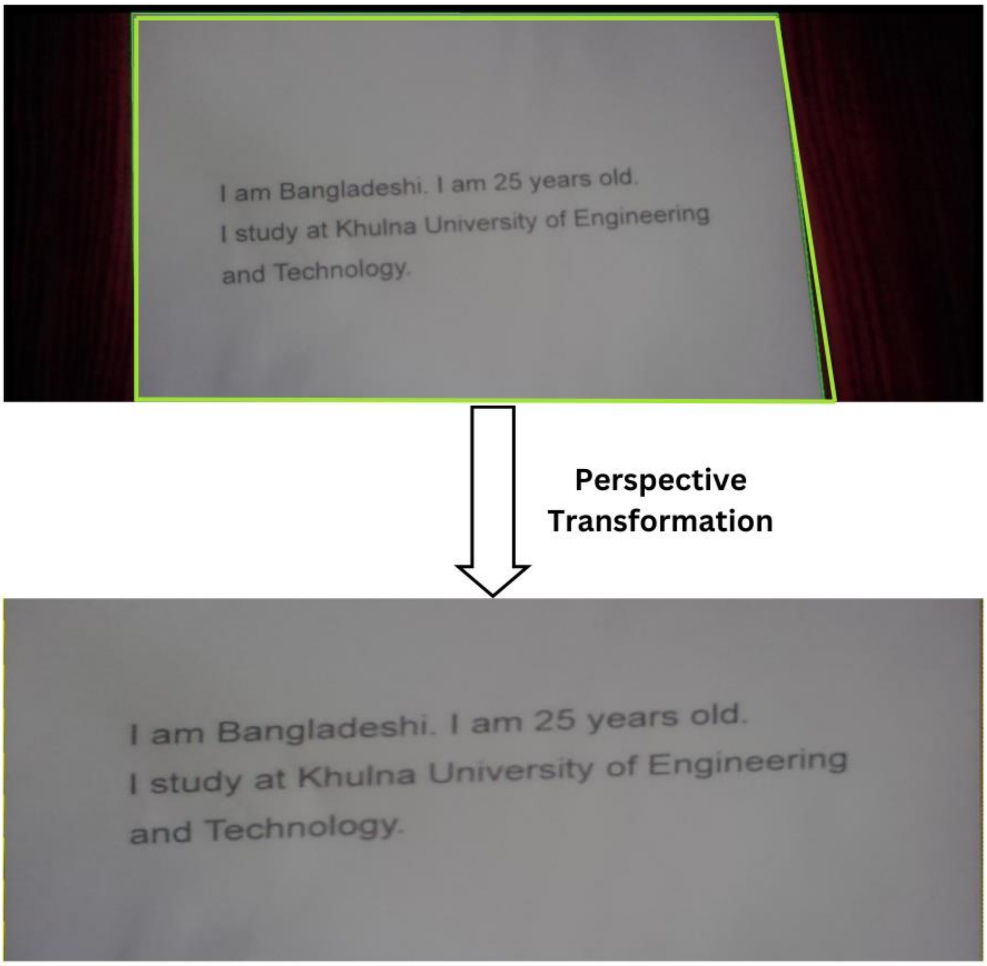


Figure 3.3: Outline of input image (green line) and image after applying perspective transformation

2. After the image has been processed, contours are extracted. A basic approximation of the contour form is derived by taking into account only the outward contours. Subject to area

thresholding, small contours are filtered out. The document outline is chosen from the largest contour as shown in Figure 3.3 (a). Then, the contour is approximated to a polygonal curve to simplify its shape using the Douglas-Peucker algorithm.

3. Perspective transformation is applied to warp the document region to a rectangular shape. This function takes an unordered set of four points: top-left, top-right, bottom-right,

Table 3.3: Pseudocode for Removing Perspective Distortion

Algorithm 3 Removing Perspective Distortion

1: Read image from image path.
2: Calculate height and width aspect ratio.
3: Resize image to 1200x800.
4: Add padding to image.
5: Convert image to grayscale format.
6: Apply Gaussian blur with a kernel size of (7, 7).
7: Set lower threshold value = 75 and upper threshold value = 200.
8: Apply Canny edge detector.
9: Apply Morphological closing operation with a kernel size of (9, 9).
10: Find contours in the image.
11: Filter out small contours and select the largest one.
12: Fit the contour into a quadrilateral shape.
13: Set top-left, top-right, bottom-left, and bottom-right of the outline as source matrix.
14: Set [0, 0], [0, Width-1], [Width-1, Height-1], [Height-1, 0] as destination matrix.
15: Get perspective transformation matrix.
16: Apply the matrix.

and bottom-left. The function takes the input image and the four points representing the region of interest (ROI). Then it computes the perspective transformation matrix required to transform the input image to a top-down view. Then The transformation is applied to the input image. The resultant image is shown in Figure 3.3 (b). A pseudocode of the procedure is given at table 3.3.

Skew Correction

Skew is adjusted by repeatedly searching a variety of angles in order to locate the one that maximizes a predetermined score. To assesses how well-aligned the text lines become at each rotation angle, the image was rotated at various degrees within a range from its initial orientation and the histogram values along each row was added. Smoother and more uniform histograms with greater overall sums signifying better alignment are characteristics of a well-aligned image. On the other hand, an image that is not properly

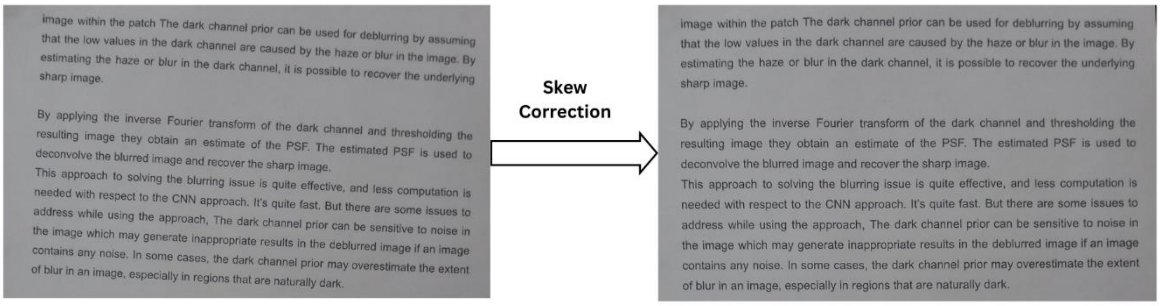


image within the patch The dark channel prior can be used for deblurring by assuming that the low values in the dark channel are caused by the haze or blur in the image. By estimating the haze or blur in the dark channel, it is possible to recover the underlying sharp image.

By applying the inverse Fourier transform of the dark channel and thresholding the resulting image they obtain an estimate of the PSF. The estimated PSF is used to deconvolve the blurred image and recover the sharp image.
This approach to solving the blurring issue is quite effective, and less computation is needed with respect to the CNN approach. It's quite fast. But there are some issues to address while using the approach, The dark channel prior can be sensitive to noise in the image which may generate inappropriate results in the deblurred image if an image contains any noise. In some cases, the dark channel prior may overestimate the extent of blur in an image, especially in regions that are naturally dark.

(a) (b)

Figure 3.4: (a) Original document (b) Skew corrected document.

aligned will provide histograms that are erratic or jagged, with smaller overall sums that indicate misalignment. By calculating a score from the sum of squared differences between successive histogram values, the alignment quality was evaluated. Better alignment will produce a smoother histogram, which will reduce squared differences and increase the total score. The script determines the ideal skew correction by comparing these scores at various rotation angles to determine the angle that maximizes alignment and minimizes histogram variations. An image of a document before and after skew correction is depicted in Figure 3.4. A pseudocode is given in table 3.4.

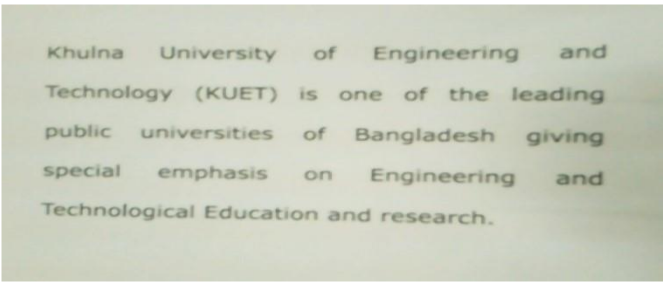
Table 3.4: Pseudocode for Skew correction

Algorithm 4 Skew Correction Algorithm

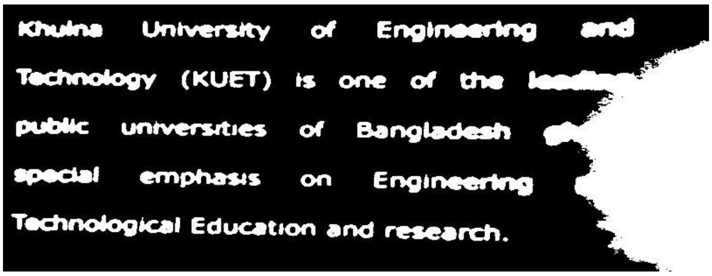
1: Generate a range of angles from $-range$ to $+range$.
2: Initialize a list.
3: **for** each angle **do**
4: Rotate the image.
5: Calculate histogram.
6: Compute the sum of squares of consecutive elements of the histogram.
7: Add the result to the list.
8: **end for**
9: Get the maximum value from the list.
10: Set the corresponding angle of the score as the desired angle.
11: Perform rotation.

