

Introduction

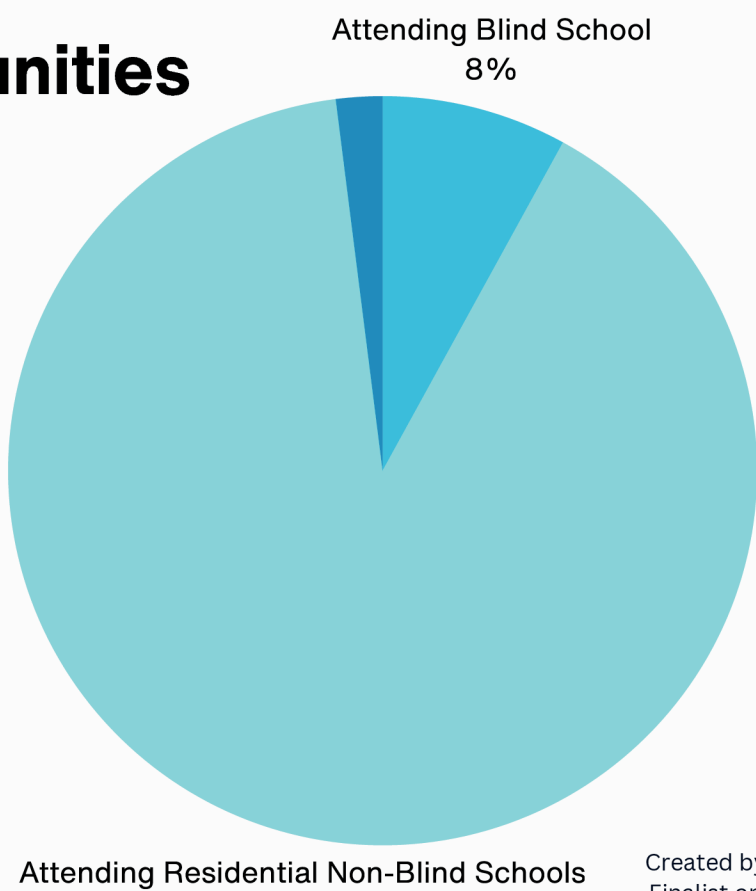
- **600,000 children** face some variety of **blindness or visual impairment** in the US (Centers for Disease Control and Prevention, 2024)
- **Less than 1-7% of books** are **accessible** to visually disabled people (World Intellectual Property Organization, n.d.)
- Only **around 50% of blind students in the US have access to assistive technologies** (Lewis & Farris, 1999)
- Studies show that **approximately 10-15% of blind children** know how to **read Braille** (National Federation of the Blind, n.d.)
- Out of all **blind students** in the US, **less than 8% attend schools specialized** for them (American Foundation for the Blind, 2023)

Importance:

- Lack of accessible educational material leads to:
 - Significant **learning gaps** compared to peers
 - **Limited access** to information
 - Potentially **limiting** future **opportunities**



Lalit, S. (2021). [Blind student in a blind school], in Where do children who are blind or visually impaired go to school? Chicago Lighthouse. <https://chicagolighthouse.org/sandys-view/where-do-children-who-are-blind-or-visually-impaired-go-to-school/>



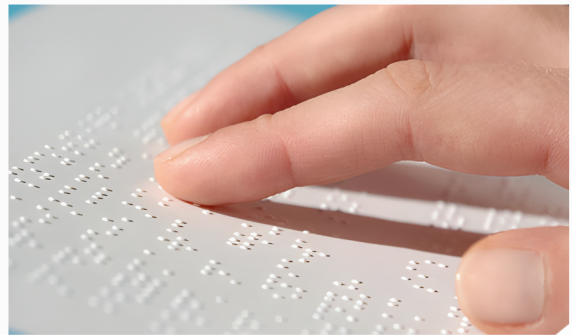
Attending Residential Non-Blind Schools 90%
Attending Blind School 8%
Created by Finalist on Canva, 2025

