

Target Audience

- **Staff of SA Deaf**
- Supporting their overall well-being
- Providing educational and training opportunities
- Advocating for their inclusion in society

They will achieve this through various means, such as:

- Accepting donations
- Fundraising activities
- Publishing educational materials
- Partnering with other organizations
- Acquiring property to support their mission

*The Singapore
Association
For The Deaf*



D



Silent Echoes: From Hand Waves to Written Phrases

GA-DSI-42 Timothy Chan

E



Agenda

01. Introduction

02. Exploratory Data Analysis

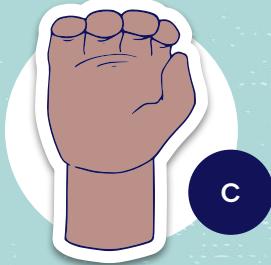
03. Modeling

04. Results

05. Limitations

A

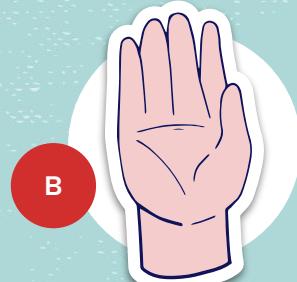




c

01.

Introduction



b

Sign Languages



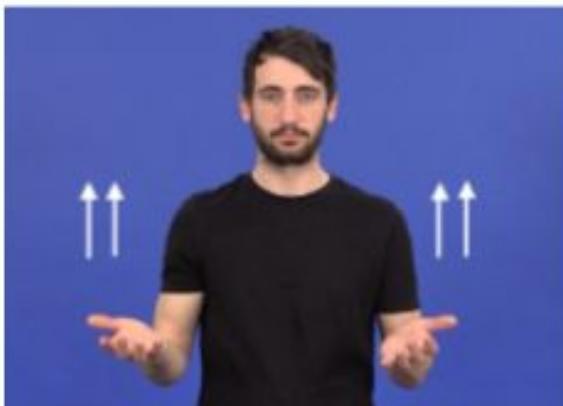
- A World of Signs: Over 300 unique sign languages exist, reflecting regional and cultural diversity
- Beyond Words: Sign languages use hand gestures to convey meaning, while facial expressions and eye movements add emotional depth



“My cat speaks sign language with her tail.”

—Robert A.M. Stern

Australia Sign Language



'not heavy'



'not dark'



'source of light'

Australian Sign Language translations for the English word "light"

American Sign Language



Light (weight) YT:
Signs



Light (illumination) YT:
Signs



Light On(illumination)
YT: Signs

Related Work 1



Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 216 (2023) 13–20



7th International Conference on Computer Science and Computational Intelligence 2022

Sign language recognition system for communicating to people
with disabilities

Yulius Obi^a, Kent Samuel Claudio^a, Vetri Marvel Budiman^a, Said Achmad^{a,*},
Aditya Kurniawan^{a,*}

^aComputer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, 11480,
Indonesia

Related Work 1



Related Work 2

Article | [Open access](#) | Published: 09 October 2023

Sign language recognition using the fusion of image and hand landmarks through multi-headed convolutional neural network

[Refat Khan Pathan](#), [Munmun Biswas](#), [Suraiya Yasmin](#), [Mayeen Uddin Khandaker](#)✉, [Mohammad Salman](#) & [Ahmed A. F. Youssef](#)

Scientific Reports **13**, Article number: 16975 (2023) | [Cite this article](#)

13k Accesses | **6** Citations | [Metrics](#)

Related Work 2



Article

**Sign
and
con**

Refat K

Ahmed

Scientif

13k Ac

e

han &

Related Work 3

Sign Language Recognition with Advanced Computer Vision

Detecting Sign Language Characters in Real Time Using MediaPipe and Keras



Mihir Garimella · Follow

Published in Towards Data Science · 8 min read · Aug 23, 2022



88



2



Related Work 3

Sign
Adv

Detecting
and Ker



88



Current Gap



- There are more models available publicly that uses alphabets to train the model
- However, words are not always spelt
- Instead an action is tagged to a word

PROBLEM STATEMENT

Creating an effective communication system between a deaf individual who uses sign language and a non-signing person presents significant challenges. The initial goal of this project is to lay the foundation for such a system by developing a solution that can accurately recognize and translate a select set of key signs – currently three common words – into written language.