Thresholding

As the document has uneven illumination and unwanted shadow, global threshold performs poorly as shown in Figure 3.5 (b). Thus, a local thresholding technique called adaptive Gaussian thresholding is used. Images with irregular light or pixel intensity can be processed using adaptive Gaussian thresholding. Because it uses an odd integer for the block size, the thresholding value for each block can vary based on the irregularities in the block. To get the desired result, the value of constant C is subtracted from the mean or weighted mean from each region (block). The resultant image of Gaussian adaptive thresholding is shown in Figure 3.5 (c).



(a)



(b)

(c)

Figure 3.5: (a) Input image (b) Image after applying global thresholding (c) Image after applying adaptive Gaussian thresholding.

Line Extraction

The line-finding algorithm is one of the few parts of Tesseract that has previously been published [43]. In various cases, uneven and curvy lines make it difficult for the OCR engine to find the correct line, thus producing errors in results. So, it's important to extract lines from the document. To do this, dilation is applied so that each line becomes a connected component as shown in Figure 3.6. Afterwards, contour finding is applied and a tight bounding box is fitted based on the contours as depicted in Figure 3.6. Each bounding box is considered as a line. A pseudocode of the procedure is given in table 3.5.

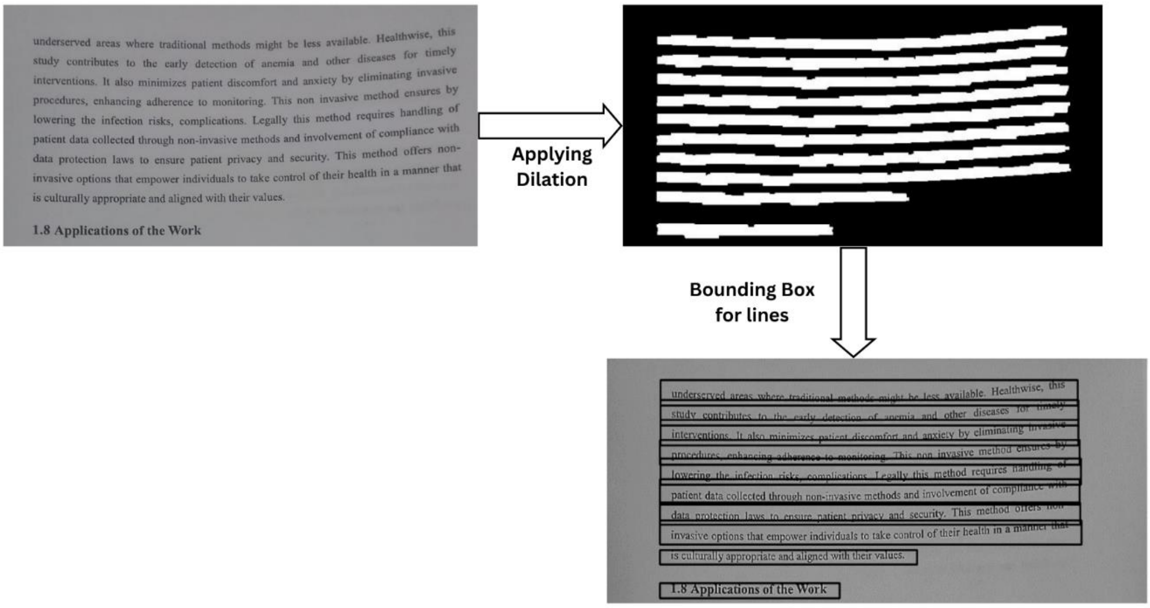


Figure 3.6: Image after performing dilation and bounding box for each line.

Table 3.5: Pseudocode for finding each line

Algorithm 5 Line Finding Algorithm

```
Initialize a structuring element of size (4, 40).
Apply morphological dilation.
Apply contour finding.
Initialize a list.
for each contour do
    Find the perimeter of the contour.
    if perimeter is inside the specified range then
        Append the boundary of the contour to the list.
    end if
end for
return the list.
```

(a)

(b)

Figure 3.7: (a) Input Image (b) Image after clearing border.

Border Clearing

The bounding boxes that have been used to find each line may contain some words from the previous and afterward lines because of uneven and curvy words. To remove these partially occupied lines, morphological reconstruction is applied. Marker is obtained along the boundary of ROI. Then, dilation is applied using the kernel until there is no new pixel is added. Finally, these resultant images are subtracted from original one to get one single line in the region of interest (ROI). A pseudocode of the procedure is given in table 3.6.

Table 3.6: Pseudocode for boundary cleaning

Algorithm 6 Boundary Cleaning Algorithm

Initialize a numpy array of zeros as the marker image.
for each pixel along the mask image border **do**
    Append marker image pixels.
end for
Initialize a 3×3 size structuring element for dilation.
while new pixels are added **do**
    Continuously dilate the marker image.
end while
Subtract the marker image from the mask image.

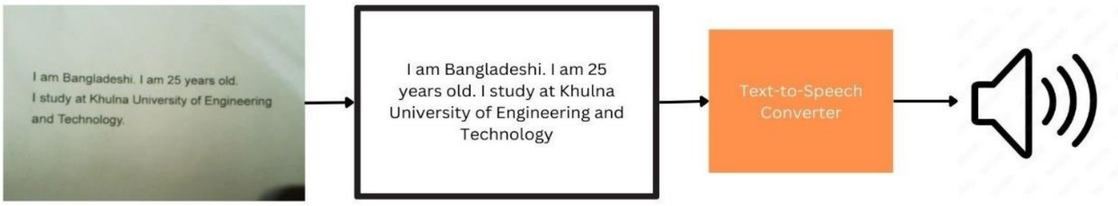


Figure 3.8: Text extraction from image and sound generation.

Character Recognition and Text-to-speech Generation

The Tesseract OCR engine is then fed the pre-processed image as a result, allowing it to predict the characters line by line. Subsequently, the document's lines are concatenated to form a single string. There's still space for improvement because some words have been misspelled. The output string is further corrected using a Textblob spell checker that has been pre-trained to identify all misspelled terms in order to increase the accuracy of the results. Prior to that, the final string was cleaned by using regex to separate each word and remove any special characters that were meaningless resulted from the undesired noises in the image The Textblob spell checker offers features for correcting misspelled words as well. The user can then hear an audible sound produced by the Google text-to-speech library thanks to the updated string being provided to it. Overall procedure is depicted in Figure 3.8.

3.3 Conclusion

To sum up, the proposed methodology offers a thorough approach for fall detection, obstacle avoidance, and optical character recognition that is specially designed for those with visual impairments. Pseudocode and schematic diagrams needed for hardware setup are included. Furthermore, the necessary picture preprocessing methods for optical character recognition (OCR) are described in detail. It is anticipated that the suggested approach will greatly improve the mobility and reading comprehension of those with visual impairments, leading to increased freedom.

Chapter 4

Implementation, Results and Discussions

4.1 Introduction

This chapter presents the findings from all of the methodologies used during the course of this thesis work. The chapter is broken down into several sections and subsections. The remainder of this chapter is structured as follows: Section 4.2 provides information about the utilized hardware and software resources. Hardware prototype design is shown in section 4.3. The dataset utilized is described in section 4.4. The numerous evaluation metrics are represented in section 4.5. Finally, the experimental results are shown in section 4.6.

4.2 Experimental Setup

The Experimental Setup section discusses useful information on the hardware specifications and software information such as toolkit, OS details, etc. used in for the experiment.

4.2.1 Hardware

Many pieces of hardware are required for the project to be completed successfully. Without the proper hardware, the software may not function properly or at all. Below is some useful information regarding the hardware utilized in the database generation and subsequent processing and detection operations.

