We-Sign

An Intermediary for Hearing Impaired Based Communication

- Western Engineering Competition
- January 10, 2021
- ***(WINKLE):*** *Izzy Alfred, Liam Briggs, Natalie Connors, Keon Jin, Waleed Sawan, Edward Zurabov*

Presentation Overview

Problem Overview

- Problem Statement
- The Numbers
- Limitations
- Other Solutions

Solution

- Function/Purpose
- Demo
- Hardware
- Image Processing
- User Feedback/Response

Concluding Statements

- Profitability Projections
- Potential Marketing Plan
- Environmental & Social Ramifications
- Applications
- Next Steps

Presentation Overview

Problem Overview

- Problem Statement
- The Numbers
- Limitations
- Other Solutions

Solution

- Function/Purpose
- Demo
- Hardware
- Image Processing
- User Feedback/Response

Concluding Statements

- Profitability Projections
- Potential Marketing Plan
- Environmental & Social Ramifications
- Applications
- Next Steps

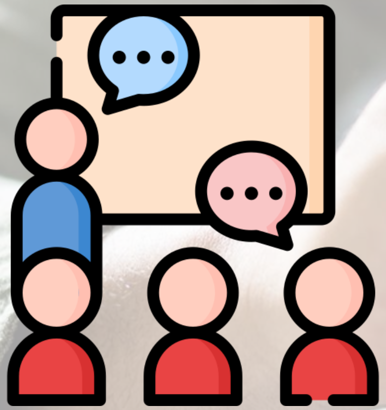
Problem Statement

Individuals with auditory disabilities face a universal challenge; their lack of audial clarity imposes natural limitations on them when in pursuit of education, employment, and communication.

Hearing Impairment Statistics

- Roughly 5% of the world's population have hearing loss
- 466 million people have hearing loss greater tan 40 dB

The Limitations



Education

Employment

Communication

Other Solutions

Interpreter

Hearing Aid

Cochlear Implant

Presentation Overview

Problem Overview

- Problem Statement
- The Numbers
- Limitations
- Other Solutions

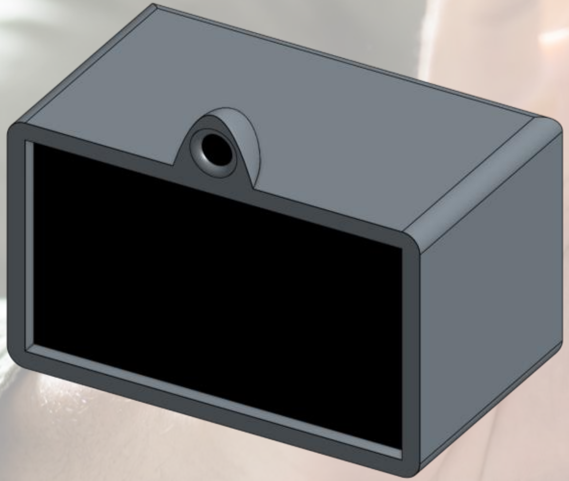
Solution

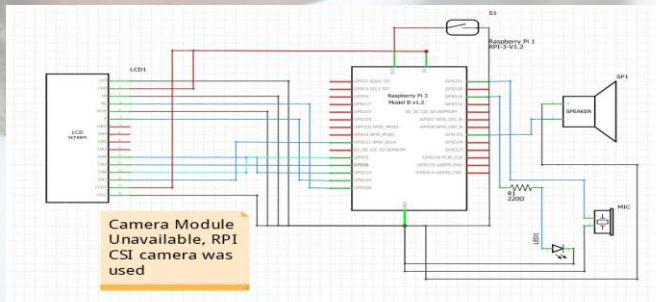
- Function/Purpose
- Demo
- Hardware
- Image Processing
- User Feedback/Response

Concluding Statements

- Profitability Projections
- Potential Marketing Plan
- Environmental & Social Ramifications
- Applications
- Next Steps

Function and Purpose



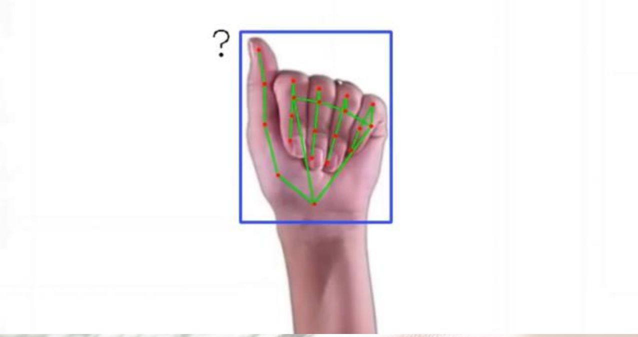


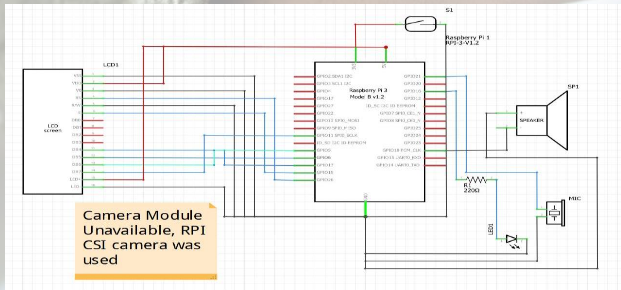
- Users with hearing impairment often face difficulty communicating with others
- Premium equipment for equipment cost incredible sums of money (e.g hearing aids)
- We-sign offers both ASL-to-speech and speech-to-text capabilities for fluid conversation