Current Solutions

- **Braille books** → heavy; **only 1%** of books are in braille
- **Audiobooks** → minimal availability, **<5%**
- **Screen Readers** → only work **online**, **inconsistent** and often **malfunctions**
- **E-pens** → have to process **line by line**; can only process physical text, accuracy heavily depends per case use
- **Outloud Scanners** → immobile, **takes 1-3 minutes** to scan a page of text, **expensive**, unable to scan material in a larger format



[Lyrio assistive reader] (n.d.). New England Low Vision and Blindness. <https://newvision.com/products/lyrio-assistive-text-to-speech-reader>



[2025, February 3]. [Blind individual reading braille] in Braille as Modern Digital Assistive Technology. New England College of Optometry. <https://www.nec.edu/news/braille-as-modern-digital-assistive-technology/>



[Image of Scanmarker Reader]. (n.d.). Scanmarker. <https://scanmarker.com/products/scanmarker-reader>



[Audiobook]. (2014, June 25). In D. L. Jacobs, Listening to audiobooks while you do something else is the ultimate in multitasking. Forbes. <https://www.forbes.com/sites/doborahjacobs/2014/06/25/summer-reading-listening-to-audiobooks-while-you-do-something-else-is-the-ultimate-in-multitasking/>

Engineering Goal

Our engineering goal is to create an **easy-to-use device** that allows **visually impaired children** to **read educational books & resources** in an **auditory form** in areas with fewer resources. We want to allow visually impaired individuals to **receive equal opportunities** compared to their peers.

Heavy and immobile	→	Portable + 3 hour use time, lightweight, usable for students
Inaccurate and slow	→	Above 80% accuracy, under 45 seconds + variable lighting
Low resources/ expensive	→	Under \$100 + works on books, signs, handwriting, etc.

Visionary: AI Glasses for Real-time Text-To-Audio Transcription to Help Visually Impaired Students