➢ Host Computer

- Processor: Intel(R) Core (TM) i5-8250U CPU
- RAM: 8GB.
- ROM: 128GB SSD.
- GPU: Intel® UHD Graphics 620

- Raspberry Pi
    - Processor: Quad core Cortex-A72 (ARM v8) 64-bit
    - RAM: 4GB

4.2.2 Software

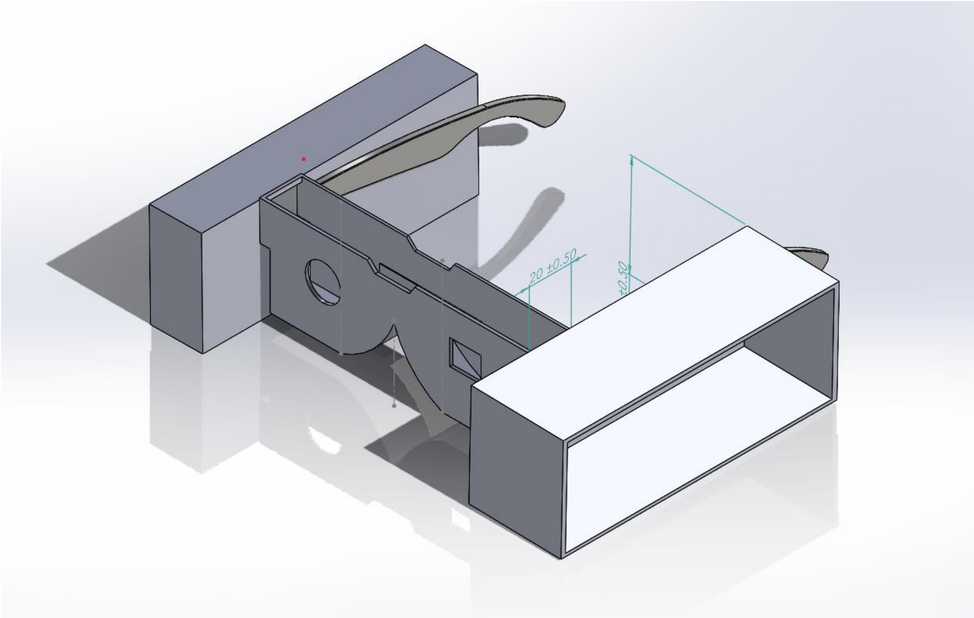
Various types of software are used to make the preprocessing, training and testing steps trouble-free. Useful information regarding the used OS information, programming language information, toolkit versions, etc. are provided below.

- Host Computer
    - Operating System: Windows 10.
    - Programming Language: Python 3.11.
    - Application type: Python Script.
    - IDE: Spyder.
    - Libraries: Scikit-learn v0.20.0, OpenCV v3.4.2, Pandas v0.23.4, Numpy v1.15.3, Seaborn v0.9.0, Matplotlib 3.8.3, pytesseract.

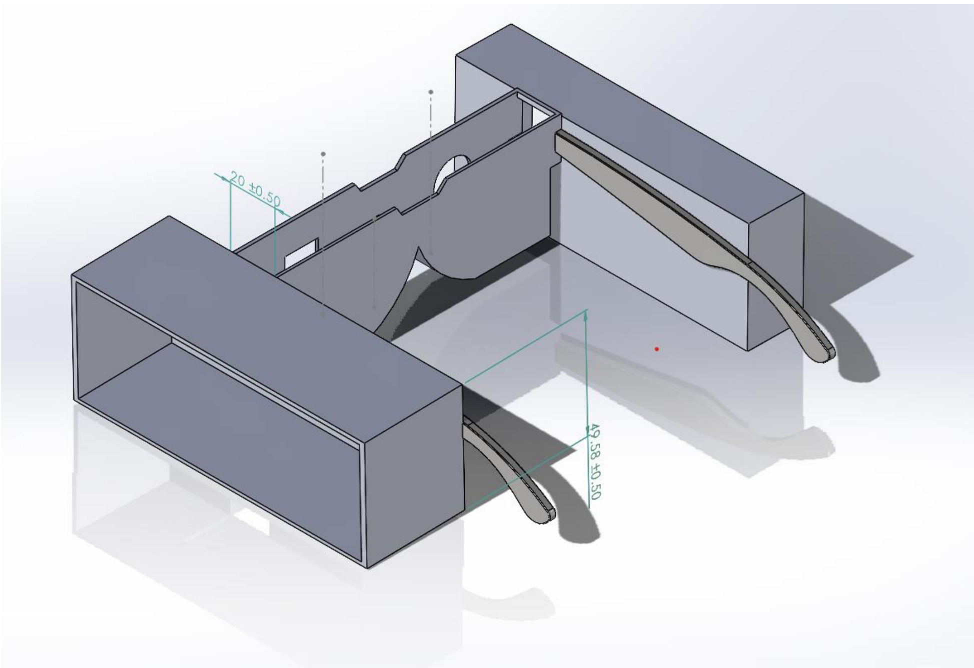
- Raspberry Pi
    - Operating System: Debian version: 12 (bookworm)
    - Programming Language: Python 3.11.
    - Application type: Python Script.
    - IDE: Genie
    - Libraries: Scikit-learn v0.20.0, OpenCV v3.4.2, Pandas v0.23.4, Numpy v1.15.3, Seaborn v0.9.0, Matplotlib 3.8.3, pytesseract.

4.3 Hardware Prototype Design

A 3D model for the hardware prototype is designed using solidworks software. All the measurements that require are carefully extracted beforehand to make the system light-weight, lucrative, smart and not bulky. A 3d modeling view of the prototype is shown in Figure 4.1.



(a) Isometric Front View



(b) Isometric Back View

Figure 4.1: 3D model for the prototype

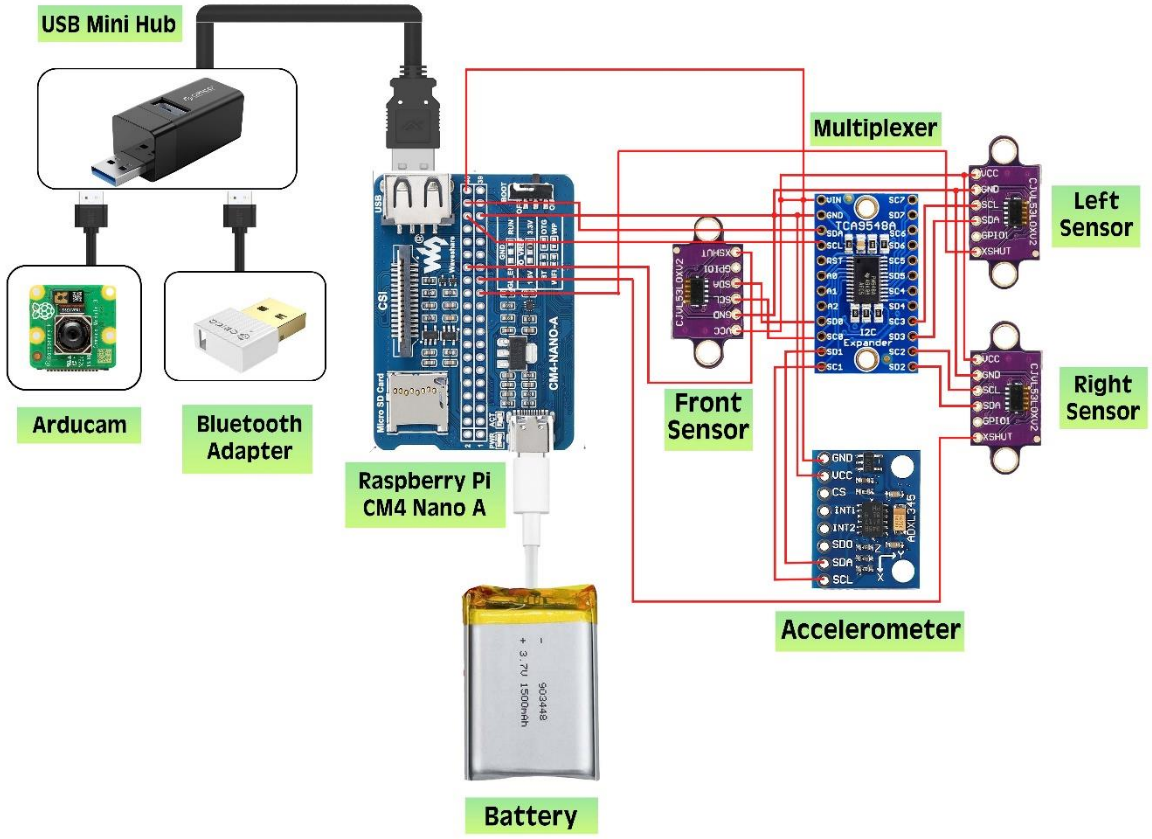
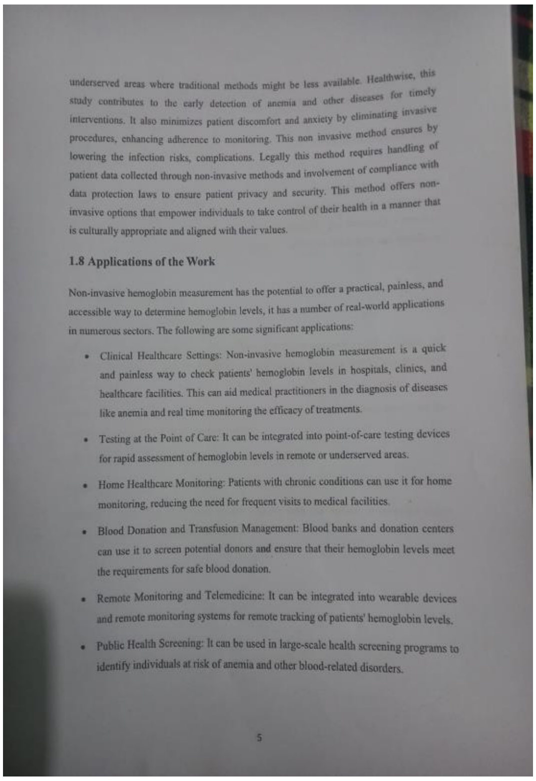


