CAT 9 INFOCOMM

Group 9-17

PROJECT BLIND EYE

Tan Kwang Yik Tremaine 4P3 (21) [Leader]
Ashley Khng 4P3 (10)

ABSTRACT

We see that Blind people generally have lower literacy rates, resulting in more problems like the high unemployment of blind adults. The cause is mainly because most books cater to the visually blessed, and there are limited scanning strategies. We aim to use computer vision to counter this problem and have come to a conclusion that our methods have succeeded after testing.

1 INTRODUCTION

We are creating a device for the visually impaired to assist them in reading text and images printed on paper. This is necessary as these people have lower literacy rates that result in higher unemployment, and this is due to academic resources not catering to them. We aim to create a tool which can aid blind people in reading content on a paper out loud, emphasis on both text and images. We are targeting the visually impaired of age above 12 as basic skills are required for use of our prototype.

2 LITERATURE REVIEW

2.1 Research On Issues Raised

(1) We found that literacy rates of legally blind people are very low. Less than 10% of legally blind people in the USA are Braille Readers, and only 10% of blind children in the US are learning Braille. This shows that not much is being done to curb the problem of low literacy rates. (Jernigan Institute, 2009)

(2) We discovered the lack of tools to educate blind people. More than 70% of blind people are unemployed, and close to 50% of blind high school students drop out. This could be due to a lack of educational resources to aid the blind. (Jernigan Institute, 2009)

(3) Literacy rates are a great consideration factor during employment. According to the US Department of Labour, literacy problems have a $225 billion cost in loss of productivity, injuries, and mistakes in businesses. This shows that low literacy rates of the blind relate to unemployment. (BrailleWorks, 2016)

2.2 Research On Object Detection Algorithm In Use

Our choice of using YOLO is summarised below, but the full research involved can be found in our github page which will be mentioned further below.

1. Why we chose YOLO (You Only Look Once)
    a. Designed for speed and real-time use, which is important for our project as we also have to perform OCR after object detection
    b. Excellent at detecting really small objects
2. Other Algorithms considered
    a. Faster RCNN
        i. Training data is too long
        ii. Training happens in multiple phases
        iii. Network is too slow at inference time
    b. SSD - Single Shot Detectors
        i. Has problems detecting small objects, which are always present in hardcopy documents

2.3 Case Studies

Below are a few projects done by others which can aid blind people in reading on the market currently.

(1) KNFB Reader is a mobile phone app which aids the visually impaired in reading text on documents using OCR technology. This app only detects text and it is very expensive despite just being a software app. (Sensotec, 2020)

(2) Subtitle Reader is an app published on hackaday.io, which uses a camera on a smartphone or tablet to take a picture of a document. It then uses OCR to detect text and read it out via a speaker. (Martin, 2017)

(3) Deaf-Blind Communication is a project made in Arduino which aids communication between deaf and blind people. It uses both a screen as well as a speaker, so the deaf person can read the text on the screen, while the speaker will read out the text for the blind person to hear. (Samir, 2017)

After analysing other projects, we think that our project is superior as they do not have object detection abilities, which is essential as pictures in a document add meaning. Furthermore, many projects are softwares in a smartphone, but they do not consider that the blind are unable to position the smartphone correctly to the paper being scanned.

3 THE STUDY & METHODOLOGY

3.1 Methodology

We have collected data from reliable online sources such as CNA, BBC as well as research papers available online. We also have a practice of making a write-up for content that is new to us, such as the object detection algorithm in use. The software is fully in python integrated with other libraries such as pytesseract for OCR, opencv for object detection and image processing, and cmudict for natural language processing to generate fluent english. Other libraries include gTTS to convert text to speech and pydub to play the audio through the speaker. Autodesk Fusion 360 is used for 3D Modelling as we think it is the most suitable 3D Modelling software for 3D printing.

3.2 Job Distribution

- Software
    - Tremaine
        - Main skeleton of code
        - Text to speech
        - Code cleanup and review
        - Object Detection (together with Ashley)
    - Ashley
        - Camera functionality
        - Optical Character Recognition
        - Object Detection (together with Tremaine)
- Hardware
    - Tremaine
        - 3D Design of scanner (including lid and case)
        - Wiring and soldering of the circuit
    - Ashley
        - 3D Design of stand

Our workflow involves using GitHub for collaboration and a modular-based coding style, where every component has a module that will be imported into the main program that runs. This allows different members to be assigned to a specific component individually and the code will not interfere with the main project. A pull request is made, which is then merged after vetting to ensure that there are no bugs.