This project aims to demonstrate the feasibility of scaling this technology to enable more comprehensive interactions and foster understanding between signing and non-signing individuals.



Leon Ng, 28, I.T Infrastructure Engineer

Attitudes & Behaviours

- Usually works overtime
- Mostly dine out due to lack of time
- Soft-spoken and kind-hearted
- Volunteers at a dog shelter during his weekends

Pain points

- Does not know how to respond in awkward and uncomfortable scenarios
- Sometimes wishes he can do something about
- Lack of available time for himself for commitments

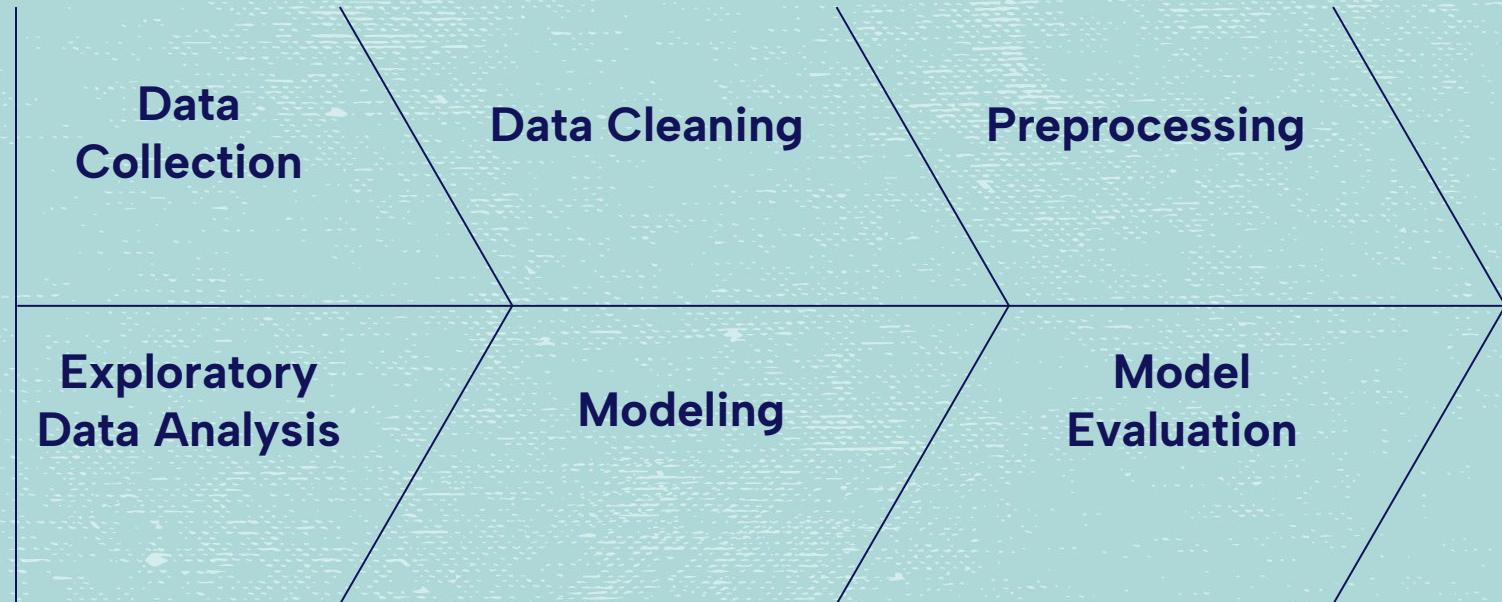
Scenario

- Dine at a new Vietnamese Restaurant at Somerset
- Did not realise it was operated by the hearing impaired

Motivation

- Liked the Vietnamese restaurant
- Watched YouTube Videos on sign languages

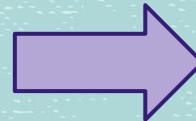
Methodology



Data Collection

Videos from:

- Microsoft American Sign Language
- World Language American Sign Language (videos included)



Video download

- Downloaded videos from Youtube
- A fraction of the videos are unavailable
- Entire length of the video

Preprocessing

Image Cleaning

- **Focus on Relevant Actions:** Crop videos to isolate key movements
- **Ensure Consistent Processing:** Resize videos to a uniform format (512x512 @ 30fps)
- **Enhance Data:** Apply video augmentation techniques to create a more robust dataset