Figure 4.2: Circuit Diagram of the system.

The hardware circuit for the system includes a raspberry pi CM4 as processor, three VL53L0X time-of-flight distance laser sensor, one ADXL345 Accelerometer, one 8 mp arducam, one battery and necessary wires. The overall circuit diagram is depicted in Figure 4.2.

4.4 Dataset

One of the goals of this thesis is to optical character recognition for visually impaired people. For this purpose, a dataset of total 30 English typed documents are created. Documents are of various size of fonts. Further, the documents are augmented using rotation and brightness decreasing to make more variation in dataset. An example of used dataset is shown in Figure 4.3.



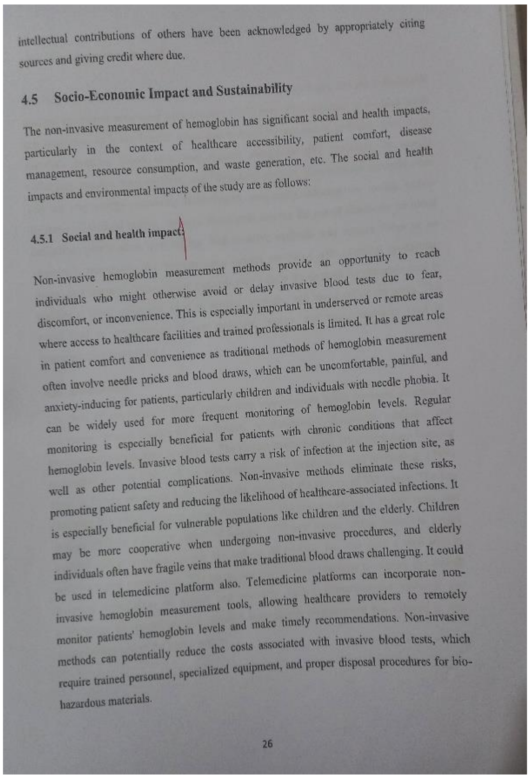
underserved areas where traditional methods might be less available. Healthwise, this study contributes to the early detection of anemia and other diseases for timely interventions. It also minimizes patient discomfort and anxiety by eliminating invasive procedures, enhancing adherence to monitoring. This non invasive method ensures by lowering the infection risks, complications. Legally this method requires handling of patient data collected through non-invasive methods and involvement of compliance with data protection laws to ensure patient privacy and security. This method offers non-invasive options that empower individuals to take control of their health in a manner that is culturally appropriate and aligned with their values.

1.8 Applications of the Work

Non-invasive hemoglobin measurement has the potential to offer a practical, painless, and accessible way to determine hemoglobin levels, it has a number of real-world applications in numerous sectors. The following are some significant applications:

- Clinical Healthcare Settings: Non-invasive hemoglobin measurement is a quick and painless way to check patients' hemoglobin levels in hospitals, clinics, and healthcare facilities. This can aid medical practitioners in the diagnosis of diseases like anemia and real time monitoring the efficacy of treatments.
- Testing at the Point of Care: It can be integrated into point-of-care testing devices for rapid assessment of hemoglobin levels in remote or underserved areas.
- Home Healthcare Monitoring: Patients with chronic conditions can use it for home monitoring, reducing the need for frequent visits to medical facilities.
- Blood Donation and Transfusion Management: Blood banks and donation centers can use it to screen potential donors and ensure that their hemoglobin levels meet the requirements for safe blood donation.
- Remote Monitoring and Telemedicine: It can be integrated into wearable devices and remote monitoring systems for remote tracking of patients' hemoglobin levels.
- Public Health Screening: It can be used in large-scale health screening programs to identify individuals at risk of anemia and other blood-related disorders.

5



intellectual contributions of others have been acknowledged by appropriately citing sources and giving credit where due.

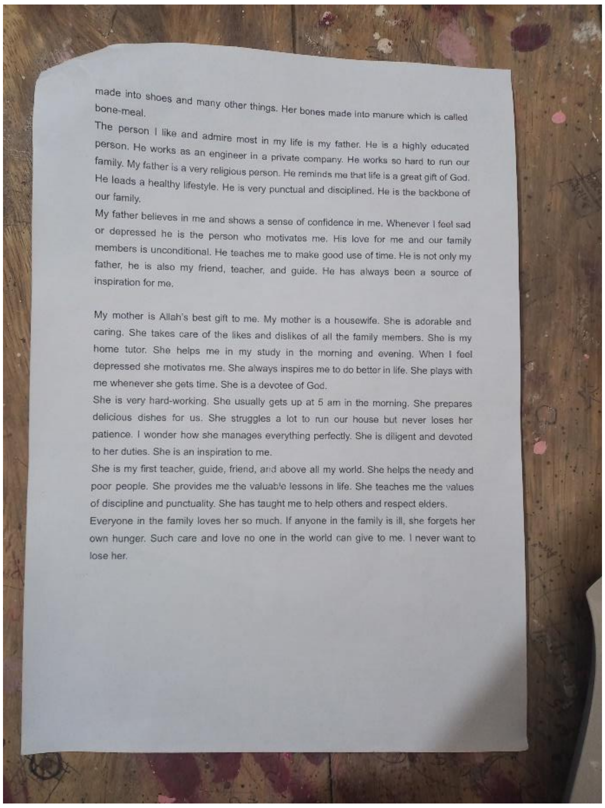
4.5 Socio-Economic Impact and Sustainability

The non-invasive measurement of hemoglobin has significant social and health impacts, particularly in the context of healthcare accessibility, patient comfort, disease management, resource consumption, and waste generation, etc. The social and health impacts and environmental impacts of the study are as follows:

4.5.1 Social and health impact

Non-invasive hemoglobin measurement methods provide an opportunity to reach individuals who might otherwise avoid or delay invasive blood tests due to fear, discomfort, or inconvenience. This is especially important in underserved or remote areas where access to healthcare facilities and trained professionals is limited. It has a great role in patient comfort and convenience as traditional methods of hemoglobin measurement often involve needle pricks and blood draws, which can be uncomfortable, painful, and anxiety-inducing for patients, particularly children and individuals with needle phobia. It can be widely used for more frequent monitoring of hemoglobin levels. Regular monitoring is especially beneficial for patients with chronic conditions that affect hemoglobin levels. Invasive blood tests carry a risk of infection at the injection site, as well as other potential complications. Non-invasive methods eliminate these risks, promoting patient safety and reducing the likelihood of healthcare-associated infections. It is especially beneficial for vulnerable populations like children and the elderly. Children may be more cooperative when undergoing non-invasive procedures, and elderly individuals often have fragile veins that make traditional blood draws challenging. It could be used in telemedicine platform also. Telemedicine platforms can incorporate non-invasive hemoglobin measurement tools, allowing healthcare providers to remotely monitor patients' hemoglobin levels and make timely recommendations. Non-invasive methods can potentially reduce the costs associated with invasive blood tests, which require trained personnel, specialized equipment, and proper disposal procedures for bio-hazardous materials.

26



made into shoes and many other things. Her bones made into manure which is called bone-meal.

