Project IRAS

ECE Group 4

Who is IRAS?

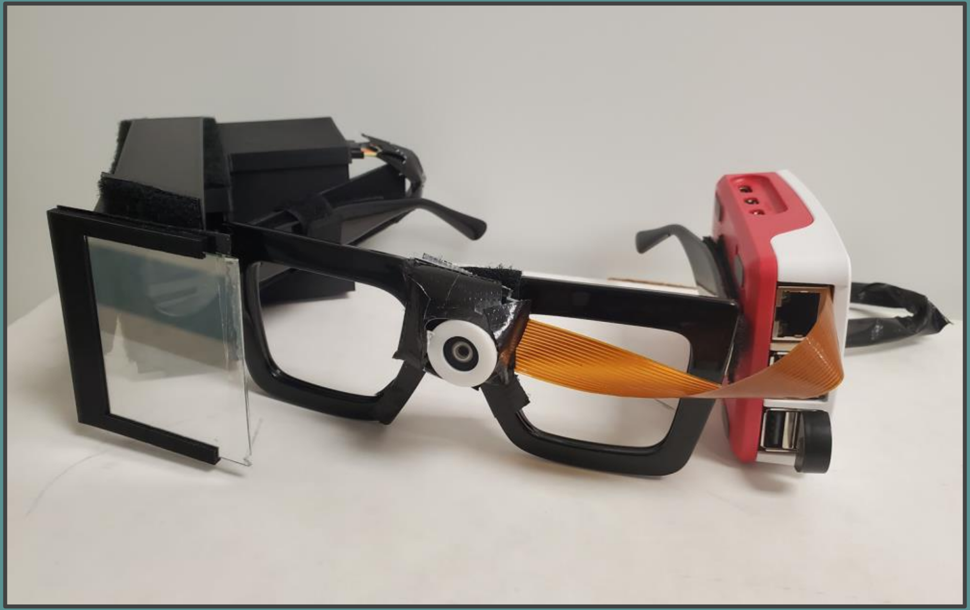
Carlos Acosta,
PSE, Project Lead

Darlandie Moise,
PSE

Kim Le,
CS

Brian U,
CS

What is IRAS?



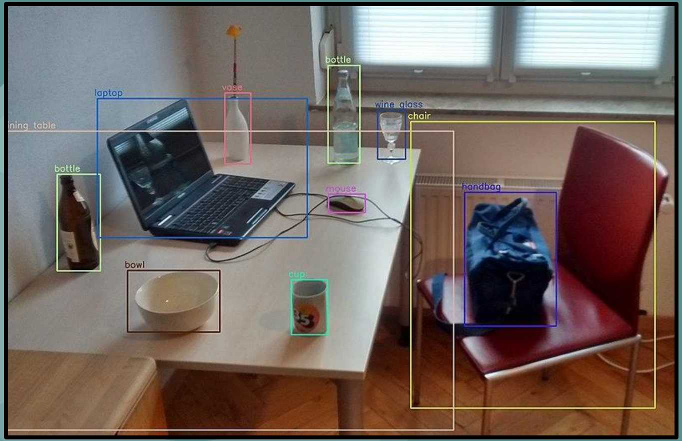
Integrated Real-Time Assisted Spectacles (Speech-to-Text and Color Detection AR Glasses)

Working towards improving the quality of life for those with color blindness, hearing issues, and to break language barriers.

Motivations

- Millions of people are affected by hearing impairment and color blindness.

- We created a wearable device aimed at helping the deaf, hard of hearing, elderly, and those who struggle with color blindness. This device enhances spoken communication comprehension through captions and translations. It also aids visual perception by using bounding boxes.



Goals

Optics

- **Basic**
  - Create the Lens System with desired FOV.
  - Create LCD Projection design.
- **Advance**
  - Keep the Lens System minimal with little distortion.
  - Allow LCD Projection design to be wearable.
- **Stretch**
  - Allow Lens System to utilize autofocus capabilities.
  - Simplify the design to minimize the overall size of LCD projection.

Computer Science

- **Basic**
  - Design an algorithm that utilizes both Speech-to-Text (StT) and Color Detection with Raspberry Pi 5.
- **Advance**
  - Apply customization to the StT and Color Detection to translate varying languages and colors.
  - Use an app to control customization and program.
- **Stretch**
  - Create our own app.
  - Implement an offline StT mode.
  - Improve Color Detection under different lighting conditions.

CS Objectives

- Software Design
  - Speech to Text and Color Detection
    - Capture audio and video
    - Implement Speechmatics and OpenCV
    - Format text printing and bounding boxes
    - Improve latency
  - Mobile App
    - Integrate with program
    - Customize appearance
  - Testing
    - Get user feedback, adjust system

Optics Objectives

★ The Parts, The Systems, and The Photonics

a. Design multiple optical systems
   - AR LCD Projection
   - Wide Angle Lens Systems