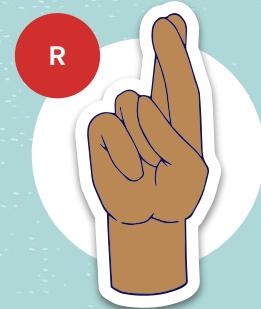
Extract Keypoints

- **Extract Meaningful Data:** Leverage MediaPipe to extract keypoint coordinates, representing crucial locations in the video
- **Isolating Finger Movements:** Focus on specific finger points to gain insights into hand gestures

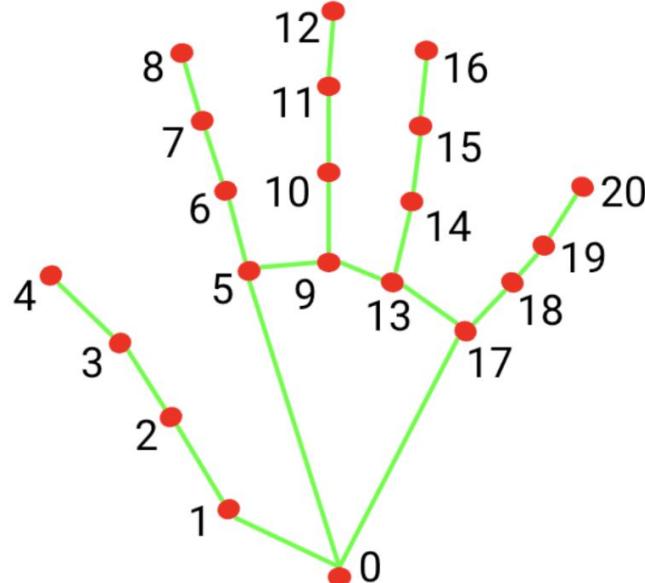


02.