Architecture

Input → Processing → Response

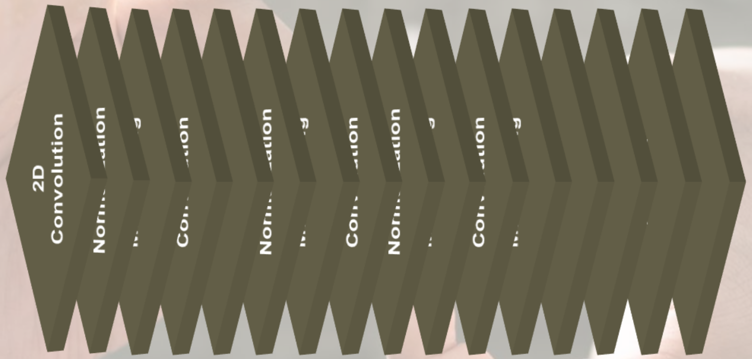
Input





Processing

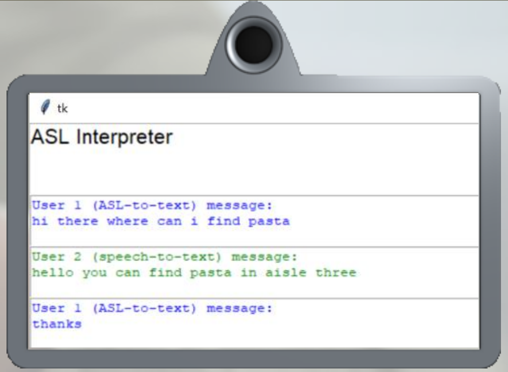
Convolutional Neural Network



- Image Processing
- ASL Data Set
- Word Inference

Response

Output