Proposed Solution

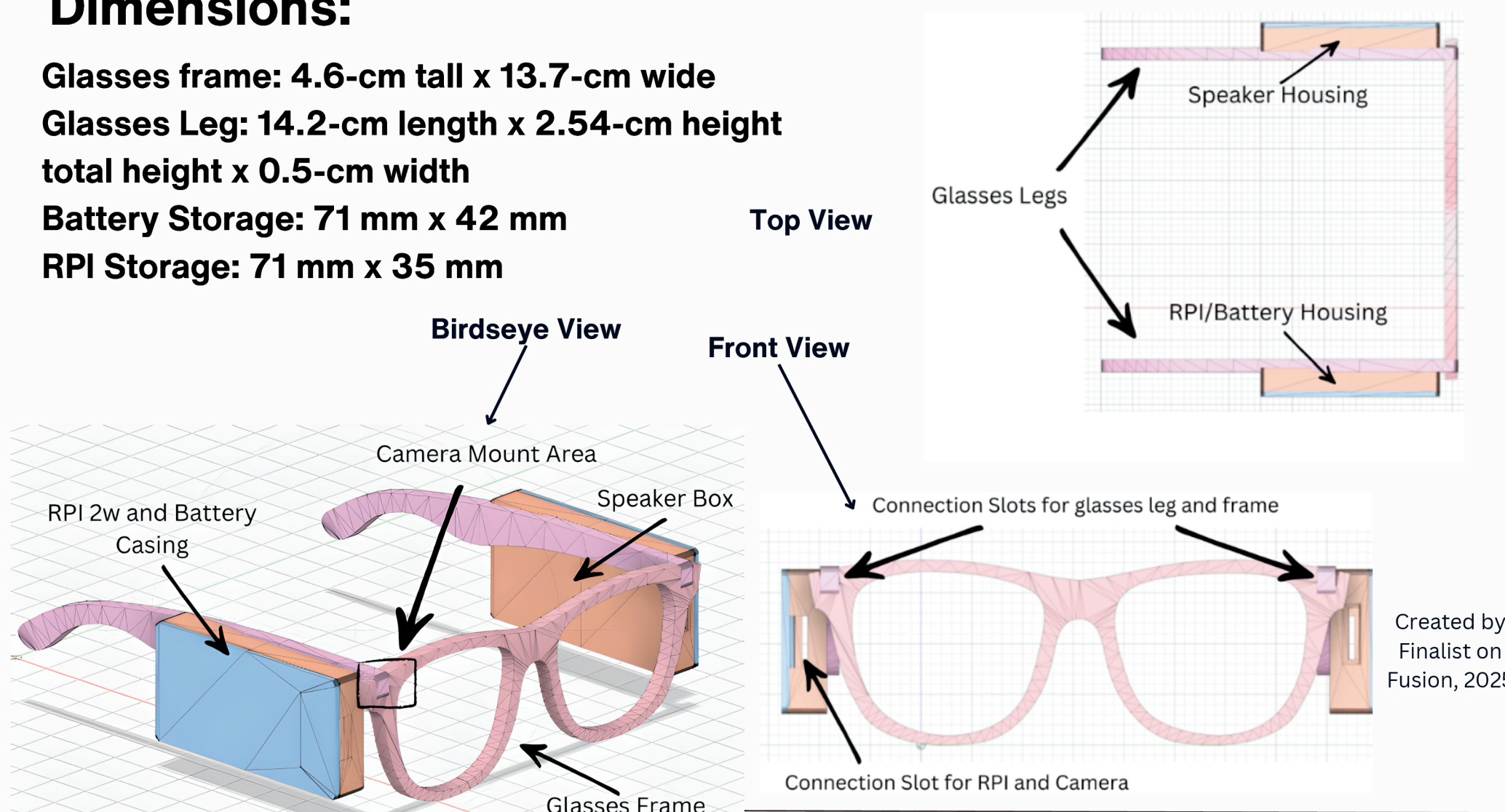
Auditory Vision Glasses

- Pair of **smart glasses** that give **real-time text-to-speech**, powered by artificial intelligence and neural networks
- Provides a **real-time audible format** of text in front of the user, such as **books, signs, and handwriting**
- The **lightweight, power-efficient** design allows it to be portable, comfortable, and running for hours
- Utilizes a **CRNN** with a **custom trained model** for **increased accuracy** of colorful and fancy fonts often used in school books
- Solves issues current solutions face; **accuracy, functionality, linear reading, portability, cost**

3D CAD Models

Dimensions:

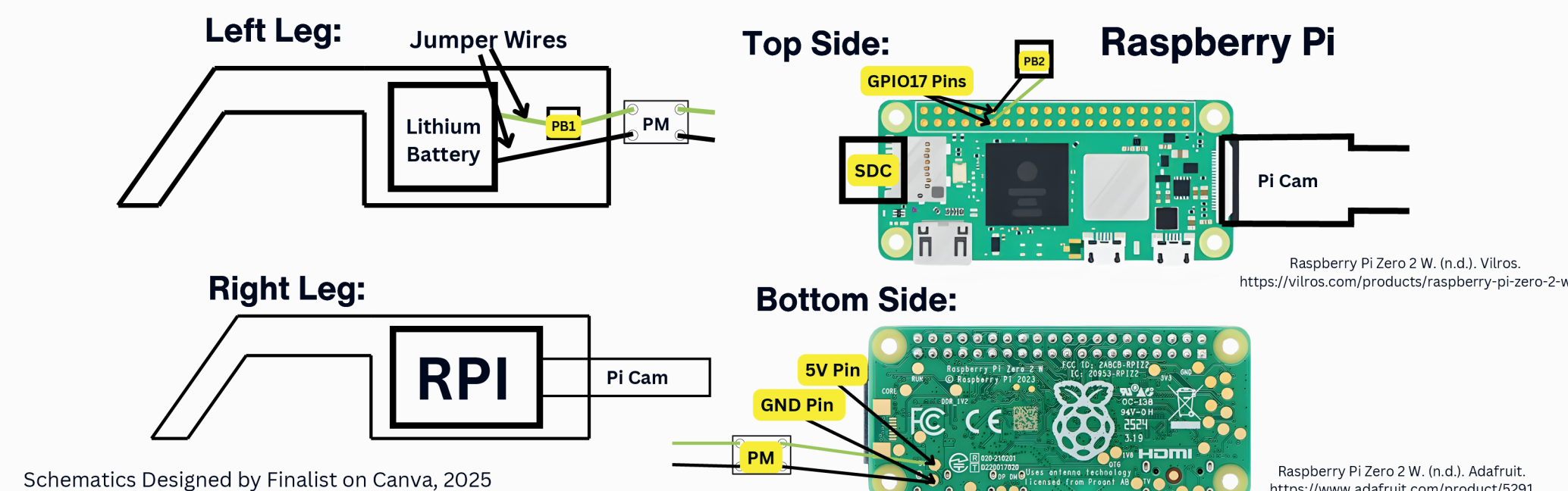
Glasses frame: 4.6-cm tall x 13.7-cm wide
Glasses Leg: 14.2-cm length x 2.54-cm height
total height x 0.5-cm width
Battery Storage: 71 mm x 42 mm
RPI Storage: 71 mm x 35 mm