The person I like and admire most in my life is my father. He is a highly educated person. He works as an engineer in a private company. He works so hard to run our family. My father is a very religious person. He reminds me that life is a great gift of God. He leads a healthy lifestyle. He is very punctual and disciplined. He is the backbone of our family.

My father believes in me and shows a sense of confidence in me. Whenever I feel sad or depressed he is the person who motivates me. His love for me and our family members is unconditional. He teaches me to make good use of time. He is not only my father, he is also my friend, teacher, and guide. He has always been a source of inspiration for me.

My mother is Allah's best gift to me. My mother is a housewife. She is adorable and caring. She takes care of the likes and dislikes of all the family members. She is my home tutor. She helps me in my study in the morning and evening. When I feel depressed she motivates me. She always inspires me to do better in life. She plays with me whenever she gets time. She is a devotee of God.

She is very hard-working. She usually gets up at 5 am in the morning. She prepares delicious dishes for us. She struggles a lot to run our house but never loses her patience. I wonder how she manages everything perfectly. She is diligent and devoted to her duties. She is an inspiration to me.

She is my first teacher, guide, friend, and above all my world. She helps the needy and poor people. She provides me the valuable lessons in life. She teaches me the values of discipline and punctuality. She has taught me to help others and respect elders.

Everyone in the family loves her so much. If anyone in the family is ill, she forgets her own hunger. Such care and love no one in the world can give to me. I never want to lose her.

Figure 4.3: Datasets used for experiment.

4.5 Evaluation Metrics

4.5.1 Evaluation Metrics of Obstacle Avoidance

In total 4 types of metrices are used for the evaluation of distance measurement performance: (1) Accuracy (2) Error rate and (3) Standard Deviation.

1.Accuracy: For a reference measurement, a total of 8 measurements and their average value is taken. Accuracy is measured by Equation (4.1).

$$Accuracy = \frac{Average\ of\ measured\ values}{Reference\ value} \times 100 \text{ .... (4.1)}$$

2. Error rate: Error rate can be calculated from accuracy by using Equation (4.2).

$$Error\ rate = 100 - Accuracy \text{..................... (4.2)}$$

3. Standard Deviation: Standard Deviation shows the variation of measured data from the mean value of measured data. Standard deviation can be calculated by Equation (4.3).

$$Standard\ Deviation = \frac{\sum(\bar{x}-x)}{n} \text{.......................... (4.3)}$$

Where,

$\bar{x}$ = mean value of measured distances

x = measured distance

n = total number of measures

4.5.2 Evaluation Metrics of OCR

Mainly two types of metric are used for evaluation of optical character recognition (OCR) performance: (1) Character Metric and (2) Word Metric. They are described in the following subsections.

A. Character Metric

1. Character Error Rate: CER calculation is based on the concept of Levenshtein distance, where we count the minimum number of character-level operations required to transform the ground truth text into the OCR output. For example:

- *Ground truth* text: 619375128
- *OCR output* text: 61g375Z8

Transformations required to transform OCR output into the ground truth are,

1. g instead of 9
2. Missing 1
3. Z instead of 2

Character error rate (CER) is defined by the Equation (4.4).

$$CER = \frac{T}{T+C} \times 100\% \quad \ldots\ldots\ldots\ldots\ldots \text{(4.4)}$$

Where,

- T is the total number of transforms
- C is the total number of correct characters (C)

2. Character Precision: Character precision are computed by Equation (4.5).

$$Precision = \frac{C}{O} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{. (4.5)}$$

Where,

- C is the total number of correct characters
- O is the total number of characters in the output

3. Character Recall: Character recall are computed by equation (4.6).

$$Recall = \frac{C}{R} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.6)}$$

Where,

- C is the total number of correct characters
- R is the total number of characters in the reference

B. Word Metric

Word Error Rate: Word error rate (WER) is calculated by dividing the number of errors (counting the number of insertions, deletions, and substitutions) between reference and predicted word divided by the total numbers of words in the reference transcription.

The math formula for calculating WER is given in Equation (4.7).

$$WER = \frac{S+D+I}{N} \text{ .... (4.7)}$$

Where,

- S (Substitution): is the total of words replaced in the output. For example, if the reference word is “cat” and the recognized word is “bat”, then there is one substitution error.
- I (Insertion): is the total of words added in the output. For example, if the reference transcription is “the book is on **the** table” and the recognized phrase is “the book is on table”, then there is one insertion error.
- D (Deletions): is the total of words omitted in the output. For example, if the reference transcription is “the book is on the table” and the recognized phrase is “the **ugly** book is on the table”, then there is one deletion error.
- N (Total Words): is the total number of words in the reference document.

2. Word Precision: Word precision are computed by Equation (4.8).

$$Precision = \frac{W}{O} \text{ .... (4.8)}$$

Where,

- W is the total number of correct words
- O is the total number of words in the output

3. Word Recall: Word recall are computed by the Equation (4.9).

$$Precision = \frac{W}{R} \text{ .... (4.9)}$$

Where,

- W is the total number of correct words
- R is the total number of words in the reference

4.6 Experimental Results

4.6.1 Obstacle Avoidance

The performance of the proposed system is evaluated in the domain of object detection and in both indoor and outdoor environments. In both environment, objects are placed at six different regions for detecting them by the laser sensors. The regions are

- Region 1: (0–30 cm)
- Region 2: (31–60 cm)
- Region 3: (61–100 cm)
- Region 4: (101–150 cm)
- Region 5: (151–180 cm)
- Region 6: (181–222 cm)

To detect objects, we have observed and collected the data from the laser sensors. The values of measured distance of the indoor environment are illustrated in Table 4.1. For each region, we have taken the measured values and made an average value from those values. These mean values of the six regions are taken for further performance evolution. The performance of the indoor object detection based on the measured distances from the indoor environment is presented in Table 4.2. The accuracy, error rate, standard deviation, and variance of the indoor environment are calculated from the average values of Table 4.1.

Table 4.1: Distance Measurements at Indoor Environment

Actual Distance (cm)	**Measured Distance (cm)**								**Average (cm)**
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
30.00	30.00	30.00	30.00	31.00	28.00	30.00	30.00	30.00	29.88
60.00	59.00	60.00	60.00	60.00	60.00	59.00	61.00	60.00	59.88
100.00	100.00	100.00	101.00	99.00	99.00	101.00	100.00	101.00	100.13
150.00	150.00	148.00	150.00	150.00	151.00	150.00	150.00	150.00	149.88
180.00	178.00	176.00	179.00	178.00	179.00	181.00	179.00	180.00	178.75
220.00	217.00	218.00	218.00	219.00	219.00	218.00	218.00	219.00	218.25

Table 4.2: Object Detection Performance of Indoor Environment

Actual Distance (cm)	Average (cm)	Accuracy (%)	Error rate (%)	Standard Deviation	Variance
30.00	29.88	99.60	0.40	0.78	0.61
60.00	59.88	99.80	0.20	0.60	0.36
100.00	100.13	99.87	0.13	0.78	0.61
150.00	149.88	99.02	0.08	0.78	0.61
180.00	178.75	99.31	0.69	1.39	1.94
220.00	218.25	99.20	0.80	0.66	0.44

The measured distances of each region in the outdoor environment are illustrated in Table 4.3. We calculated the average value from the samples of each region. The performance of the object detection in the outdoor environment is shown in Table 4.4. which represents the accuracy, error rate, standard deviation, and variance of the object detection module for the outdoor environment.

Table 4.3: Distance Measurements at Outdoor Environment

Actual Distance (cm)	Measured Distance (cm)								Average (cm)
	1	2	3	4	5	6	7	8	
30.00	29.00	30.00	28.00	31.00	29.00	30.00	29.00	30.00	29.50
60.00	58.00	59.00	58.00	59.00	61.00	59.00	62.00	58.00	59.25
100.00	100.00	101.00	101.00	103.00	99.00	98.00	100.00	98.00	100.00
150.00	152.00	147.00	148.00	152.00	150.00	148.00	149.00	148.00	149.25
180.00	176.00	181.00	184.00	177.00	179.00	183.00	182.00	181.00	180.38
220.00	219.00	222.00	224.00	223.00	218.00	221.00	221.00	222.00	221.25

Table 4.4: Object Detection Performance of Outdoor Environment

Actual Distance (cm)	Average (cm)	Accuracy (%)	Error rate (%)	Standard Deviation	Variance
30.00	29.50	98.33	1.67	0.87	0.75
60.00	59.25	98.75	1.25	1.39	1.94
100.00	100.00	100.00	0.00	1.58	2.50
150.00	149.25	99.50	0.50	1.79	3.19
180.00	180.38	99.79	0.21	2.64	6.98
220.00	221.25	99.43	0.57	1.85	3.44

the comparison between the indoor and outdoor environment about the measured distance is depicted in Figure 4.1. From the figure, it can be observed that the measured distances of each region are so close in both environments. In regions 5 and 6, the difference between the two environments is 1.63 and 3 cm respectively.