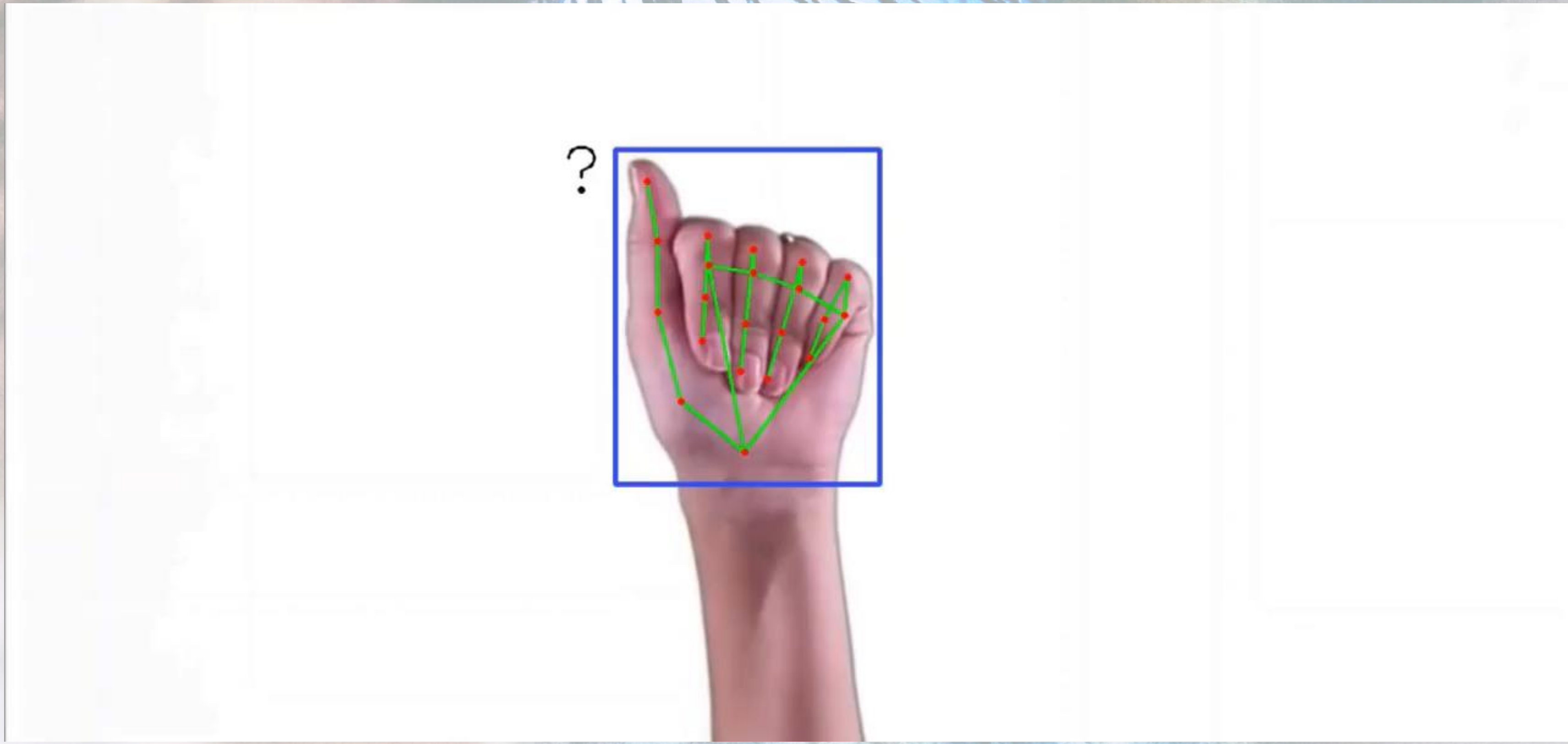


pyttsx3
Text-to-Speech

Voice Recognition

Check out the Video in the README to see this

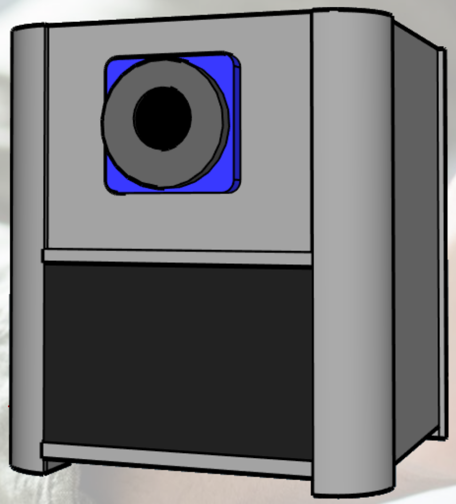
Visual Demonstration



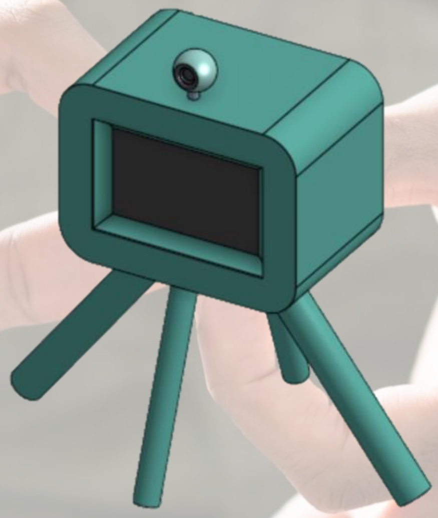
- Hardware
- Image Processing
- User Feedback/Response