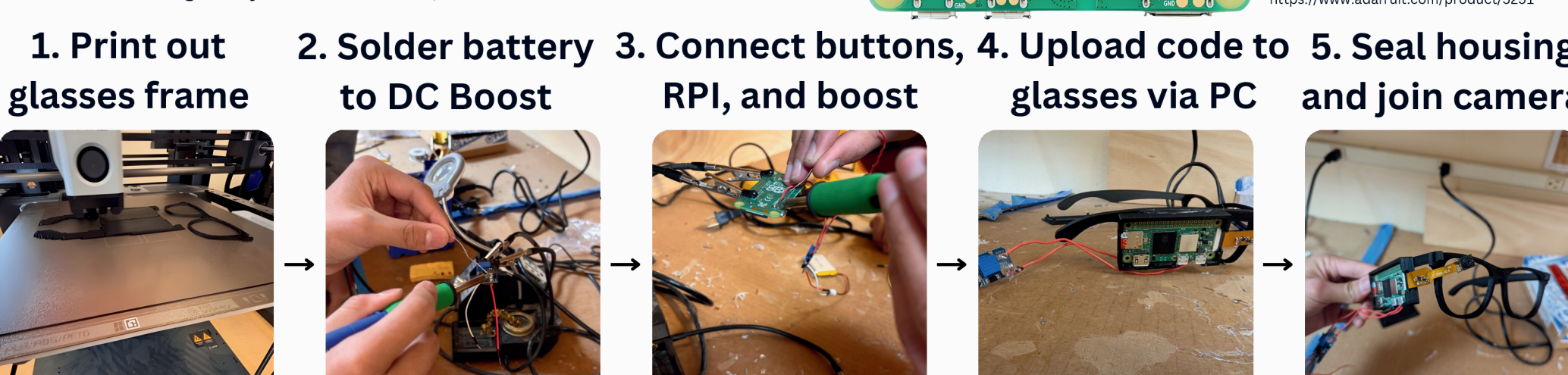
Created by Finalist on Fusion, 2025

Building Procedures

Hardware Design:



Schematics Designed by Finalist on Canva, 2025



Images Captured by Finalist, 2025

Materials

- Raspberry Pi Zero 2W
- Micro Rapsberry Pi Camera Module
- 3.7 Volt Lithium Polymer Battery
- MT3608 DC-DC Step-up Boost Converter
- PLA Filament
- Mini Tactile Pushbuttons
- Jumper Wires