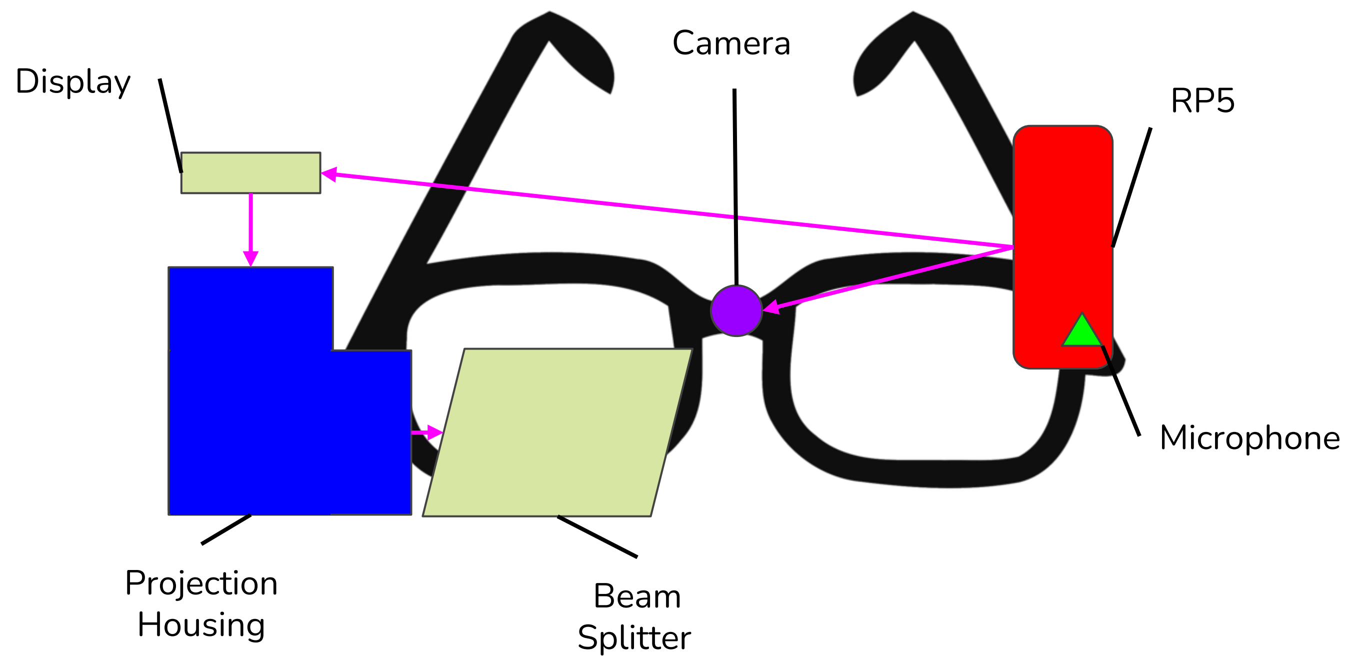
b. Gather parts provided

c. Perfect each optical design with trial and errors
   - Zemax simulations

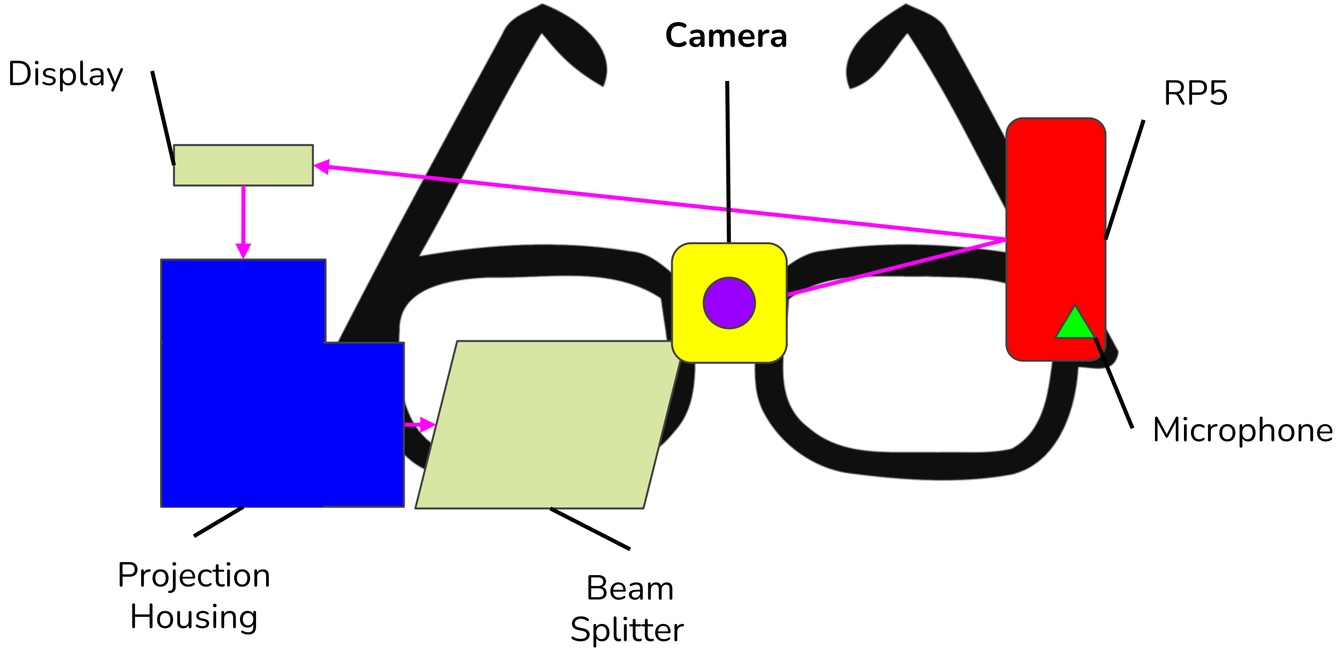
The Engineering Specifications

Components	Parameters	Specifications
Antenna: Raspberry Pi 5	Transmission Range	50m-100m
Camera: OV5640	Capture Quality (Resolution and FOV)	Between 2-5 MP. FOV; ~60 - 80 Degrees
LCD TFT Display	Text and Color Accuracy	>80%
LCF TFT Display	Text and Color Latency	<2s
Frame of Lens Designs	Weight	<500 grams
Camera Lens System	Widen FOV	~80 Degrees Horizontally
LCD Projector	Transparency/Visibility	70% Transparent