Hardware

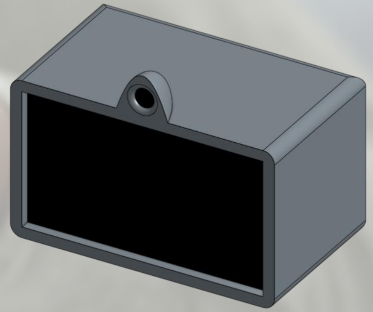
Packaging



Iteration #1



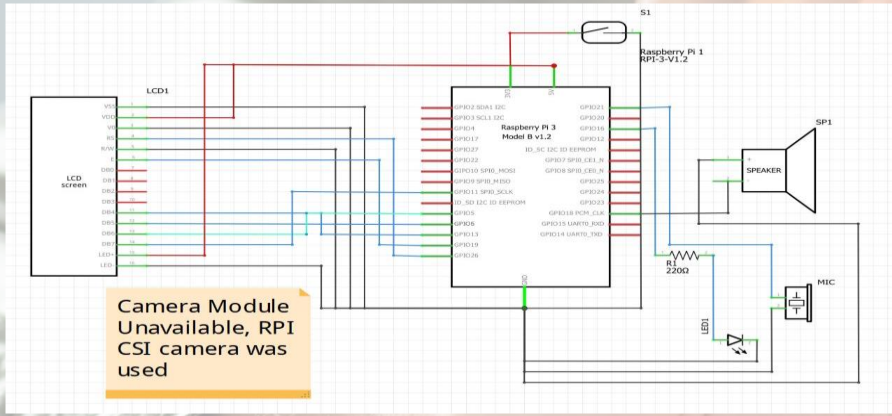
Iteration #2



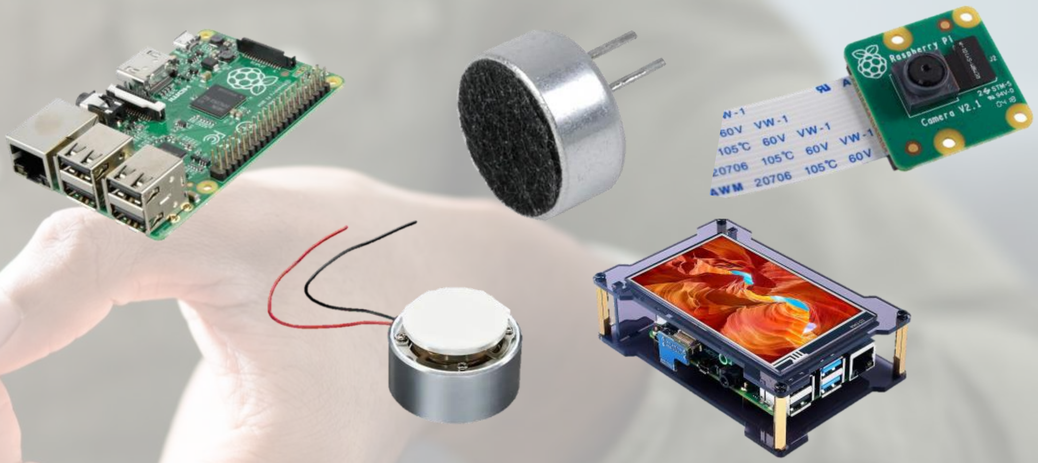
Iteration #3

Hardware

Schematic



Components



Pricing Chart

	Cost of Production	
Item	Prototyping (Source: Amazon)	Manufactering (Source: Alibaba)
Raspberry Pi	$99.99	$31.00
Power Supply Adapter	$11.85	$1.02
4-Inch LED Screen	$44.99	$23.28
Microphone	$11.00	$2.56
Camera	$39.99	$8.77
Packaging	$15.00	$5.67
Speaker	$24.71	$3.50
Total	**$247.53**	**$75.80**

$75.80

Prototype Cost: $247.53

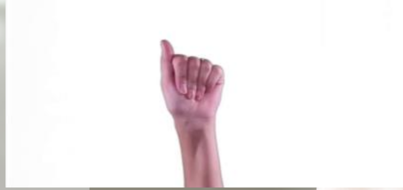
- Hardware
- Image Processing
- User Feedback/Response

Image Processing

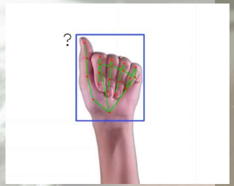
Our Approach

Pre-Classification Image Processing

Read in Input



Hand Location



Background Subtraction

Image Crop & Grayscale

Neural Network

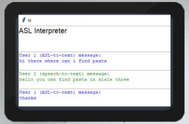
Conv. Neural Network

Image Separation

Feedback

Word Inference

Text to Speech



Speech to Text

Image Processing

Motion-Based Background Subtractor

Our Approach:

- We further advanced the device's image processing capabilities to distinguish hands from its environment using background subtractor techniques
- This was done by extracting the moving foreground (hand) from the static background while minimizing the detection of interference/noise

Our Implementation:

- OpenCV, a library that houses multiple image processing functions, was used to implement background subtraction

Visual Demonstration of Background Subtraction Capabilities

Image Processing

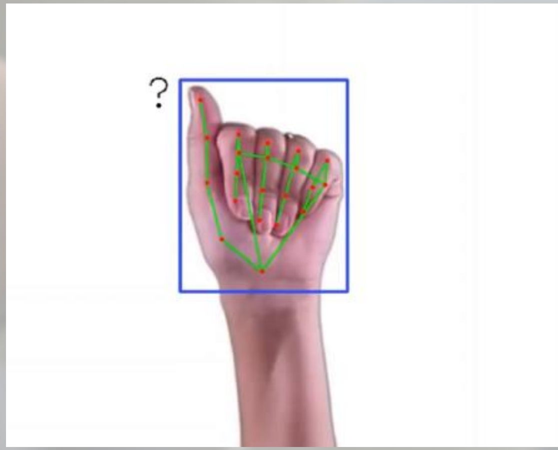
Hand Tracking

Our Approach:

- The need to locate where exactly the hand is in the frame is obviously crucial to ensure that the algorithm is successful
- Locating the hand was done by plotting the location of parts of the hand and using a set to determine a box around

Our Implementation:

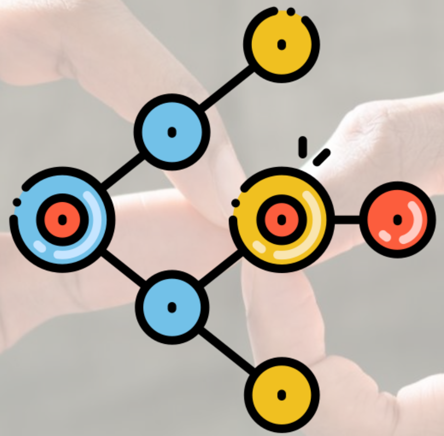
- MediaPipe, Google's popular heavily customized ML library, allows us to plot exactly the hand placement
- Additionally, this implementation allows us to expand into more complex phrases in ASL in the future



Visual Demonstration of the ML hand tracking

Conv. Neural Network

Training Data (MINST)