Link to GitHub Repository: https://github.com/tremainetan/ProjectBlindEye

3.3 Timeline

- January & February
  - Group formation and decided on the aim of project and target audience
- March
  - Read up projects on computer vision to aid the blind and drafting proposal for mentor approval
- April
  - Conduct needs analysis (Was not possible to reach our target audience of blind people due to COVID-19 restrictions)
  - Complete Proposal Evaluation Slides
- May
  - Complete prototype based on judges' comments
- June
  - Start working on potential features
  - Start working on Written Report
- July
  - Complete prototype and do beta testing
- August
  - Complete Final Evaluation Slides

4 OUTCOMES, ANALYSIS & DISCUSSION

The product of the project is a scanner that captures a picture of a piece of paper/book, and performs OCR and Object Detection on the image to speak the contents out via a speaker for the blind user. The scanner is placed on a 3D printed stand that allows a paper to sit in a capture area.

4.1 Software Product

1. A button connected to the GPIO pins of the Raspberry Pi is pressed
2. Camera captures a picture
   a. Uses Raspberry Pi Built in Python picamera library

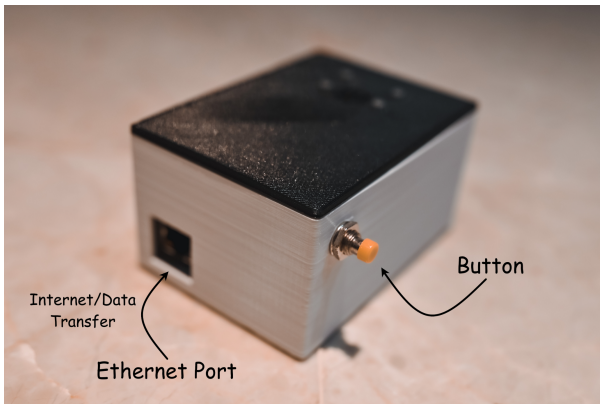
b.
```python
import time
from picamera import PiCamera
CAMERA = PiCamera()

def take_picture():
    #Takes Picture and saves to path pic.jpg
    time.sleep(2)
    CAMERA.resolution = (1024, 768)
    CAMERA.capture("pic.jpg")
```

3. The picture is processed by our program in 2 sections
    a. Object Detection - OpenCV
        i. Library Used: opencv-contrib-python version 4.1.0.25 (YOLO Library)
        ii.
```python
import numpy as np
import cv2
import os

def add_object(objDict, newObj):
    if newObj in objDict.keys():
        # Object already exists
        objDict[newObj] += 1
    else: objDict[newObj] = 1

    return objDict

def get_image(imagePath):
    #Load Input Image using CV2 with Path
    image = cv2.imread(imagePath)
    return image

def scan_image(image, libPath, minConfidence=0.5, threshold=0.3):

def show_image(image):
```
        iii. Full Code can be found in GitHub Page
    b. Optical Character Recognition - Tesseract
        i. Library Used: pytesseract python library
        ii.
```python
import pytesseract
import cv2

#Preprocessing of Image
def get_grayscale(image):
    return cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

#Main Get Text Function
def get_text(image):
    image = get_grayscale(image)
    text = pytesseract.image_to_string(image)

    return text
```

4. Objects detected by OpenCV is converted into a sentence that describes what objects are in the image
    a. Example: “The image contains a banana, two bottles, and an umbrella”
    b. Library Used: cmudict python library to differentiate between when to use “a”, and when to use “an”
    c.
```
import cmudict

def use_an(word, pronunciations=cmudict.dict()):

def generate_read_string(objects):
```
5. Both detected OCR Text and sentence generated above are converted into a stereo audio file using TTS technology
    a. Library Used: gTTS python library is used
    b. Though pyttsx3 is mostly used, we chose gTTS because the pyttsx3 voice on RPi sounds robotic
    c. Lastly, the audio file is read out through the speaker using the I2S Interface
    d.
```
import os, shutil, vlc, time
from pydub import AudioSegment
from gtts import gTTS

vlc_instance = vlc.Instance('--aout=alsa')
audio_player = vlc_instance.media_player_new()

def read(text):
```
6. Main Skeleton
    a. Import Various Modules
    b. Multi-Threading (Speeds up processing by 2 times)
    c.
```
import objectdetect
import strgen
import ocr
import tts
import camera
import os
import time
import threading
import RPi.GPIO as GPIO

OCR_TEXT = ""
OBJECT_TEXT = ""

def main():

def scan():

def ocr_thread(image):

def object_detect_thread(image):

def welcome():

def init_gpio():

if __name__ == "__main__":
```