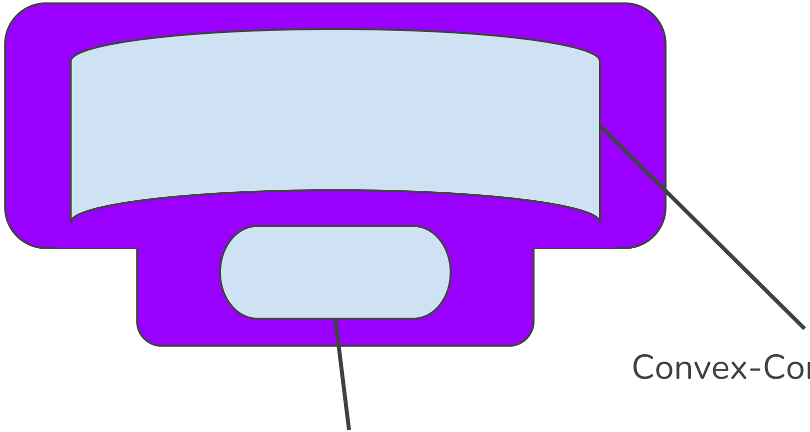
Hardware Design



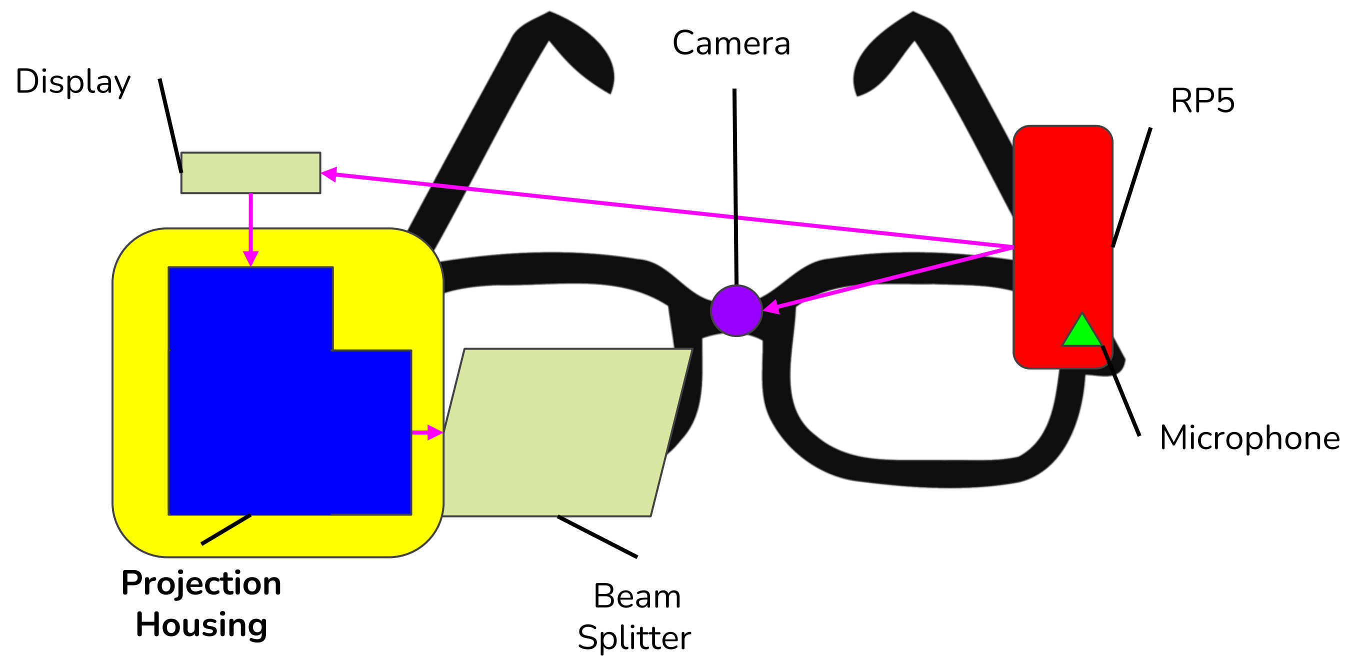
Hardware Design



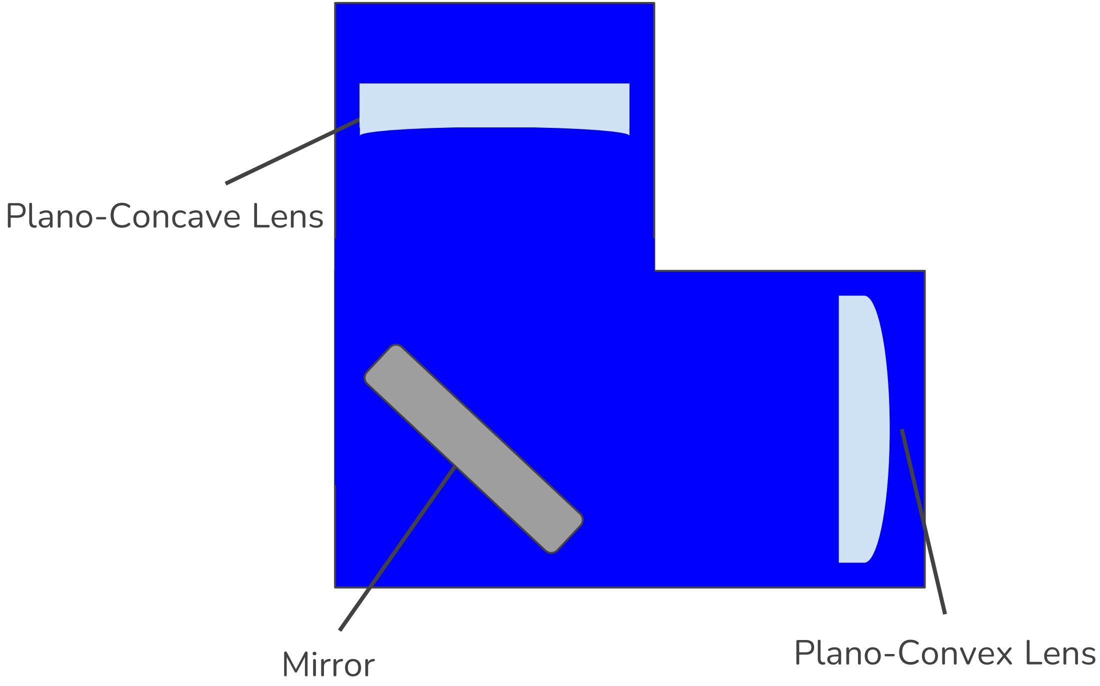
Camera Hardware Design



Hardware Design



Projection Hardware Design

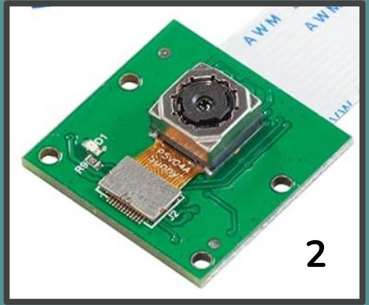


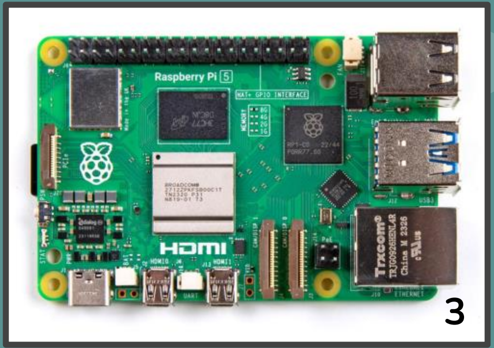
Current Technological Comparison

	Colorblind Glasses	Transcribe Glass	XRAI Glasses
Details	Effective for those with certain color blindness. Lacks customization and environmental factors.	Easy to use. Has closed captions. MUST have frame. Low quality.	AI Assistance. Customizable on smartphone. Lack of other accessibility accommodations.
Tweaks	Expand customizability for color blindness.	Enhance flexibility for non glasses users.	Make sure both hearing impaired and color blind users are accommodate.
Imagery			

Hardware Components

1. Microphone
    a. SunFounder USB 2.0 Mini Microphone
2. Ov5640 5MP Camera
    a. Compatible with RP5
    b. Extra features
        i. 60° FOV
        ii. Auto-Focus
3. Raspberry Pi 5
    a. Python Library
    b. Main SBC
4. 1.69” LCD TFT Display
    a. Compatible with Raspberry Pi 5
    b. Bright





Parts/Technical Reasons

- Commercially available
  - These components are what's available on the market.
    - Lenses are NOT cheap
  - The boards, display, and camera can easily be switched out.
    - Compatibility
- Flexibility
  - Allow any user to wear these glasses.
  - Everyday use.
- Simplicity
  - Capitalize on using 2 models in social and environmental interactions.
  - Easy for anyone to start using and customizing these glasses.
- Affordability
  - Less complexity = Cheaper usage.
- Minimalistic
  - Keep the design small and simple, as long as it works!

Successes/Difficulties

Successes

- Camera
  - Livestream with wide FOV
- LCD TFT
  - Power it on and see texts and images
  - Display text on a reflector
- SBC Compatibility
  - Both the camera, mic, and LCD TFT are connected to one board
  - Color Detection and Speech to text are functioning
- Mobile App

Difficulties

- Expensive
  - Lenses & Displays
- Size/Weight
  - Large housing all the components
- WiFi Connection
  - Delayed connection
- Power Connection
  - Power bank vs outlet

Impact of Design Constraints

Manufacturability

- Difficulty finding compatible parts and small in size.