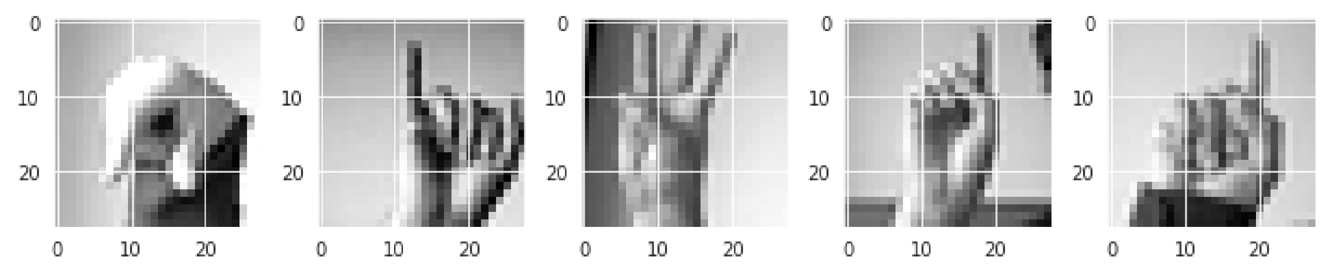
Sequential Layers

Model Accuracy

Conv. Neural Network

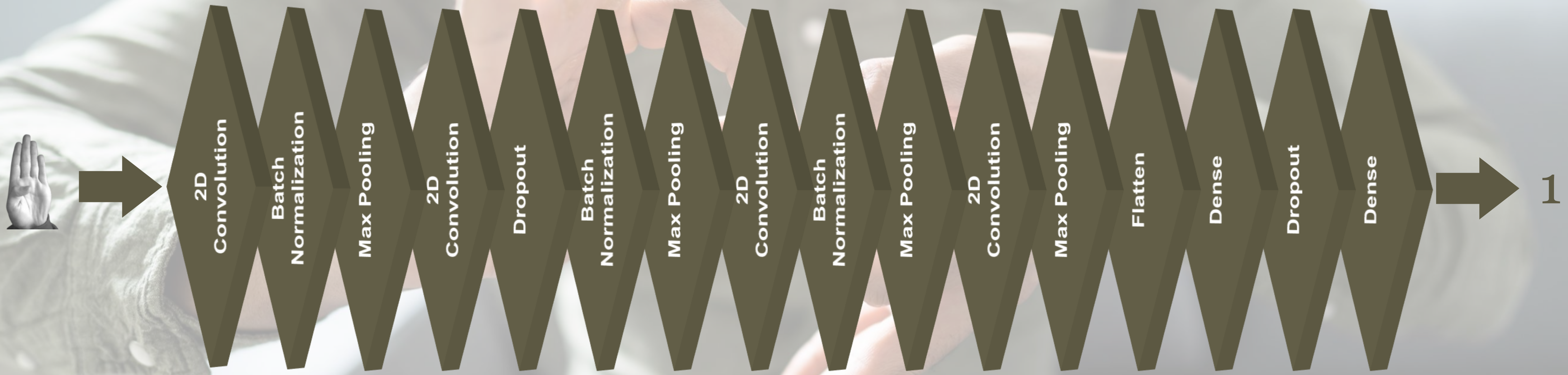
Training Data

- Used a publicly available ASL training dataset (Sign Language MINST)
- Over 27,000 of training images, 7,000 of testing
- Representing multiple people repeating the standard gestures against different backgrounds