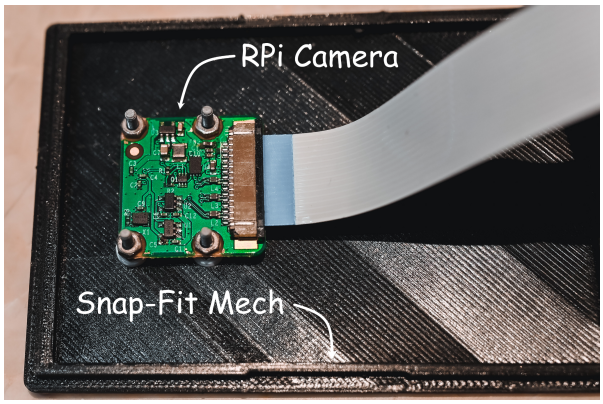
4.2 Hardware Product

Camera: Capture Picture

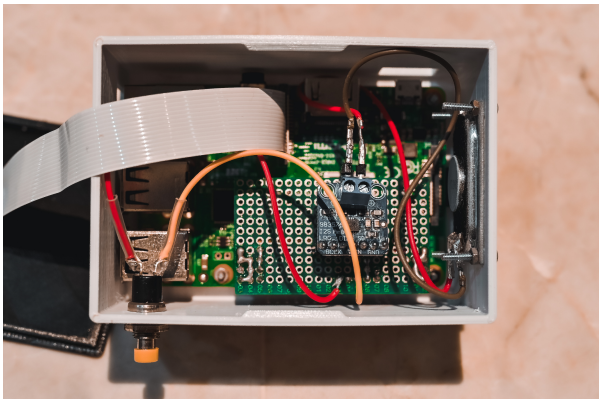


Button: Press to Take Picture

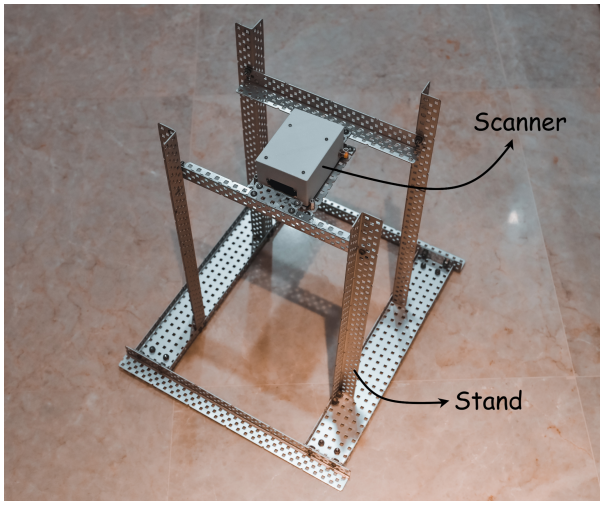
Speaker: Play Audio



Snap-Fit Mech: Sleek Snap-In Lid Design

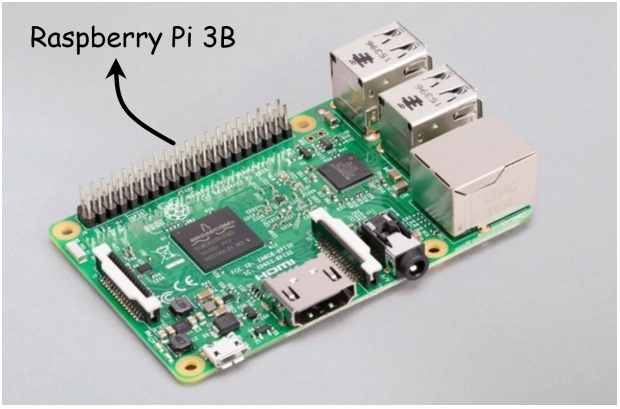