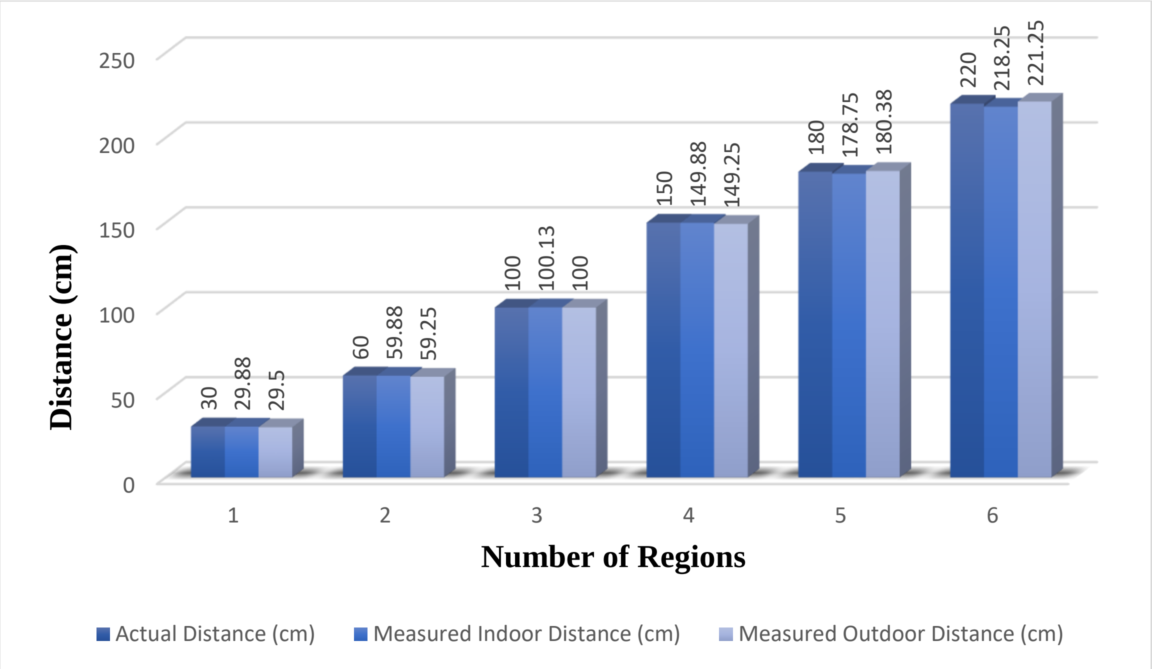


Figure 4.4: Comparison of measured distance between indoor and outdoor environment.

4.6.2 Optical Character Recognition

A total of 25 images are used for testing purpose. Total character count in testing documents is 21852, out of which 20367 were identified correctly. Table 4.5 and Table 4.6 shows the statistics of experiment for the character metrices. Overall character error rate (CER) is 6.97% with precision, recall and F1-score are 94.45%, 96.07% and 95.25% respectively.

Table 4.5: Character counter summary

Correct Characters	Transformed Characters	Total Characters
20367	1485	21852

Table 4.6: Performance evaluation of character recognition

CER	Accuracy	Precision	Recall	F1-score
6.97	93.03	94.45	96.07	95.25

Table 4.7: Word count summary

Correct Words	Substitution	Insertion	Deletion	Total Words
4058	319	23	82	4482

Table 4.8: Performance evaluation of word recognition

WER	Accuracy	Precision	Recall	F1-score
9.46	90.54	91.62	94.22	92.90

4.7 Conclusion

In conclusion, the experimental results offer valuable insights into the performance and effectiveness of the proposed methodology. Within this methodology, objects are detected using laser sensors positioned in the front, left, and right directions. The results clearly demonstrate the efficacy and accuracy of the object detection module. Furthermore, the proposed methodology functions as a real-time smart reader, providing crucial assistance to visually impaired individuals with remarkable accuracy and precision.

Chapter 5

Societal, Health, Environment, Safety, Ethical, Legal and Cultural Issues

5.1 Intellectual Property Considerations

In our work, we address intellectual property issues beyond developing novel methodologies to develop the system. Methods for preparing data, image pre-processing techniques, a lightweight and portable prototype construction and optimization strategies specifically created for the aid of visually impaired comprise significant intellectual property. To protect these assets, thorough records and documentation have been maintained throughout the study process. We aim to safeguard the innovative aspects of our work and facilitate its application in other research components, by paying anticipatory attention to these intellectual property issues. These datasets may contain carefully chosen image collections intended for real-time optical character recognition. Intellectual property also includes the development of software tools, scripts, or libraries to support model testing, training, and evaluation. These assets are protected by appropriate measures like version control, access restrictions, and usage guide documentation. Moreover, because the research is interdisciplinary, collaborating with experts from different fields may yield fresh perspectives or ideas that lead to cooperative agreements about intellectual property ownership or licensing. Our aim is to ensure proper and ethical practices in intellectual property management while optimizing the value and impact of our research products. We fully recognize and manage these intellectual property issues in order to achieve this.

5.2 Ethical Considerations

Our approach to develop an obstacle avoidance and smart reading system for visually impaired is heavily influenced by ethical considerations. Our primary concern is the ethical treatment of data, and we ensure that all data collection and usage adhere to the principles of informed consent, confidentiality, and privacy protection. To preserve people's privacy, this means obtaining consent before using any sensitive or personally identifiable information and, whenever possible, anonymizing data. To further promote justice and fairness in the study's findings, we are committed to reducing any possible biases in our datasets and algorithms. Additionally, we uphold ethical standards in the manner in which we conduct our research, ensuring transparency, veracity, and honesty in all aspects of the study as well as equity and respect in all of our endeavors. Last but not least, by establishing appropriate safety protocols and ethical standards, we prioritize the health and safety of everyone involved in our research, including participants and collaborators. To put it briefly, we want to pursue our research in an ethical and socially responsible manner while upholding the principles of fairness, integrity, and respect in whatever we do.

5.3 Safety Considerations

Safety is a top priority in our work. Primarily, we safeguard the security of both participants and researchers by implementing stringent safety protocols throughout experimental processes, encompassing the handling of software, hardware and other equipment. This means making sure that all relevant safety laws and regulations are observed as well as providing researchers with the necessary training on how to use equipment properly. In addition, we use security measures to lessen the risks associated with data handling and storage. Examples of these measures include protections against unauthorized access to or disclosure of private or sensitive data. In all human studies and experiments, we put the safety and well-being of the participants first, making sure that their rights and dignity are upheld at every stage of the investigation. In addition, we keep emergency protocols and backup plans in place to handle any unanticipated circumstances or mishaps that might arise throughout the study. By taking proactive measures to address safety concerns, we hope to establish a safe and supportive environment in which to carry out our study, ensuring the wellbeing of all parties concerned.

5.4 Legal Considerations

We consider compliance with data protection and intellectual property rules, as well as possible liability and risk management issues, in our thesis. We assess the possible legal consequences related to the use of software and hardware protocols for experiments, or methods for processing data. To minimize any legal conflicts or objections, we keep records of all of thesis operations and guarantee transparency in reporting procedures and outcomes. This entails determining and minimizing any risks related to model biases, algorithmic mistakes, or unexpected repercussions of the study's findings. By proactively addressing legal problems, we hope to reduce legal obligations, ensure accountability for our actions, and conduct our research in a responsible and ethical manner.

5.5 Impact of the Project on Societal and Health Issues

The impact of this research on Societal, Health, and Cultural Issues is discussed below:

Societal impact:

- Contributes to advancements in technology, fostering innovation and progress in image processing and computer vision fields.
- Potential application to improve the mobility and reading capability of the visually impaired and make them more independent.

Health Impact:

- Users can be alert of any upcoming obstacles and avoid any fatal accident.
- In case of any outdoor accidents like fall care-taker of the user can be notified immediately.
- Reading can be beneficial for mental health for the visually impaired.

5.6 Impact of Project on the Environment and Sustainability

Here's the impact of the thesis on the environment and sustainability mentioned below:

- **Environmental Impact**:
  - **Reduction in energy consumption**: By using raspberry pi cm4 (nano version) our thesis reduces the battery power consumption and also provides lower energy consumption in computing systems.
  - **Minimization of waste**: Our thesis presents a solution that utilizes a low-power-consuming processor and rechargeable battery, effectively reducing the generation of waste and promoting sustainability.