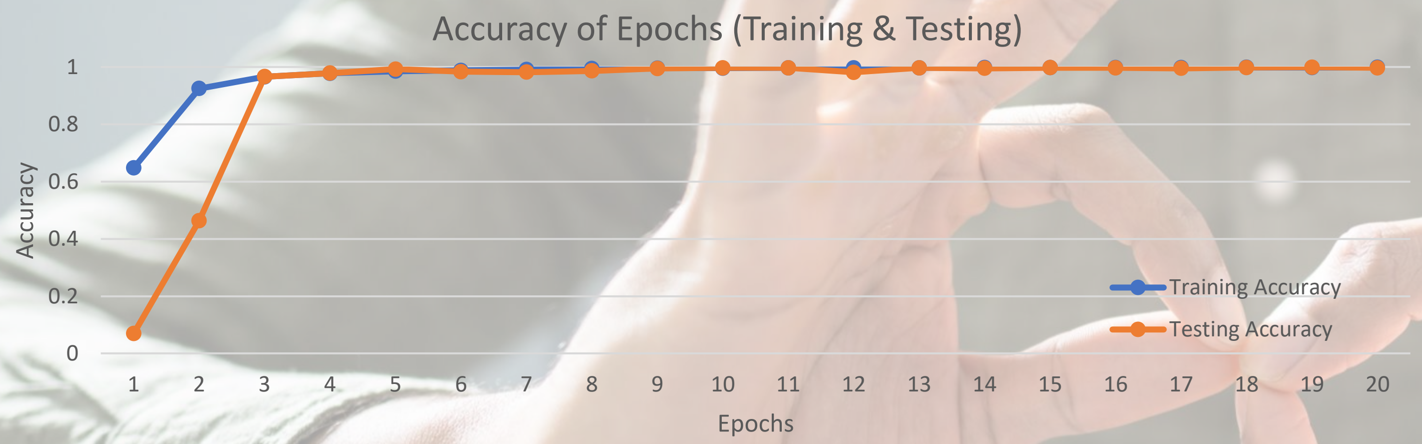
Conv. Neural Network

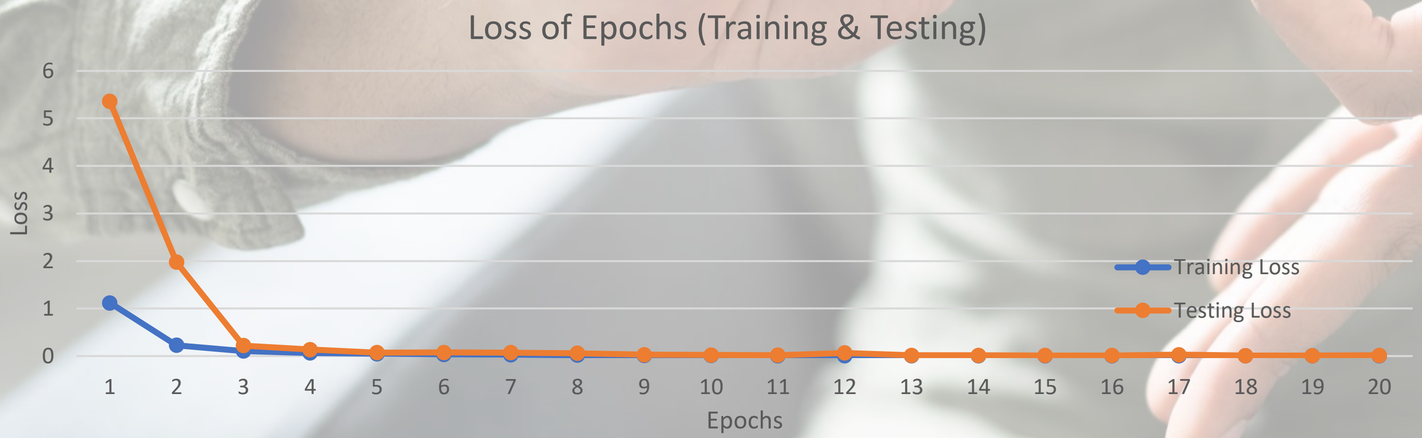
Sequential Layers

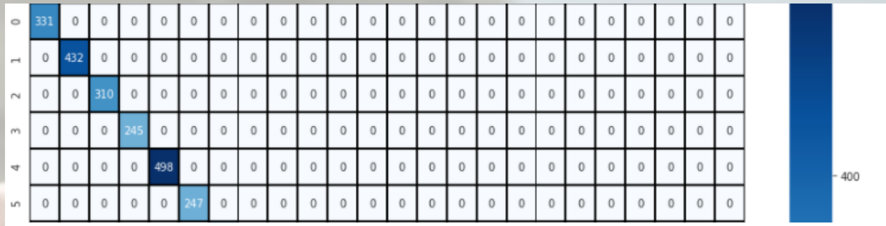


Conv. Neural Network

Model Accuracy

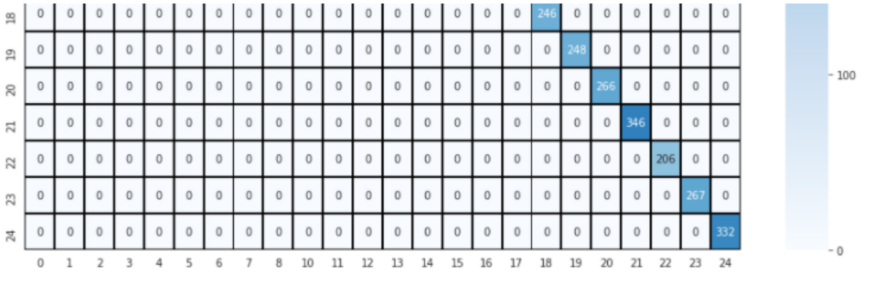






100%

7172/7172
of test cases



- Hardware
- Image Processing
- User Feedback/Response

User Feedback/Response

Word Inference

Character String Transformation

Our Approach:

- The ASL Frame classifier returns a continuous string of characters – string meaning cannot be determined without inferring spaces

Our Implementation:

- Use Zipf's Law – which determines the probability of a word occurrence based on frequency – to construct a cost dictionary of words