Health & Safety

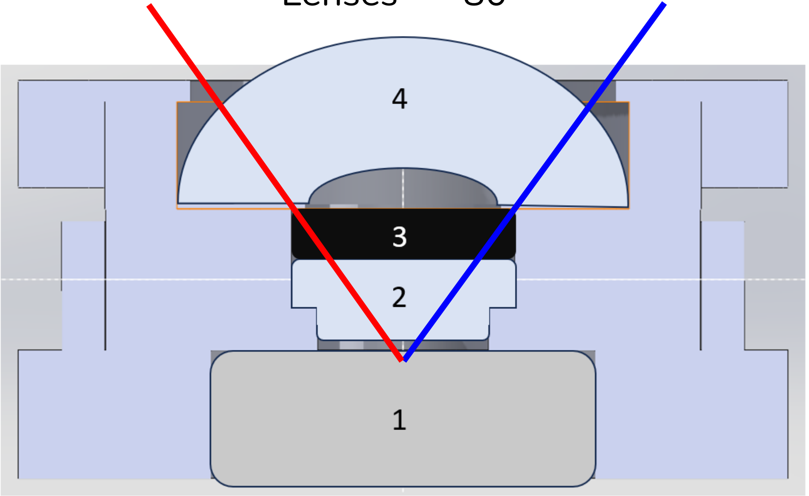
- Safety concerns with wearing electronic components.
- Risk of display flickers that may cause issues to the user.

Economic

- Parts can be very expensive to purchase
- Speechmatics testing costs money.