- **Sustainability Impact**:
    - **Sustainable resource management**: By extending the lifespan of hardware resources and processor, this thesis promotes sustainable resource management.
    - **Encouraging eco-friendly practices**: By employing energy-efficient components and sustainable materials, this thesis aligns with environmentally conscious principles, promoting a greener approach to technology for the benefit of both users and the planet. This emphasis on eco-friendly design reflects our commitment to sustainable innovation and responsible use of resources.

Chapter 6

Addressing Complex Engineering Problems and Activities

6.1 Complex engineering problems associated with the current thesis

Complex engineering problems associated with the current thesis are as follow

1. **Wearable lightweight device construction:** To construct the prototype it requires various hardware components like sensors, cameras, etc. along with the microprocessor. It is quite challenging to construct a user friendly, lightweight and portable system that is not bulky.

2. **Real-time data acquisition:** Various sensors are employed to gather real-time data from the environment. However, obtaining accurate and precise real-time data can pose challenges due to various sources of noise and other factors. Capturing images in real-time may not always meet expectations, as images can frequently suffer from blurriness, unwanted illumination, and persistent noise that is difficult to remove.

3**. Generalization and robustness**: Image pre-processing techniques for optical character recognition (OCR) can vary significantly from one image to another. Developing a universal and robust procedure that applies effectively to a large volume of images poses a considerable challenge. This variability arises due to diverse factors such as image quality, lighting conditions, noise levels, and variations in text styles and sizes. As a result, designing an adaptable pre-processing pipeline capable of handling this variability is a complex and ongoing task in the field of OCR research.

4. **Handling complex scene geometries**: Designing algorithms that can precisely handle intricate object motions and scene geometry, which may include complicated interactions, occlusions, or reflection.

5. **Hardware and Software co-ladder:** The integration of developed software systems can often pose challenges due to mismatches in processor configurations. In real-time

scenarios, accuracy can frequently diminish as a result of factors such as noisy data, sluggish processors, and other related issues. These challenges highlight the importance of thorough testing, optimization, and adaptation to ensure seamless integration and sustained performance of the software system in various environments.

6.2 Complex engineering activities associated with the current thesis

1. **Prototype Design:** Development of circuits and their corresponding control programs in an efficient and optimized manner. The design process also includes utilizing 3D printing technology and meticulous planning for the placement of hardware components. The goal is to create a lightweight, user-friendly system that prioritizes accuracy and efficiency. Through careful consideration of hardware layout and selection of materials, the aim is to achieve a well-balanced system that meets the needs of visually impaired users while optimizing performance and usability.

2. **Data preprocessing and augmentation:** Implementing complex data preprocessing and augmentation techniques to enhance the quality and diversity of datasets, including image normalization, augmentation, and synthetic data generation to improve model generalization and robustness.

3. **Performance evaluation and analysis**: Conduct comprehensive performance evaluation and analysis of including quantitative metrics such as accuracy, precision, recall and F1-score, as well as qualitative analysis of visual outputs and error analysis to identify the procedure's strengths, weaknesses, and areas for improvement.

4. **Deployment and integration**: Integrating the model into microprocessor or any software systems, involving tasks such as model deployment, API development, compatibility testing, and system integration to ensure seamless operation and interoperability with other components.

5. **Continuous improvement and optimization**: Iteratively refining and optimizing image pre-processing techniques for optical character recognition (OCR) and functionalities of the sensors based on feedback from performance evaluation and user feedback, to achieve incremental improvements in performance and usability over time.

Chapter 7

Conclusions

7.1 Summary

In this thesis, we have taken a comprehensive approach to developing a solution tailored for individuals with visual impairments, incorporating modules for obstacle avoidance, fall detection, and optical character recognition (OCR). The obstacle avoidance module utilizes VL53L0X time-of-flight sensors managed by a TCA9548A multiplexer known for its high accuracy and robustness to noise. Through experimentation and iterative refinement, we carefully selected the appropriate threshold for obstacle detection distance. An additional feature of the obstacle avoidance module is its ability to alert users through text-to-speech when obstacles are detected. Fall detection, crucial for the safety of visually impaired and elderly individuals, relies on an ADXL345 accelerometer. Upon detecting a fall, the system promptly notifies guardians. By setting precise thresholds and conditions for what constitutes a fall event, the system aims to minimize instances where false alarms are triggered. Through these measures, the system enhances its accuracy in identifying true fall events while reducing the likelihood of false alerts, thereby ensuring the reliability and effectiveness of the fall detection module for the safety of visually impaired and elderly individuals. The OCR module, essential for improving reading capabilities, undergoes a series of image preprocessing steps including perspective distortion removal, skew correction, and adaptive thresholding to ensure precise character recognition. The experimental results demonstrate the efficacy of our method to significantly enhance the mobility, safety, and reading comprehension of its users, promoting greater independence and accessibility.

7.2 Limitations

While our thesis presents promising results, several limitations are needed to be addressed. Some of them are as follows.

1. Kernel manipulation during the pre-processing section may not fully represent all types real-world document.
2. Challenges with image capturing, such as noisy or blurry images, can impair the real-time performance of optical character recognition.
3. Laser sensors can sometimes function slowly due to internal voltage drop.
4. Despite taking necessary measures to minimize the occurrence of false positives in fall detection, they may still occur.

7.3 Future Works

There are some areas where the works can be extended, some of them are as follows.

1. Developing an adaptive dilation kernel that can automatically detect the line gap to choose appropriate structuring elements.
2. Optical character recognition can be expanded to include the extraction of text from diverse domains, such as reading information from medicine boxes.

REFERENCES

[1] H.-A. Shahriari, S. Izadi, M.-R. Rouhani, F. Ghasemzadeh, and A.-R. Maleki, “Prevalence and causes of visual impairment and blindness in Sistan-va-Baluchestan Province, Iran: Zahedan Eye Study,” *British Journal of Ophthalmology*, vol. 91, no. 5, pp. 579–584, May 2007, doi: 10.1136/bjo.2006.105734.

[2] R. R. A. Bourne *et al.*, “Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The Right to Sight: An analysis for the Global Burden of Disease Study,” *Lancet Glob Health*, vol. 9, no. 2, pp. e144–e160, Feb. 2021, doi: 10.1016/S2214-109X(20)30489-7.

[3] K. F. Tabbara, H. F. El-Sheikh, and S. S. Shawaf, “Pattern of childhood blindness at a referral center in Saudi Arabia,” *Ann Saudi Med*, vol. 25, no. 1, pp. 18–21, Jan. 2005, doi: 10.5144/0256-4947.2005.18.

[4] B. Thylefors and A. D. Négrel, “The global impact of glaucoma.,” *Bull World Health Organ*, vol. 72, no. 3, pp. 323–6, 1994.

[5] G. Brian and H. Taylor, “Cataract blindness--challenges for the 21st century.,” *Bull World Health Organ*, vol. 79, no. 3, pp. 249–56, 2001.

[6] J. L. Leasher *et al.*, “Global Estimates on the Number of People Blind or Visually Impaired by Diabetic Retinopathy: A Meta-analysis From 1990 to 2010.,” *Diabetes Care*, vol. 39, no. 9, pp. 1643–9, Sep. 2016, doi: 10.2337/dc15-2171.

[7] Q. Lu, “Feasibility Study of a ‘Smart’ Aid for the Visually Impaired and Blind’s Independent Mobility in Outdoor Environments,” 2018. [Online]. Available: https://api.semanticscholar.org/CorpusID:52243960

[8] R. S. Mulky, S. Koganti, S. Shahi, and K. Liu, “Autonomous Scooter Navigation for People with Mobility Challenges,” in *2018 IEEE International Conference on Cognitive Computing (ICCC)*, IEEE, Jul. 2018, pp. 87–90. doi: 10.1109/ICCC.2018.00020.

[9] F. Shaikh, V. Kuvar, and Mohd. A. Meghani, “Ultrasonic sound based navigation and assistive system for visually impaired with real time location tracking and Panic button,” in *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, IEEE, Oct. 2017, pp. 172–175. doi: 10.1109/CESYS.2017.8321259.

[10] R. F. Olanrewaju, M. L. A. M. Radzi, and M. Rehab, “iWalk: Intelligent walking stick for visually impaired subjects,” in *2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)*, IEEE, Nov. 2017, pp. 1–4. doi: 10.1109/ICSIMA.2017.8312000.