- Find lowest-cost string by backtracking cost arrays of different word possibilities

```python
from math import log

# Build a cost dictionary, assuming Zipf's law and cost = -math.log(probability).
words = open("word-frequency.txt").read().split()
wordcost = dict((k, log((i+1)*log(len(words)))) for i,k in enumerate(words))
maxword = max(len(x) for x in words)

def stringToWords(s):
    # dynamic programming to infer location of spaces in string

    # find best match, given for i characters given i-1 cost
    # returns a pair (match_cost, match_length).
    def best_match(i):
        candidates = enumerate(reversed(cost[max(0, i-maxword):i]))
        return min((c + wordcost.get(s[i-k-1:i], 9e999), k+1) for k,c in candidates)

    # build the cost array.
    cost = [0]
    for i in range(1,len(s)+1):
        c,k = best_match(i)
        cost.append(c)

    # backtrack to recover the minimal-cost string.
    out = []
    i = len(s)
    while i>0:
        c,k = best_match(i)
        assert c == cost[i]
        out.append(s[i-k:i])
        i -= k

    return " ".join(reversed(out))
```

Word inference from ASL character string code component

Text-to-Speech

Character String Transformation

Our Approach:

- A text-to-speech component from user ASL input encourages seamless and convenient communication between two parties

Our Implementation:

- Pyttsx3 – a text-to-speech Python library – is implemented to express user ASL input in voice-form

pyttsx3 2.90

```python
import pyttsx3

# initialize engine
engine = pyttsx3.init()

# set properties
engine.setProperty('rate', 150) # speed percent
engine.setProperty('volume', 0.9) # Volume 0-1

# Text-to-speech function
def speak(s):
    engine.say(s)
    engine.runAndWait()
```

Text to Speech Component

Voice Recognition

User Interface

Our Approach:

- Main criteria in designing the user interface consisted of enabling a seamless interaction between the two users
- A speech-to-text peripheral enables the device to be efficiently used and avoids interference with ongoing tasks

Our Implementation:

- Using the speech-to-text Google Cloud API, our device can accurately pick up on verbal phrases to be displayed on the screen towards the impaired user

Google Cloud

```
import speech_recognition as sr

r = sr.Recognizer()
speech = sr.Microphone(device_index=1)
with speech as source:
    print("say something!…")
    audio = r.adjust_for_ambient_noise(source)
    audio = r.listen(source)
try:
    recog = r.recognize_google(audio, language = 'en-US')

    print("You said: " + recog)
except sr.UnknownValueError:
    print("Google Speech Recognition could not understand audio")
except sr.RequestError as e:
    print("Could not request results from Google Speech Recognition service; {0}".format(e))
```

Speech to Text Component

Presentation Overview

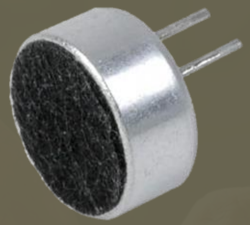
Problem Overview | Solution | Concluding Statements

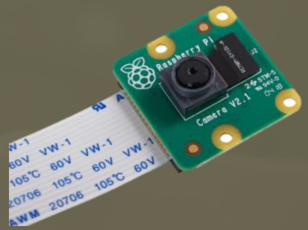
Problem Overview
- Problem Statement
- The Numbers
- Limitations
- Other Solutions

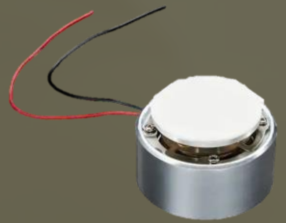
Solution
- Function/Purpose
- Demo
- Hardware
- Image Processing
- User Feedback/Response

Concluding Statements
- Profitability Projections
- Potential Marketing Plan
- Environmental & Social Ramifications
- Applications
- Next Steps

Cost Breakdown







Key Components

	Cost of Production	
Item	Prototyping (Source: Amazon)	Manufactering (Source: Alibaba)
Raspberry Pi	$99.99	$31.00
Power Supply Adapter	$11.85	$1.02
4-Inch LED Screen	$44.99	$23.28
Microphone	$11.00	$2.56
Camera	$39.99	$8.77
Packaging	$15.00	$5.67
Speaker	$24.71	$3.50
Total	**$247.53**	**$75.80**

Manufacturing:
$75.80/Unit
Prototyping:
$247.53/Unit

Projected Profitability

We-Sign Inc. Projected Income Statement For Years Ending 2021 & 2022			
Item	**Assumption**	**2021**	**2022**
Revenue			
Net Sales	2021: Sold 10 Units 2022: Sold 50 Units @ $300/Unit	$3,000	$15,000
COGS	$75.80/Unit (See Cost Breakdown of Manufactering)	$758	$3,790
Gross Profit		$2,242	$11,210
Expenses			
General & Administrative Expense	5% of Net Sales	(150)	(750)
Utilities Expense	3% of Net Sales	(90)	(450)
Advertisment Expense	30% of Net Sales	(900)	(4,500)
Research & Development Expense	10% of Net Sales (Continuous Prototyping)	(300)	(1,500)
Total Expenses		(1,440)	(7,200)
Income Before Tax		$802	$4,010
Income Tax	30%	(241)	(1,203)
Net Income		$561	$2,807

Key Assumptions:

Unit Price: Approximately manufacturing cost per unit x4

Units Sold: Based on potential partnerships with small-medium institutions

Expenses: Small portions of revenue would be allocated to basic operating expenses, with an additional large focus on advertisement

Income Tax: A conservative yet common rate placed on most corporate entities

Marketing Plan

Price:
$300/Unit

Promotion/Growth:
Partnerships with Franchises

Target Market:
Retail and Service Centres

Placement:
B2B Distribution

Product:
Seamless Use and Effectiveness

Environmental & Social Ramifications

Production

Delivery

End of Useful Life

Mitigations

Environmental Mitigations

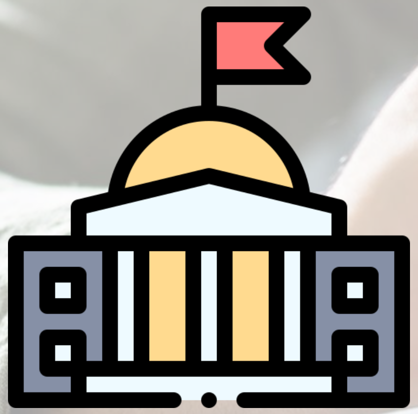
- Bringing production processes to Canada specifically in locations that use primarily renewable energy, this would also simultaneously lower We-sign's transportation/shipping footprint
- Create a safe disposal program to deal with the hardware after it has reached its useful life

Social Mitigations

- Work with associations such as the Ontario Association of the Deaf to determine those that need the device the most

ONTARIO ASSOCIATION OF THE DEAF

Applications



Government Service Centers

Retail Stores

Education

Next Steps

Expanded Functionality:

The integration of the ML hand detection algorithm to enables a unique opportunity to increase the functionality to include more mobile phrases

Educational Capabilities:

Implement a learning program alongside the device so that education providers can use the tool to promote ASL learning

We-Sign

An Intermediary for Hearing Impaired Based Communication

- Western Engineering Competition
- January 10, 2021
- ***(WINKLE):*** *Izzy Alfred, Liam Briggs, Natalie Connors, Keon Jin, Waleed Sawan, Edward Zurabov*

References

[1] "Deafness and hearing loss", *Who.int*, 2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss. [Accessed: 08- Jan- 2021].

[2] "How Deaf Children Are Being Locked Out of Language", *OZY*, 2021. [Online]. Available: https://www.ozy.com/news-and-politics/how-deaf-children-are-being-locked-out-of-language/89643/. [Accessed: 08- Jan- 2021].

[3] *Deafjobwizard.com*, 2021. [Online]. Available: https://www.deafjobwizard.com/post/unemployment-in-the-deaf-community-barriers-recommendations-and-benefits-of-hiring-deaf-employees. [Accessed: 08- Jan- 2021].

[4] *Mayo Clinic*. [Online]. Available: https://www.mayoclinic.org/tests-procedures/cochlear-implants/about/pac-20385021 [Accessed: 08-Jan-2021].

[5] "Hearing Aids," *National Institute of Deafness and Other Communication Disorders*, 14-Dec-2020. [Online]. Available: https://www.nidcd.nih.gov/health/hearing-aids. [Accessed: 08-Jan-2021].

[7] B. Garcia and S. A. Viesca, "Real-time American Sign Language Recognition with Convolutional Neural Networks," *Stanford University*, 27-Nov-2012. [Online]. Available: http://cs231n.stanford.edu/reports/2016/pdfs/214_Report.pdf. [Accessed: 08-Jan-2021].

[8] K. Suri and R. Gupta, "Convolutional Neural Network Array for Sign Language Recognition Using Wearable IMUs," *2019 6th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2019.

[9] S. Mathur and P. Sharma, "Sign Language Gesture Recognition using Zernike Moments and DTW," *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, 2018.

[10] "Sign Language Recognition using Hybrid Neural Networks," *International Journal of Innovative Technology and Exploring Engineering Regular Issue*, vol. 9, no. 2, pp. 1092–1098, 2019.

[11] "Find quality Manufacturers, Suppliers, Exporters, Importers, Buyers, Wholesalers, Products and Trade Leads from our award-winning International Trade Site. Import & Export on alibaba.com," *Alibaba*. [Online]. Available: https://www.alibaba.com/. [Accessed: 08-Jan-2021].

[12] "Projecting Income Statement Line Items - Step by Step Guide," *Corporate Finance Institute*, 08-Apr-2020. [Online]. Available: https://corporatefinanceinstitute.com/resources/knowledge/modeling/projecting-income-statement-line-items/. [Accessed: 08-Jan-2021].

[13] Tecperson, "Sign Language MNIST," *Kaggle*, 20-Oct-2017. [Online]. Available: https://www.kaggle.com/datamunge/sign-language-mnist. [Accessed: 08-Jan-2021].

[14] "Flaticon, the largest database of free vector icons," *Flaticon*. [Online]. Available: https://www.flaticon.com/. [Accessed: 08-Jan-2021].