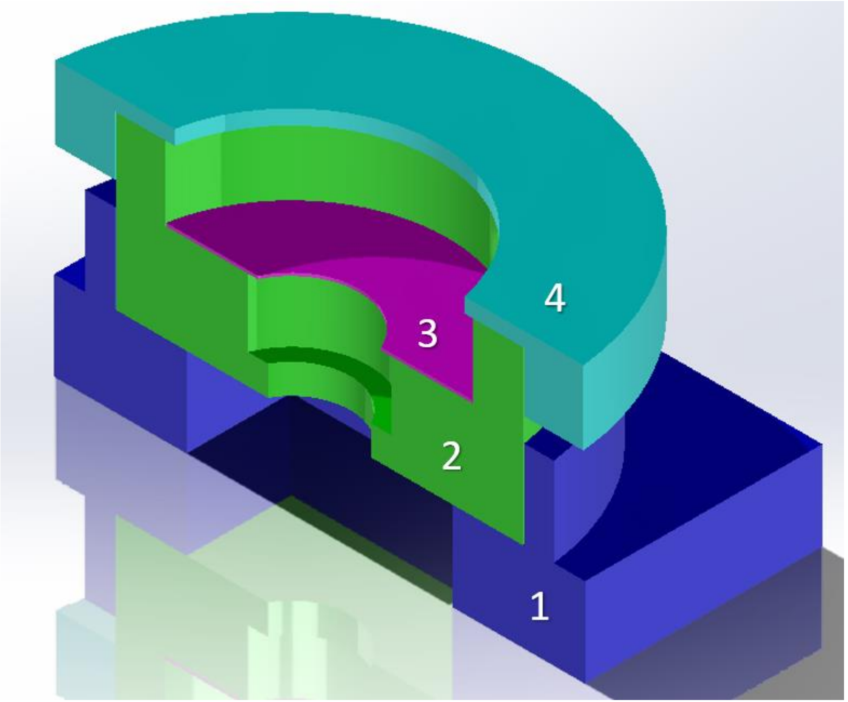
Lens System Design Schematic

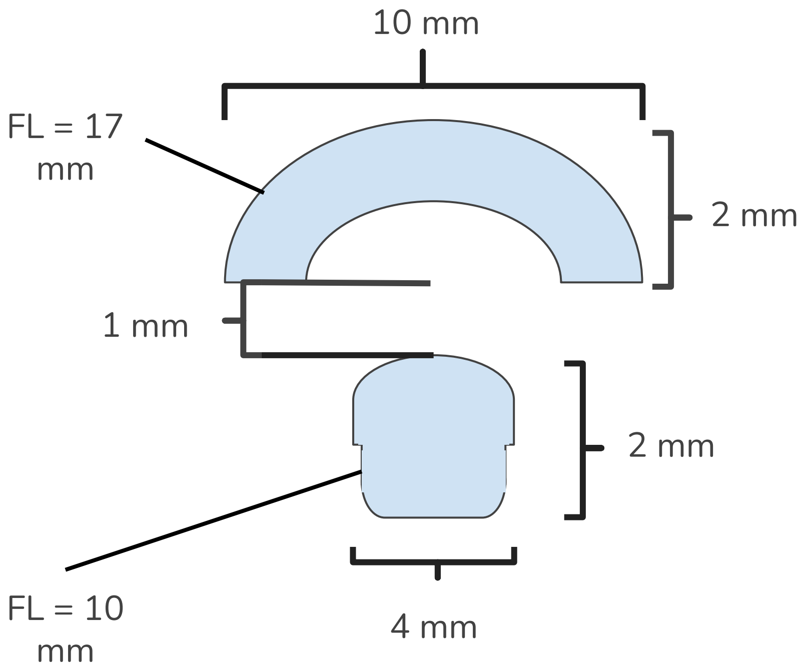
2-4: Wide FOV Lenses = ~80°



1: Camera

Lens System Design Hardware

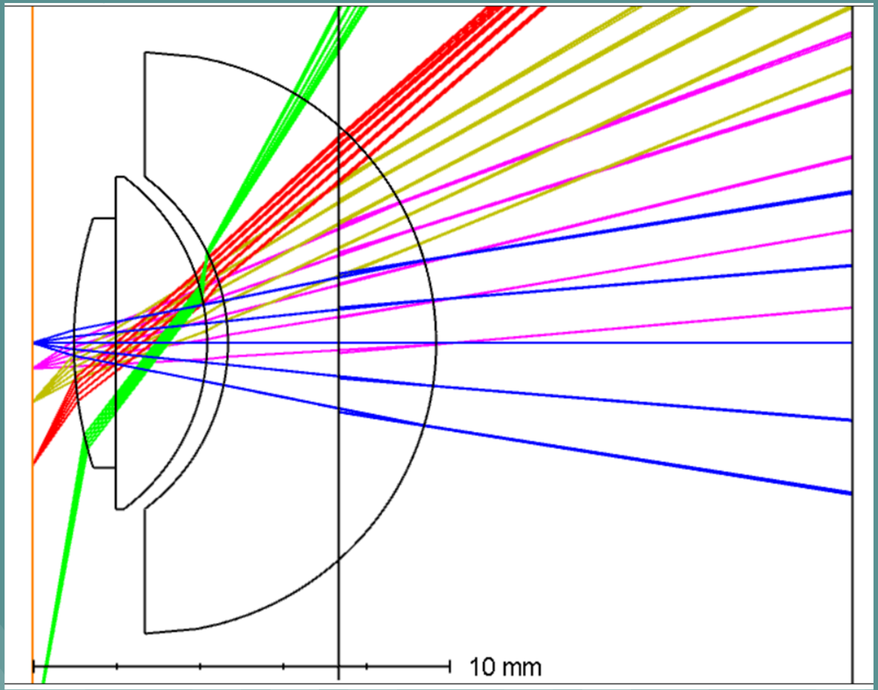


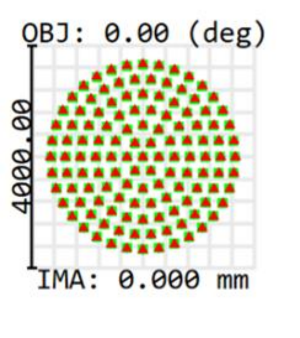


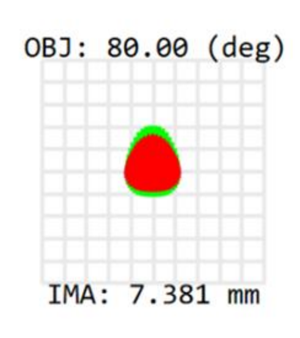
Lens System Design Parts List

Part	Specification	Purpose
Lens System (19.2)	Pinhole diameter - 4 mm with a thickness of about 1 mm. Small housing diameter 5.2 mm and the depth is about 2.3 mm. Larger housing diameter - 10.6 mm and a depth of about 2.5 mm. Outer diameter - 14 mm with a height of about 5.8 mm	Houses the optical design of the camera
Lens Cap (19.4)	Outer diameter - 18mm, Larger Inner diameter - 16mm, Smaller Inner diameter - 10 mm. Length of Lens Cap - 2.5 mm with the inner thickness being about 1 mm	Caps the lenses within the lens system and keeps them from falling out
Plastic Spacer (19.3)	Inner diameter - 4.6 mm Outer diameter of 10.6 mm, with a thickness of about .1 mm	Separates and holds the lenses in place
Metallic Spacer (20.3)	Diameter - 5 mm, Thickness - 1 mm	Keeps the smaller lens in place and separate from the larger lens, little distortion
Smaller Lens (20.2)	Diameter - 3 mm and 4 mm, Thickness - 2 mm, Focal Length - 10 mm	The entry lens for the optical design
Larger Lens (20.4)	Diameter - 10 mm, Thickness - 2 mm, Focal Length - -17 mm	The main lens providing the FOV of the lens system
Lens Mount (19.1)	Base - 18 mm x 18 mm Inner diameter - 14 mm Outer diameter - 16 mm	Holds camera and lens system together
Camera (20.1)	8.5 mm x 8.5 mm x 5.4 mm	The main source for visual input of our project

Zemax

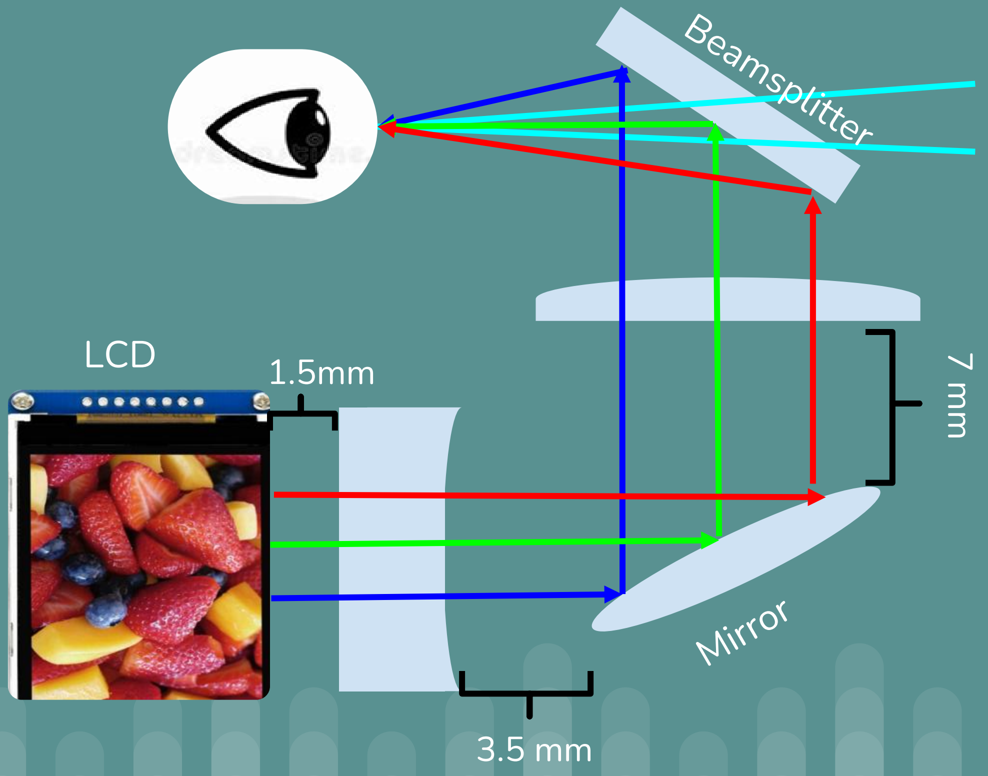

10 mm


OBJ: 0.00 (deg)
4000.00
IMA: 0.000 mm


OBJ: 80.00 (deg)
IMA: 7.381 mm


0.55
0.4
0.7

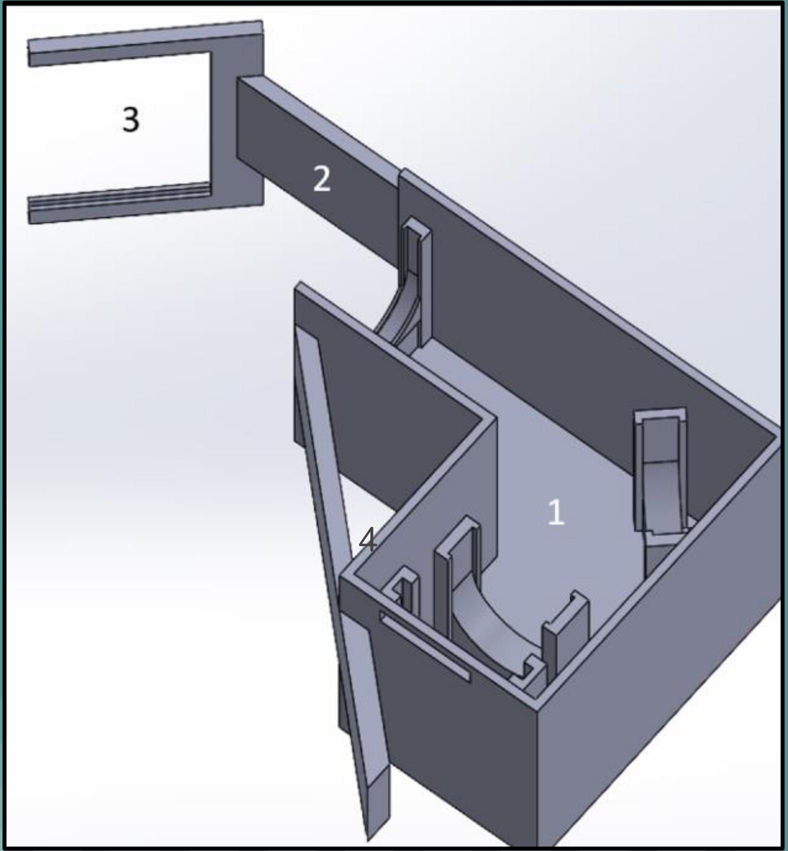
Lens Text Projection Optics Schematics



Components:

- LCD TFT
  - 1.69 in Display
  - 400 nits
- Plano Concave Lens
  - Focal Length: -50 mm
  - Diameter: 25.4 mm
- Mirror
  - Silver Coated
  - Diameter: 25.4 mm
- Plano Convex Lens
  - Focal Length: 25 mm
  - Diameter: 25.4 mm
- Beamsplitter
  - Anti-Reflection Coating
  - 70T/30R
  - 40 x 30 x 1 mm

Near Eye Projection Setup



(1) Main Housing (LCD, lenses, mirror)
Designed to fit each component properly in place with correct distances.