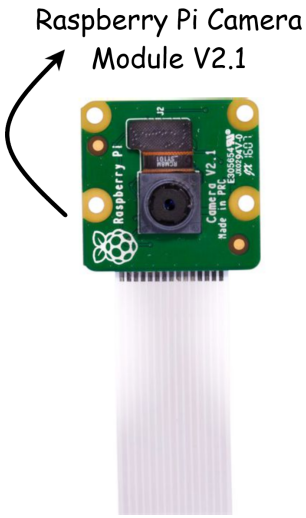
Blue Amplifier Board: Amplify Audio
Circuit Board: Wire Electrical Components

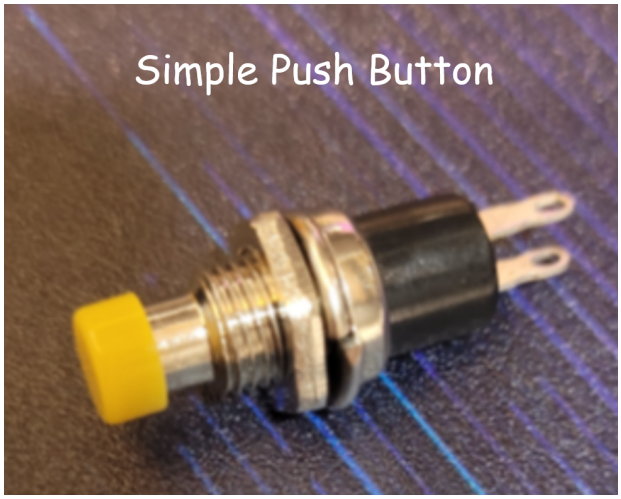


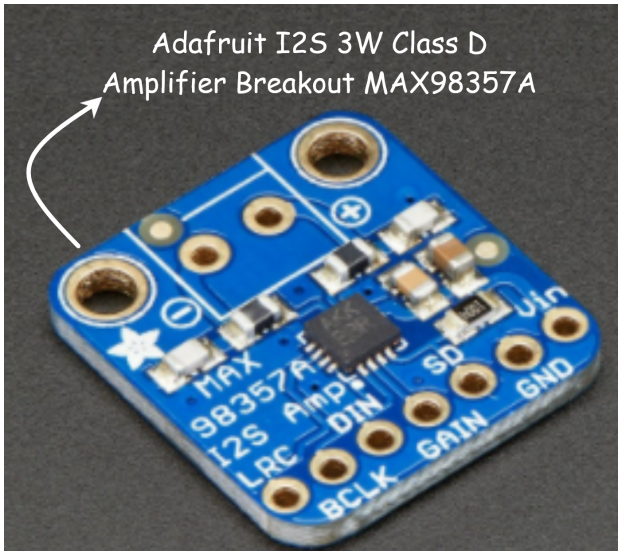
Stand: Helps Blind User in Positioning Scanner

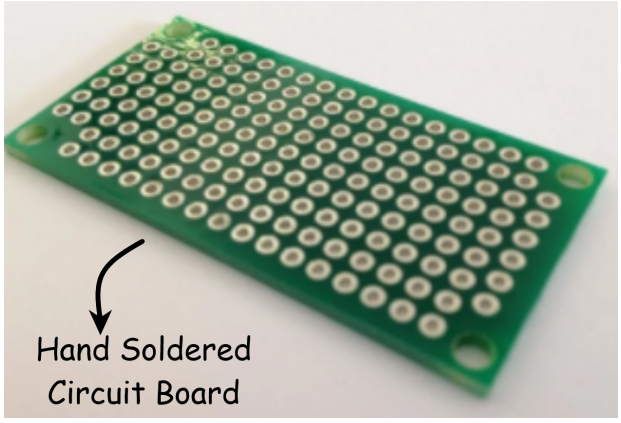
4.3 Hardware Components:





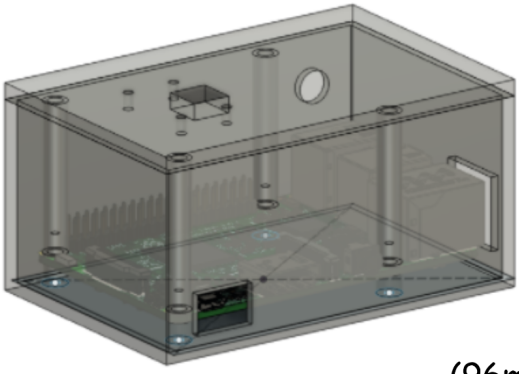




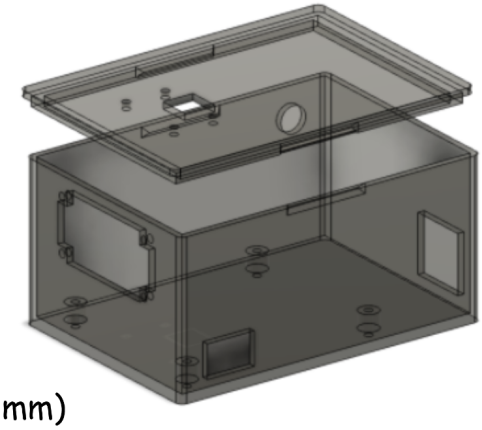


3D Printed Scanner Component

First Prototype



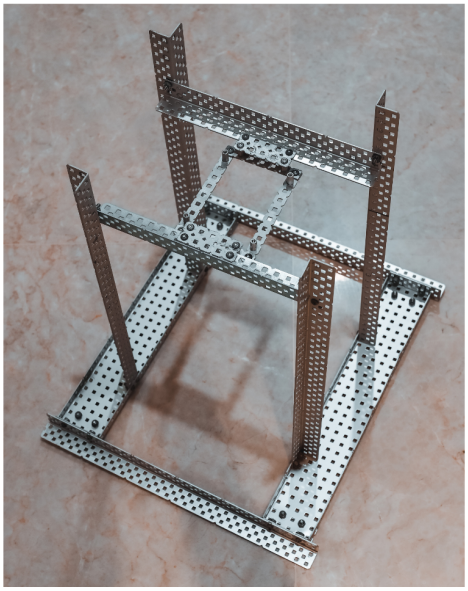
Second Prototype



(96mm x 68mm x 49mm)