Total Cost:
\$68.92

Software Design Process

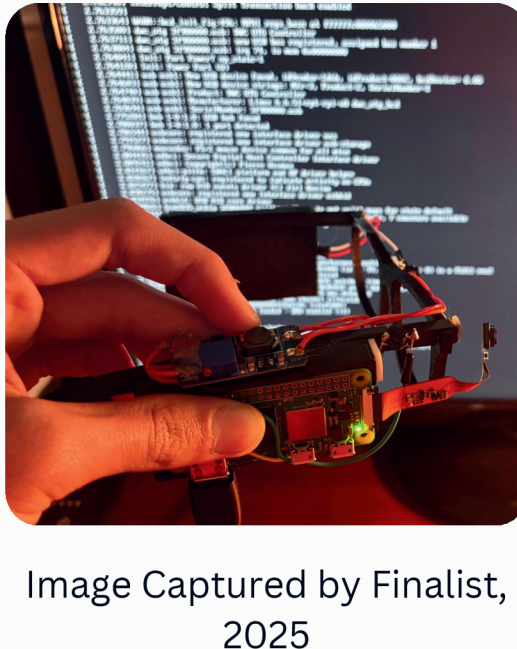
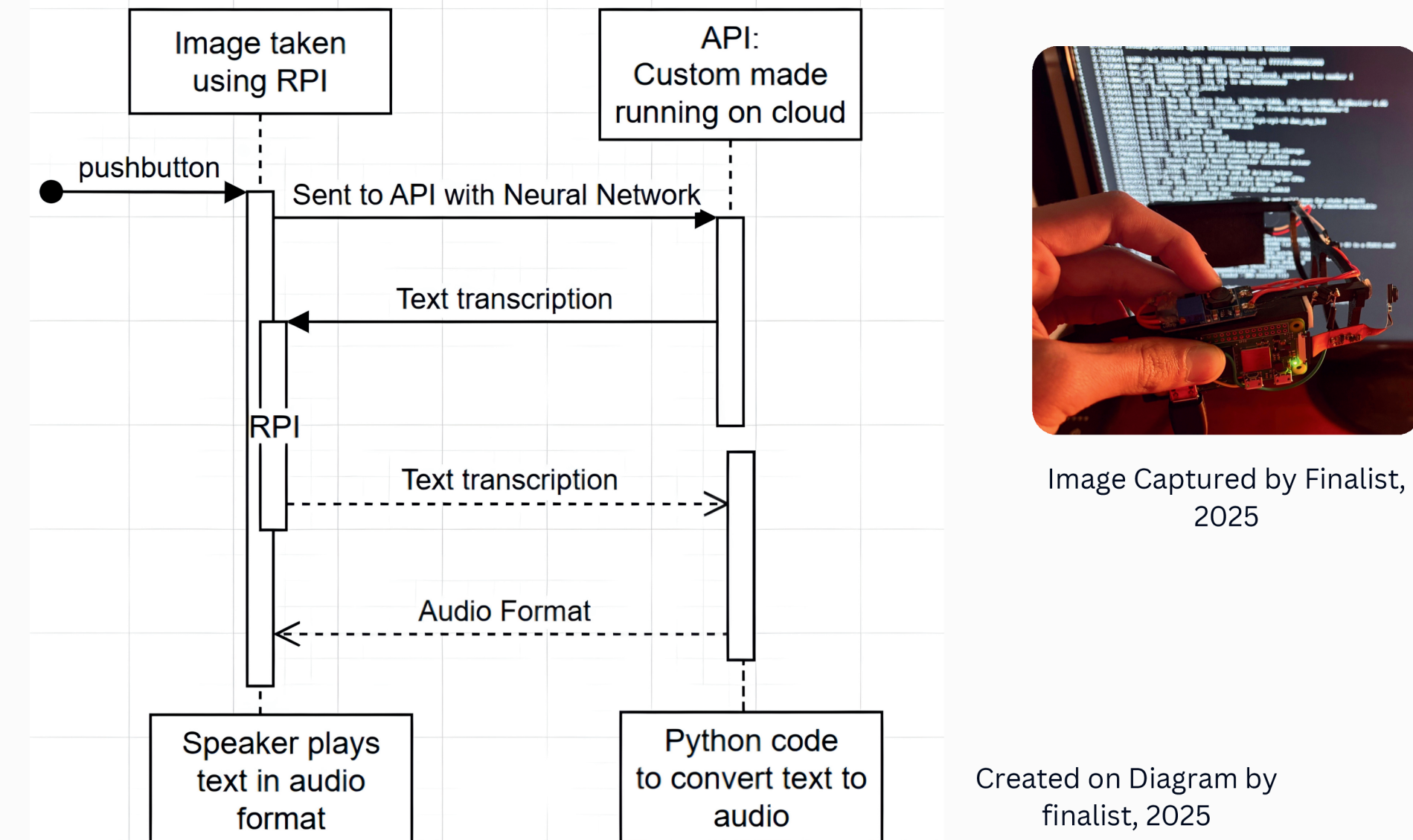
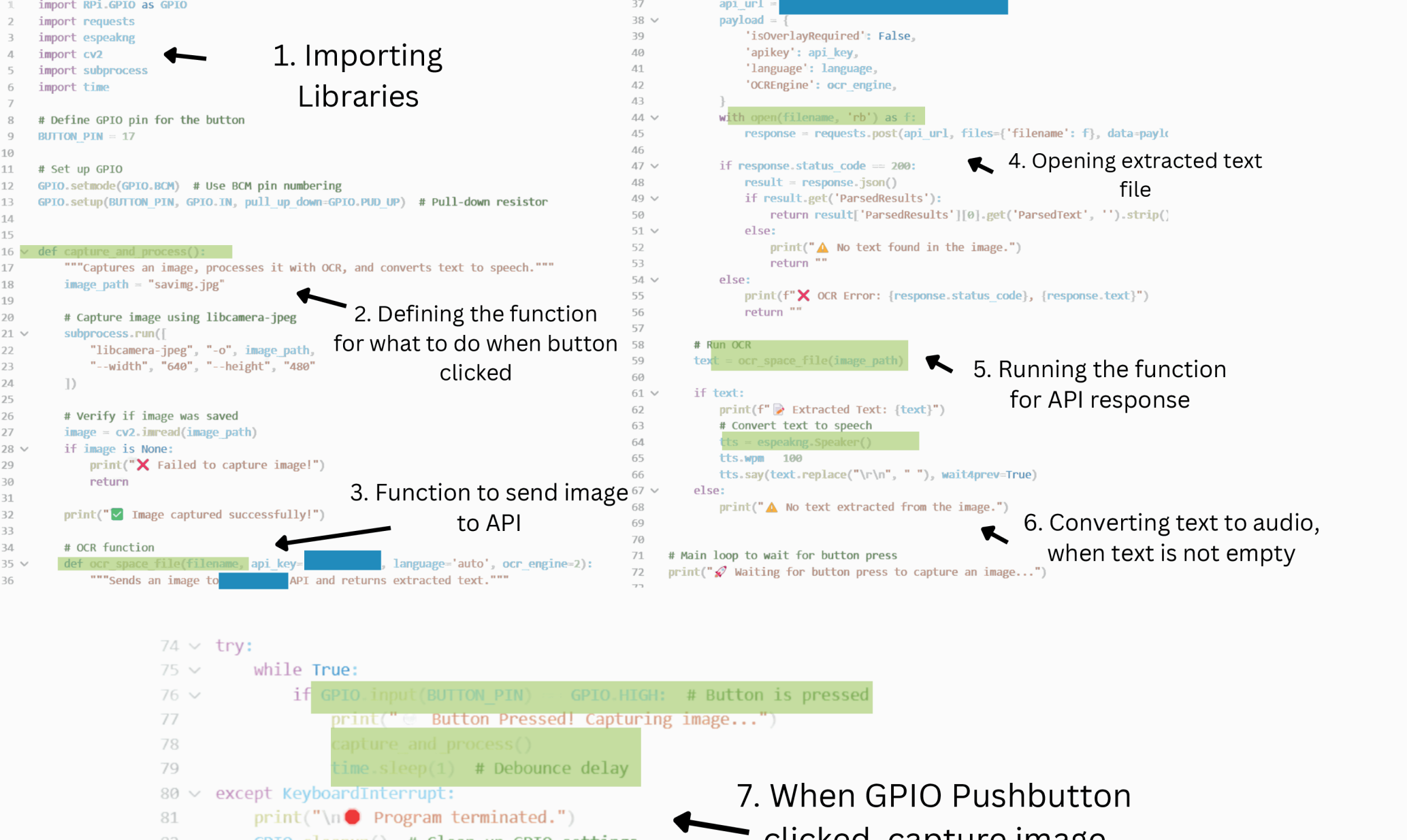