(1) Beamsplitter Support
Attached to the beamsplitter holder
Adjustable/customizable based on the user

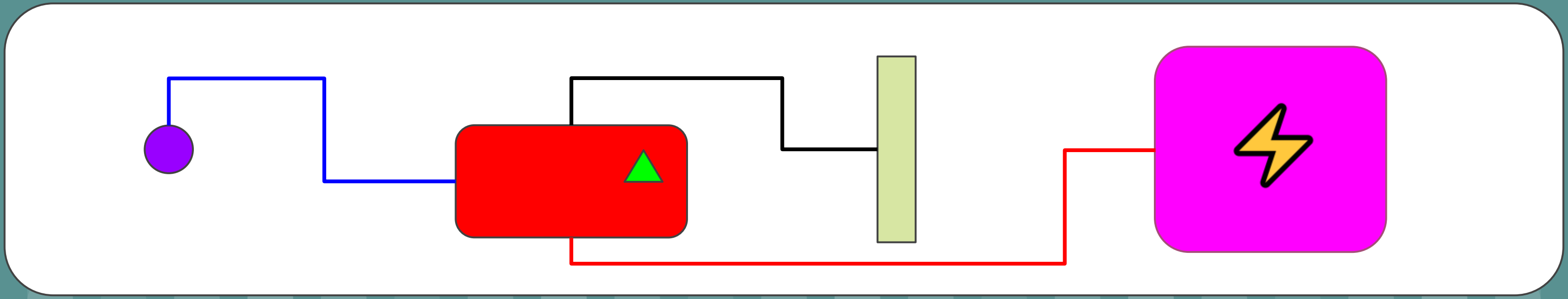
(1) Beamsplitter Holder
Also has flexibility to be adjusted

(1) Glasses Support Beam
Enables us to directly attach to the side of a pair of glasses frame

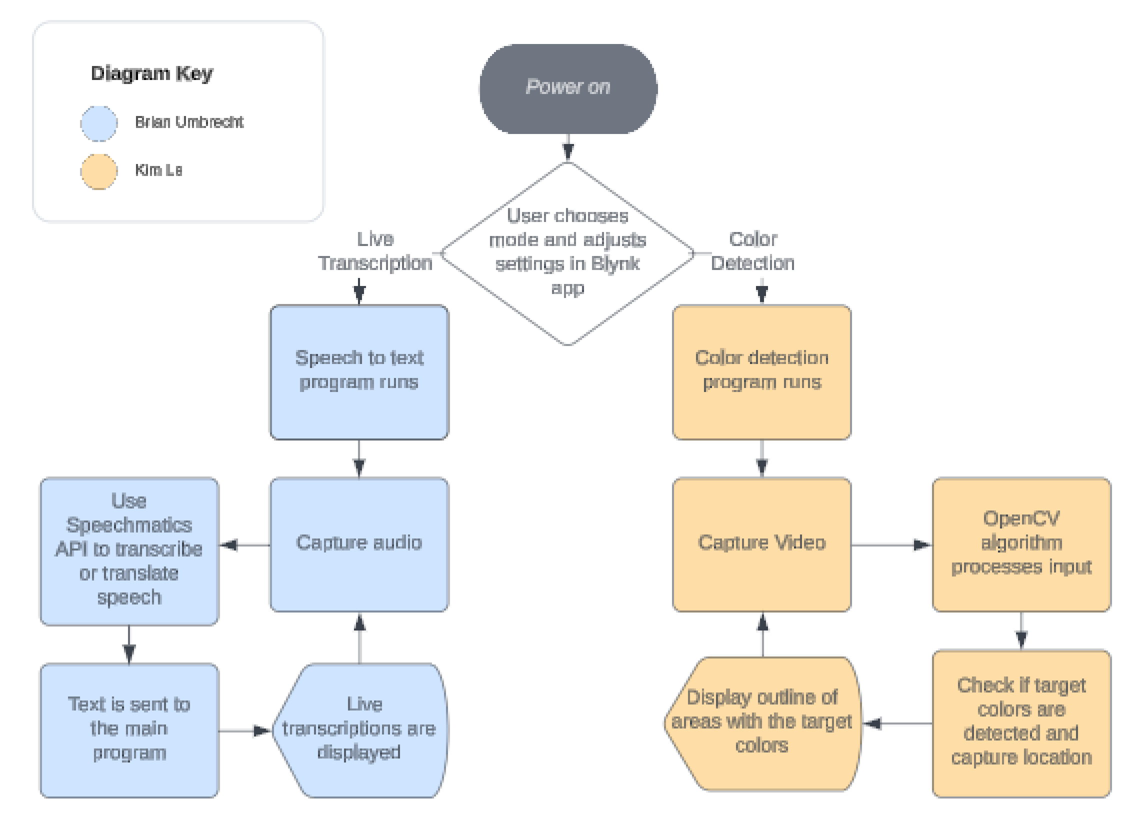
(1) Additional Support
Ensures security the beamsplitter remain in place

No ECE? No Problem!

- No PCB Requirement
- Simple Connections
  - Wireless communication with SBC
  - Direct connections from USB mic and replaceable camera/display.
- Simple framework for device
  - Optics housing
- Challenge: Power Supply...