3D Printed Stand Component

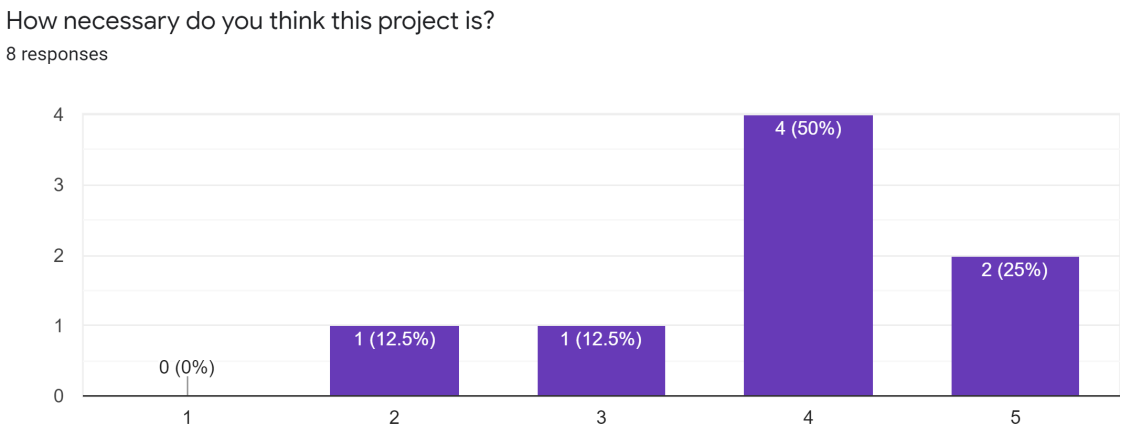
First Prototype

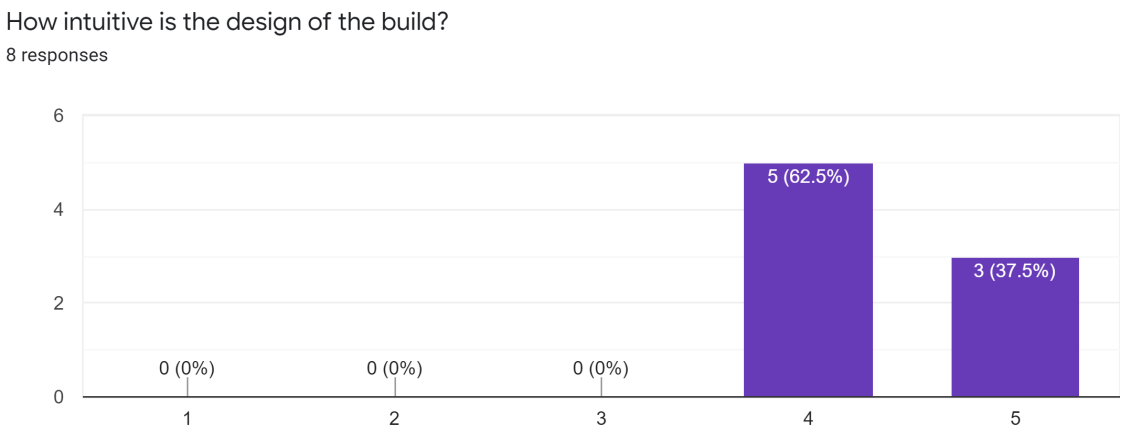


Second Prototype

4.4 User testing

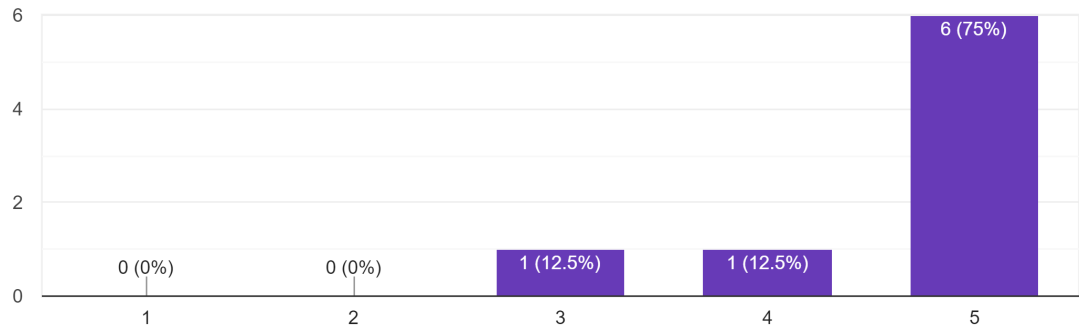
Due to the COVID-19 restrictions, we only conducted physical pilot testing on our immediate family members. The pilot testing was conducted with the stand being the first prototype, but we allowed the volunteers to have a look at the virtual image of our 3D Modelled Stand for them to judge the outlook of the stand.





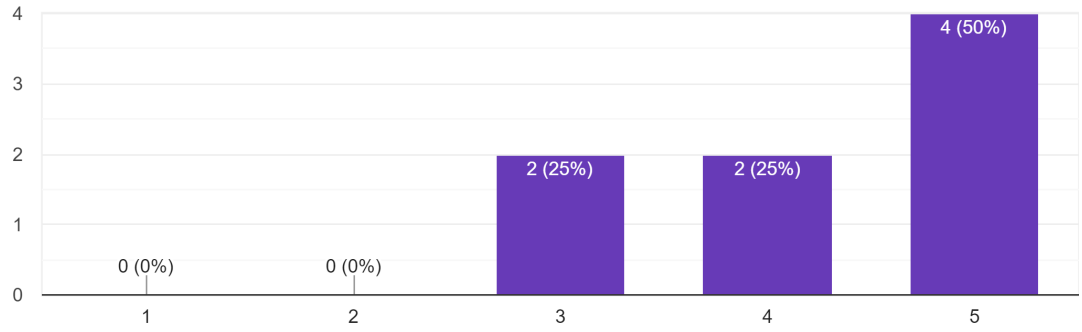
How effective is the scanner in conveying the text message on the paper to the user?