Image Captured by Finalist, 2025

Created on Diagram by finalist, 2025

1. Collected a dataset of images from textbooks and books in the classroom ~ **800 images**, manually annotated.
2. Custom-trained an **Convolutional Recurrent Neural Network** to extract text; trained our own model as existing ones couldn't read **fancy fonts + colorful backgrounds**
3. Model then run using **Python** code + **Pytorch**, with **OpenCV**, could open images and extract the text on a local machine. Using **Ngrok**, we were able to have the **local server** then become a **cloud server**, using the **HTTPS** protocol.
4. Separate code running on glasses worked with **custom API** to do **text-to-speech** with Espeak

Software Code Design



Pictures captured on Visual Studio Code by finalist, 2025

Methodology

1. Test our **CRNN detection system**; accuracy rate must be **higher than 80%**, measured by formula (correct characters/total characters) * 100
2. Test if the software can convert the detected text into a broadcast-able **MP3 file** with a **95% success rate**.
3. Test if the glasses can read text in different lightings; **sunlight, dimly lit room, classroom**; using lux meter
4. Time how long it takes for glasses to **power on**, and then provide audio out; must be **under 45 seconds**.
5. Test the software on the physical prototype to see if it will **broadcast** the text to our **speaker system** (must work **95% of the time**; reliable). We will hold up our prototype while testing.

Software Results #1

Document/Font Size	Lighting	Accuracy	MP3	Start Up Time (sec.)	Audio Played
Mirrors and Windows ELA (12pt)	Low Lighting (250 lux)	90%	✓	14.8	✓
Mirrors and Windows ELA (24pt)	Classroom (500 lux)	98%	✓	13.2	✗
Mirrors and Windows ELA (36pt)	Daylight (5000 lux)	94%	✓	14.0	✓
Mirrors and Windows ELA (48pt)	Classroom (500 lux)	99%	✓	12.9	✓
Big Ideas Math (12pt)	Low Lighting (250 lux)	91%	✓	14.8	✓
Big Ideas Math (24pt)	Classroom (500 lux)	97%	✓	13.5	✓

Graph created by Finalist using Canva, 2025

Hardware Results #2

Document/Font Size	Lighting	Accuracy	MP3	Start Up Time (sec.)	Audio Played
Big Ideas Math (36pt)	Daylight (5000 lux)	93%	✓	13.9	✓
Big Ideas Math (48pt)	Classroom (500 lux)	98%	✓	13.0	✓
Inspire Science (12pt)	Daylight (5000 lux)	92%	✓	14.5	✓
Inspire Science (24pt)	Classroom (500 lux)	97%	✓	13.4	✓
Inspire Science (36pt)	Low Lighting (250 lux)	90%	✓	14.9	✓
Inspire Science (48pt)	Classroom (500 lux)	99%	✓	13.1	✓

Graph created by Finalist using Canva, 2025

Analysis/Conclusion

- Achieved **96% CRNN accuracy**; best in classroom light
- Text → MP3 conversion success: **100%** (MP3 files created reliably).
- Audio **playback success: 91.6%** (works reliably).
- Lighting performance: reliably detects text in direct **sunlight, classroom lighting, and dim lighting**.
- Startup → audio latency: average **13 seconds** (greatly **beating set criteria**)

In conclusion, we **successfully** created a device that **reliably** can **scan and extract text** from in front of the user, convert it to a written text file, **convert** the text file into an MP3 file, and then **broadcast** the MP3 on the glasses.