[11] Z. Saquib, V. Murari, and S. N. Bhargav, “BlinDar: An invisible eye for the blind people making life easy for the blind with Internet of Things (IoT),” in *2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, IEEE, May 2017, pp. 71–75. doi: 10.1109/RTEICT.2017.8256560.

[12] A. Dastider, B. Basak, Md. Safayatullah, C. Shahnaz, and S. A. Fattah, “Cost efficient autonomous navigation system (e-cane) for visually impaired human beings,” in *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, IEEE, Dec. 2017, pp. 650–653. doi: 10.1109/R10-HTC.2017.8289043.

[13] S. Rao and V. M. Singh, “Computer Vision and Iot Based Smart System for Visually Impaired People,” in *2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, Jan. 2021, pp. 552–556. doi: 10.1109/Confluence51648.2021.9377120.

[14] S. Choudhary, V. Bhatia, and K. R. Ramkumar, “IoT Based Navigation System for Visually Impaired People,” in *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, IEEE, Jun. 2020, pp. 521–525. doi: 10.1109/ICRITO48877.2020.9197857.

[15] A. G. Sareeka, K. Kirthika, M. R. Gowthame, and V. Sucharitha, “pseudoEye — Mobility assistance for visually impaired using image recognition,” in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, IEEE, Jan. 2018, pp. 174–178. doi: 10.1109/ICISC.2018.8399059.

[16] Dr. Y. Perwej, S. A. Hannan, A. Asif, and A. Mane, “An Overview and Applications of Optical Character Recognition,” *International Journal of Advance Research In Science And Engineering (IJARSE)*, vol. Vol. 3, p. Pages 261-274, Feb. 2014.

[17] L. Eikvil, “Optical character recognition,” *citeseer. ist. psu. edu/142042. html*, vol. 26, 1993.

[18] A. Habib, Md. Islam, M. Kabir, M. Mredul, and M. Hasan, “Staircase Detection to Guide Visually Impaired People: A Hybrid Approach,” *Revue d’Intelligence Artificielle*, vol. 33, no. 5, pp. 327–334, Nov. 2019, doi: 10.18280/ria.330501.

[19] Md. M. Kamal, A. I. Bayazid, M. S. Sadi, Md. M. Islam, and N. Hasan, “Towards developing walking assistants for the visually impaired people,” in *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, IEEE, Dec. 2017, pp. 238–241. doi: 10.1109/R10-HTC.2017.8288947.

[20] N. S. Khan, S. Kundu, S. Al Ahsan, M. Sarker, and M. N. Islam, “An Assistive System of Walking for Visually Impaired,” in *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, IEEE, Feb. 2018, pp. 1–4. doi: 10.1109/IC4ME2.2018.8465669.

[21] E. Șipoș, C. Ciuciu, and L. Ivanciu, “Sensor-Based Prototype of a Smart Assistant for Visually Impaired People—Preliminary Results,” *Sensors*, vol. 22, no. 11, p. 4271, Jun. 2022, doi: 10.3390/s22114271.

[22] W.-J. Chang, L.-B. Chen, M.-C. Chen, J.-P. Su, C.-Y. Sie, and C.-H. Yang, “Design and Implementation of an Intelligent Assistive System for Visually Impaired People for Aerial Obstacle Avoidance and Fall Detection,” *IEEE Sens J*, vol. 20, no. 17, pp. 10199–10210, Sep. 2020, doi: 10.1109/JSEN.2020.2990609.

[23] T. Froneman, D. van den Heever, and K. Dellimore, “Development of a wearable support system to aid the visually impaired in independent mobilization and navigation,” in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, Jul. 2017, pp. 783–786. doi: 10.1109/EMBC.2017.8036941.

[24] Md. A. Rahman and M. S. Sadi, “IoT Enabled Automated Object Recognition for the Visually Impaired,” *Computer Methods and Programs in Biomedicine Update*, vol. 1, p. 100015, 2021, doi: 10.1016/j.cmpbup.2021.100015.

[25] M. F. Zamir, K. B. Khan, S. A. Khan, and E. Rehman, “Smart Reader for Visually Impaired People Based on Optical Character Recognition,” 2020, pp. 79–89. doi: 10.1007/978-981-15-5232-8_8.

[26] A. Goel, A. Sehrawat, A. Patil, P. Chougule, and S. Khatavkar, “Raspberry Pi Based Reader for Blind People,” pp. 2356–2395, Feb. 2022.

[27] G. Vaidya, K. Vaidya, and K. Bhosale, “Text Recognition System for Visually Impaired using Portable Camera,” in *2020 International Conference on Convergence to Digital World - Quo Vadis (ICCDW)*, IEEE, Feb. 2020, pp. 1–4. doi: 10.1109/ICCDW45521.2020.9318706.

[28] M. Rajesh *et al.*, “Text recognition and face detection aid for visually impaired person using Raspberry PI,” in *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*, IEEE, Apr. 2017, pp. 1–5. doi: 10.1109/ICCPCT.2017.8074355.

[29] H. Jiang, T. Gonnot, W.-J. Yi, and J. Saniie, “Computer vision and text recognition for assisting visually impaired people using Android smartphone,” in *2017 IEEE International Conference on Electro Information Technology (EIT)*, IEEE, May 2017, pp. 350–353. doi: 10.1109/EIT.2017.8053384.

[30] A. Vijayanarayanan, R. Savithiri, P. Lekha, and R. S. Abbirami, “Image Processing Based on Optical Character Recognition with Text to Speech for Visually Impaired,” *Journal of Science, Computing and Engineering Research*, vol. 6, no. 4, pp. 68–73, 2023.

[31] C. W. Zhao, J. Jegatheesan, and S. C. Loon, “Exploring iot application using raspberry pi,” *International Journal of Computer Networks and Applications*, vol. 2, no. 1, pp. 27–34, 2015.

[32] N. Lakovic, M. Brkic, B. Batinic, J. Bajic, V. Rajs, and N. Kulundzic, “Application of low-cost VL53L0X ToF sensor for robot environment detection,” in *2019 18th International Symposium INFOTEH-JAHORINA (INFOTEH)*, IEEE, Mar. 2019, pp. 1–4. doi: 10.1109/INFOTEH.2019.8717779.

[33] H. Chen and P. Li, “Intelligent Positioning Boots for Miners’ Fall Monitoring Based on ZigBee Communication Protocol and ADXL345 Acceleration Sensor,” in *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)*, IEEE, Apr. 2022, pp. 304–307. doi: 10.1109/IPEC54454.2022.9777577.

[34] B. Shi, X. Chen, Z. He, H. Sun, and R. Han, “Research on Gesture Recognition System Using Multiple Sensors Based on Earth’s Magnetic Field and 1D Convolution Neural Network,” *Applied Sciences*, vol. 13, no. 9, p. 5544, Apr. 2023, doi: 10.3390/app13095544.

[35] “TCA9548A I2C Multiplexer: ESP32, ESP8266, Arduino | Random Nerd Tutorials.” Accessed: Feb. 21, 2024. [Online]. Available: https://randomnerdtutorials.com/tca9548a-i2c-multiplexer-esp32-esp8266-arduino/

[36] M. Zhang, A. Joshi, R. Kadmawala, K. Dantu, S. Poduri, and G. S. Sukhatme, “Ocrdroid: A framework to digitize text using mobile phones,” in *International Conference on Mobile Computing, Applications, and Services*, 2009, pp. 273–292.

[37] C. Patel, A. Patel, and D. Patel, “Optical Character Recognition by Open source OCR Tool Tesseract: A Case Study,” *Int J Comput Appl*, vol. 55, no. 10, pp. 50–56, Oct. 2012, doi: 10.5120/8794-2784.

[38] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin, Heidelberg: Springer-Verlag, 2003.

[39] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, “Morphological reconstruction,” *Digital image processing using MATLAB, MathWorks*, 2010.

[40] J. Canny, “A Computational Approach to Edge Detection,” *IEEE Trans Pattern Anal Mach Intell*, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986, doi: 10.1109/TPAMI.1986.4767851.

[41] N. M. Wanas, D. A. Said, N. H. Hegazy, and N. M. Darwish, “A study of local and global thresholding techniques in text categorization,” *ausdm*, vol. 6, pp. 91–101, 2006.

[42] A. K. Chaubey, “Comparison of the local and global thresholding methods in image segmentation,” *World J Res Rev*, vol. 2, no. 1, pp. 1–4, 2016.

[43] R. Smith, “An Overview of the Tesseract OCR Engine,” in *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007) Vol 2*, IEEE, Sep. 2007, pp. 629–633. doi: 10.1109/ICDAR.2007.4376991.