8 responses



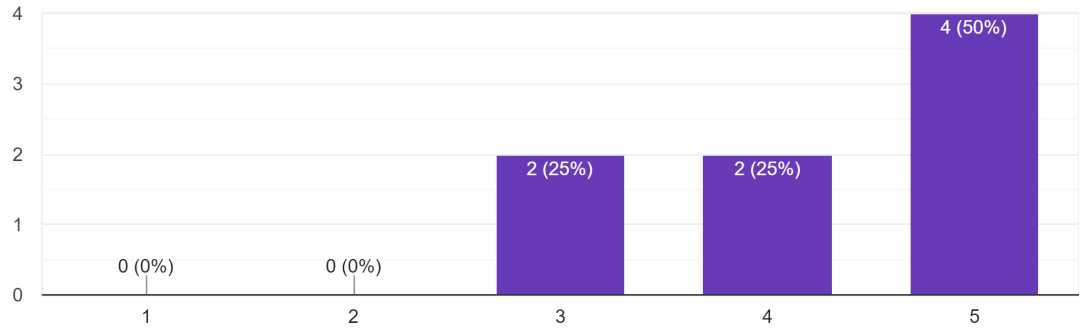
How effective is the scanner in enabling the user to visualize the pictures on the paper?

8 responses



How was your experience using Project Blind Eye?

8 responses



Which feature in the prototype do you like the most?

8 responses

- it can detect images on the paper
- I like how it can form a sentence with the images, makes it more clear and understandable than just listing
- the stand
- the text to speech sounds quite good
- reading text
- the detection of pictures on the paper is quite cool
- i think the design of the thing that takes pictures was done well
- i like the design

Areas for Improvement:

6 responses

- maybe can make into an app
- Ability to store images
- maybe a physical device is not necessary
- repeat button for the text in case they miss it the first time

We can hence conclude that many are amazed at the ability of AI in terms of detecting both text and objects. Thus, we can say that although there is room for improvements, our project has mostly succeeded in attaining the objective of the project to help the blind.

5 IMPLICATIONS & RECOMMENDATIONS

We noted down the areas for improvement of this project after seeking comments from the judges, through pilot testing and getting user feedback. Below are some of the areas for improvements:

1. Adding a headphone jack to the scanner

2. Adding a feature to store images taken
3. Adding a repeat button for the user to relisten to the speech

6 REFLECTIONS

Tan Kwang Yik Tremaine

I learnt through our research and study that not much has been done to aid the blind and also learnt a lot about computer vision algorithms which were new to me. As a leader, I had responsibilities regarding distribution of work, planning the next steps that the group should take, and communicating with other group members.

Ashley Khng

I have learnt that although people are more accepting of people with disabilities, many things can still be done to help them. I have learnt more about object recognition algorithms as well as OCR algorithms. I have also gained more knowledge in 3D modelling and design. I have also learnt how to collaborate with others better in projects, especially in programming.

7 BIBLIOGRAPHY

The Braille Literacy Crisis in America. (2009). Retrieved 25 January 2021, from https://www.nfb.org//images/nfb/documents/pdf/braille_literacy_report_web.pdf

Braille Literacy Statistics and How They Relate to Equality. (2016). Retrieved 25 January 2021, from https://brailleworks.com/braille-literacy-statistics/

McDonnall, M., & Cmar, J. L. (2019, October 14). Blind people have increased opportunities, but employers' perceptions are still a barrier. The Conversation. Retrieved 25 January 2021, from https://theconversation.com/blind-people-have-increased-opportunities-but-employers-perceptions-are-still-a-barrier-124977

Sensotec. (2020, November 9). Version (2.9.1). KNFB Reader. https://play.google.com/store/apps/details?id=com.sensotec.knfbreader&hl=en_SG&gl=US

Martin. (2017, April 19). Subtitle Reader. Subtitle reader | Hackaday.io. computer software, hackaday.io. https://hackaday.io/project/21294-subtitle-reader.

Samir, A. (2017, March 23). Deaf-blind CommunicationVersion (1.0). Deaf-blind Communication with 1Sheeld/Arduino - Arduino Project Hub. Arduino. https://create.arduino.cc/projecthub/skyseeker/deaf-blind-communication-with-1sheeld-arduino-bb3362?ref=tag&ref_id=blind&offset=4.