Software Flowchart



Live Speech to Text

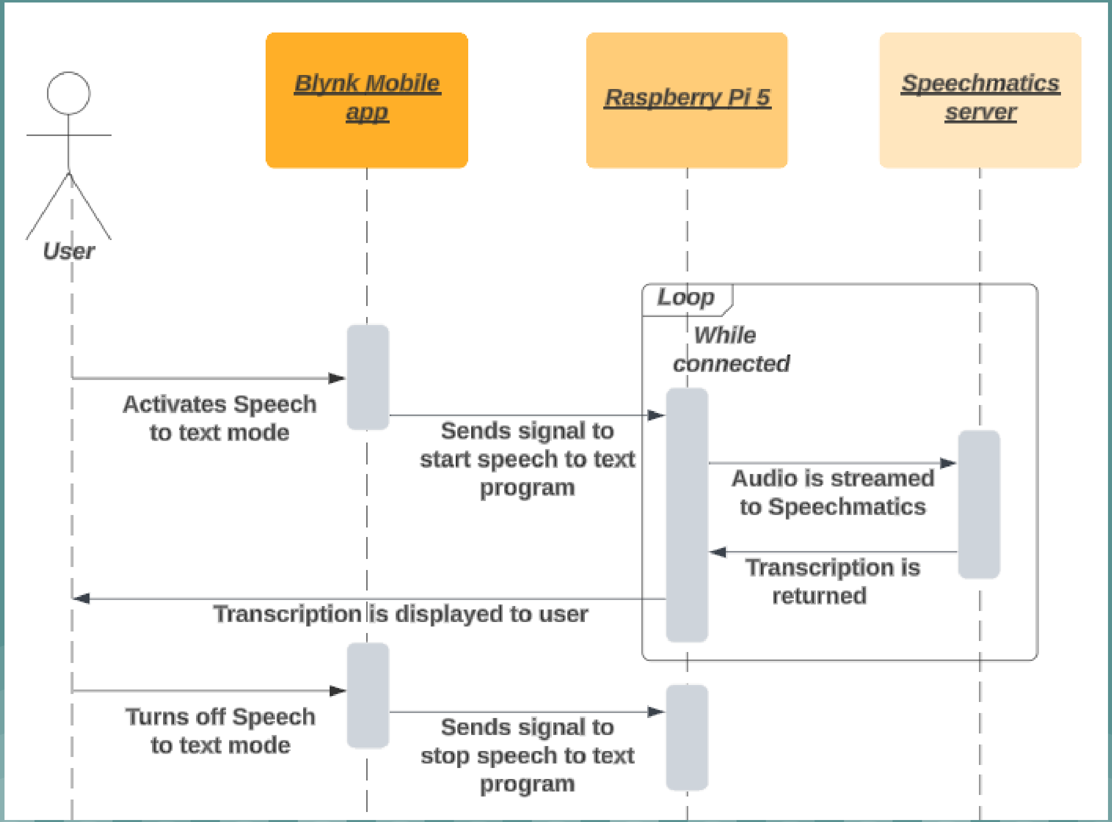
- We're using **Speechmatics** for transcribing speech.
- Speechmatics offers accurate transcriptions, multiple languages, translation, and a free trial.
- Different types of transcription; batch and real-time.
- We'll use real-time to allow transcribing speech as it comes in.
- Uses **WebSocket** API.
- Easy to implement in Python.

Transcription

- Our Raspberry Pi 5 use Speechmatics' API to stream audio to their server.
- Received text, is sent back to main program.
- Text is displayed.
- The user can activate this mode using the Blynk app.
- Only English can be transcribed.

- Program is implemented as a subprocess in a separate program.
- Integration issues; library compatibility, GPIO pins in use.
- Latency is <2s.

Activity Diagram



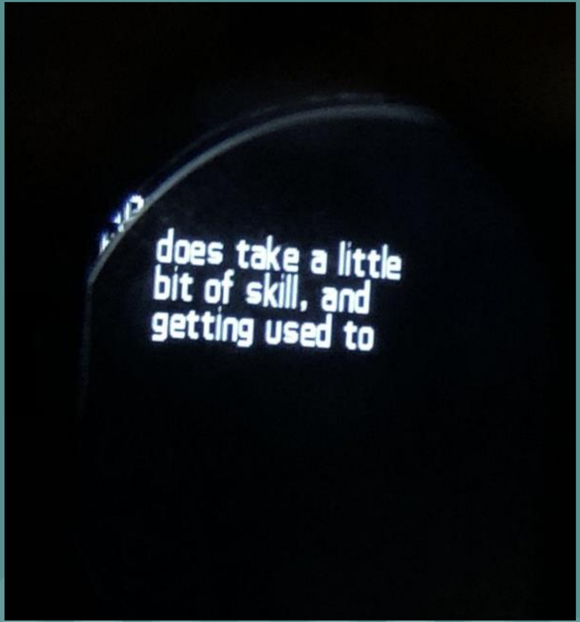
Translation

- Implementation is the same as in regular transcription.
- Over 30 languages.
- User chooses language and mode in Blynk app.
- Chosen language is passed as an argument to the subprocess.

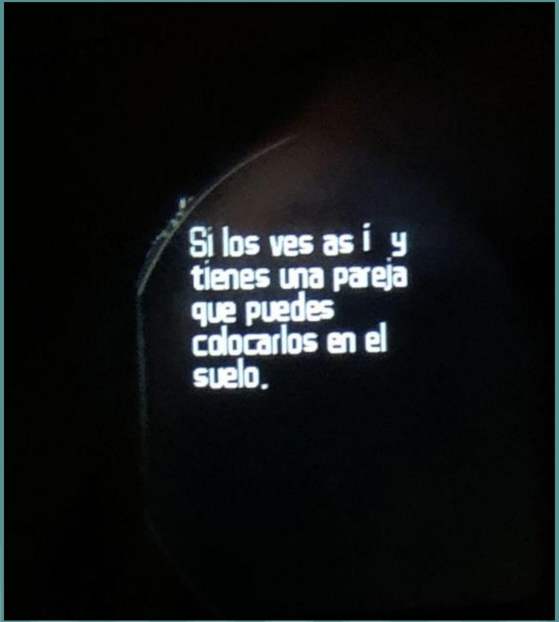
- Speechmatics only allows translation from English to other languages, and vice versa.
- Translations takes longer to start.
- Translations are often longer.

Text Display

English Transcription



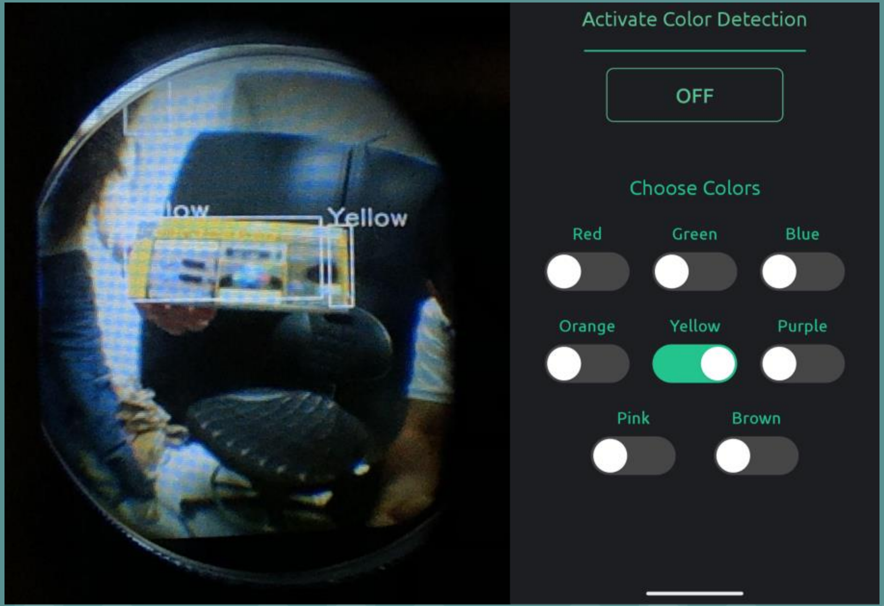
English Translation to Spanish



B

Live Color Detection

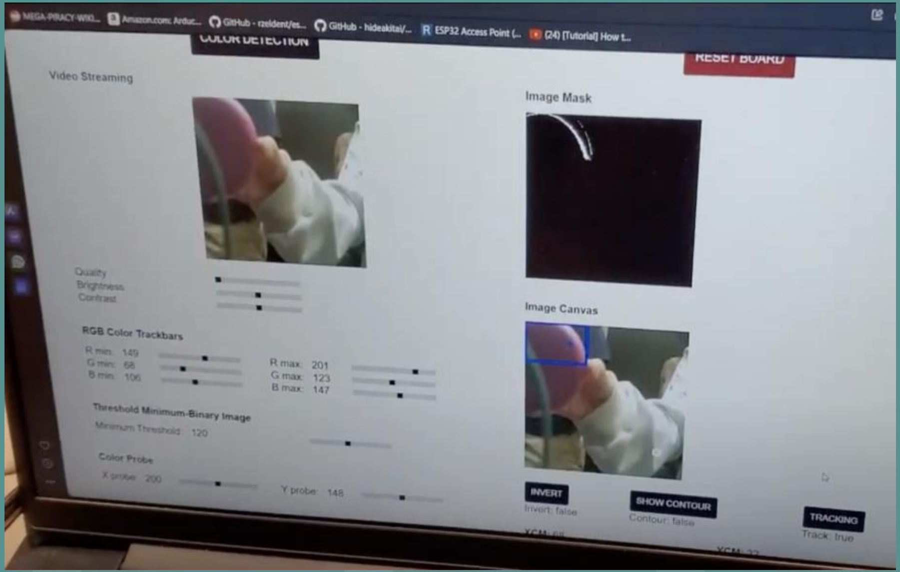
- **OpenCV** is a versatile computer vision library for image and video analysis.
- Designed to handle real-time processing of video streams and frames.
- Optimized for resource-constrained platforms like microcontrollers.
- Popular with a strong community and continuous support.



Live Color Detection

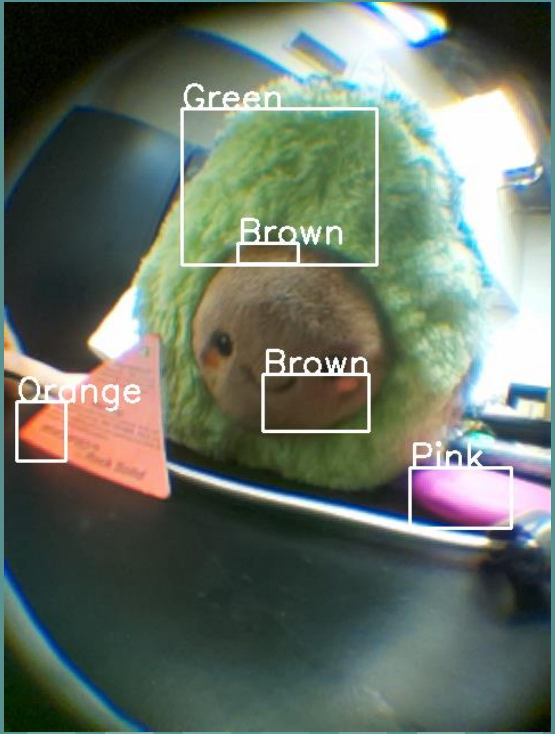
Challenges with ESP32 and OpenCV.js:

- Performance:
  - Less powerful
- Internet Dependency
- Limited Color Detection



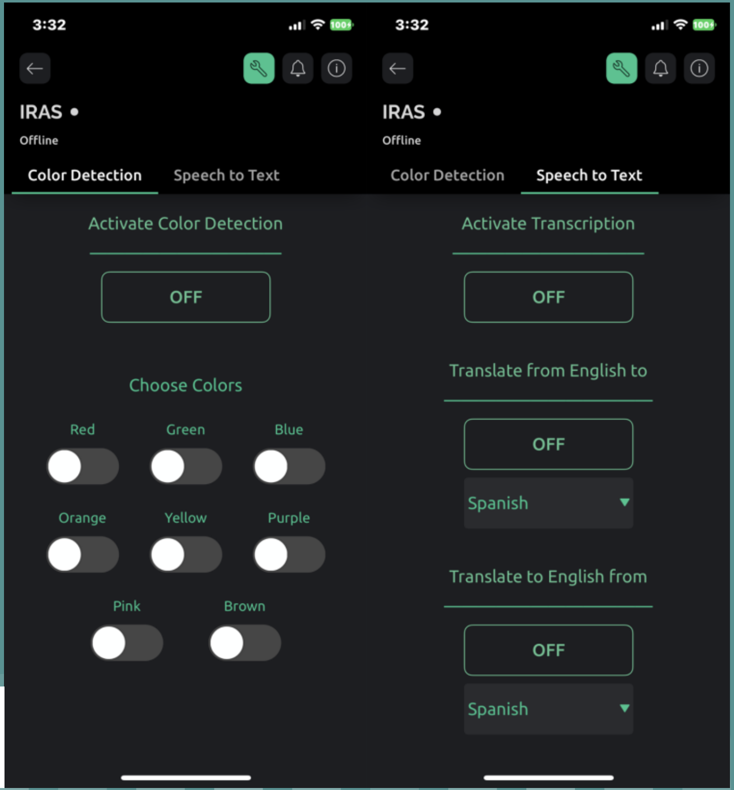
GUI Design

- Minimalist Design: Focuses on essential information and functions to avoid overwhelming users.
- Optimized Text Sizes: Carefully selected font sizes for easy readability.
- High Contrast for Readability: Ensures clear visibility under different lighting conditions.



Mobile App

- User-Friendly Interface:
  - Eliminates the need for physical buttons
- Extended Functionality:
  - Personalized settings
- Over-the-Air (OTA) Updates
- Current Mobile App Options:
  - Blynk, Ubidots, ThingSpeak, Cayenne, and MQTT Dash.
- Considerations for Future Development:
  - Custom Mobile App



Blynk

Work Distribution

- **Camera Lens Design**
  - Carlos Acosta
- **Near Eye Projection Design**
  - Darlandie Moise
- **Framework**
  - Carlos Acosta, Darlandie Moise
- **Live Transcription and Translation**
  - Brian Umbrecht
- **Live Color Detection**
  - Kim Le
- **Mobile app**
  - Brian Umbrecht, Kim Le

Administrative Content

Budget

- Raspberry Pi 5 Kit
  - $120
- OV5640 Camera
  - $25
- LCD Display
  - $6
- Frame
  - $10
- Optical Components
  - $150
- App
  - $7
- Velcro/Tape
  - $25
- Speechmatics
  - $40

Total: $383

Progress

- Purchased and received all components
  - 100% completed
- Color Detection
  - 97% completed
- Live Transcription / Translation
  - 95% completed
- Mobile App
  - 100% completed
- Camera Lens Design
  - 95% completed
- Near Eye Projection Design
  - 98% completed

Plans for Completion

- Finish testing device and implement changes
- Meetup to prepare for final presentations

Thanks for Watching!