Exploratory Data Analysis



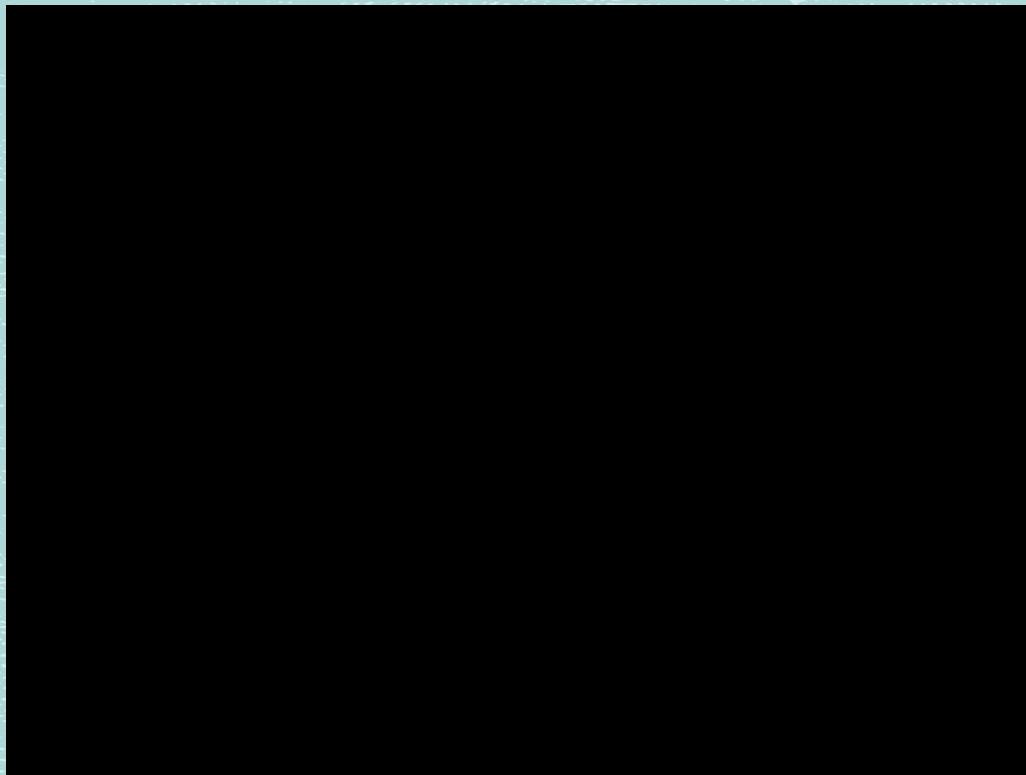
Hand Landmarks



- | | |
|-----------------------|-----------------------|
| 0. WRIST | 11. MIDDLE_FINGER_DIP |
| 1. THUMB_CMC | 12. MIDDLE_FINGER_TIP |
| 2. THUMB_MCP | 13. RING_FINGER_MCP |
| 3. THUMB_IP | 14. RING_FINGER_PIP |
| 4. THUMB_TIP | 15. RING_FINGER_DIP |
| 5. INDEX_FINGER_MCP | 16. RING_FINGER_TIP |
| 6. INDEX_FINGER_PIP | 17. PINKY_MCP |
| 7. INDEX_FINGER_DIP | 18. PINKY_PIP |
| 8. INDEX_FINGER_TIP | 19. PINKY_DIP |
| 9. MIDDLE_FINGER_MCP | 20. PINKY_TIP |
| 10. MIDDLE_FINGER_PIP | |



'Hello' There





'Hello' There

Frame 29



Frame 30



Frame 31

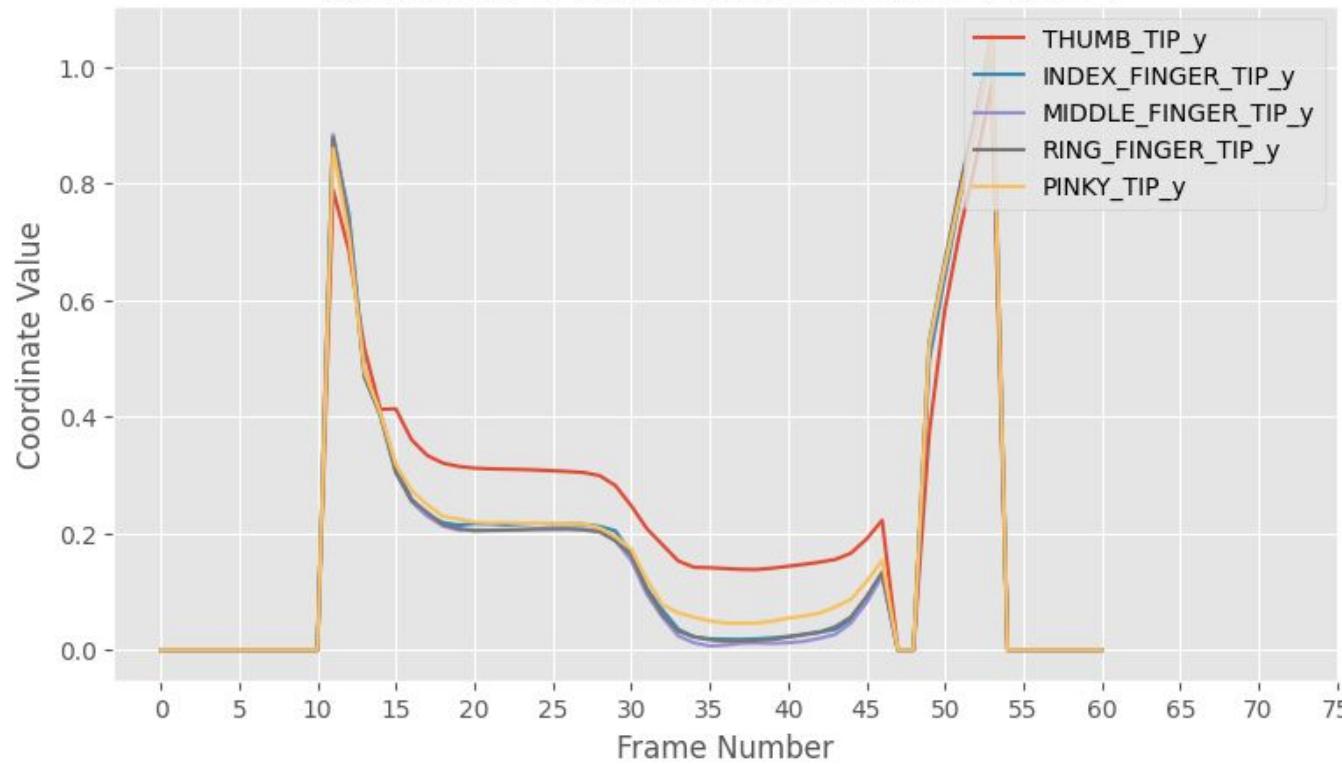


Frame 32



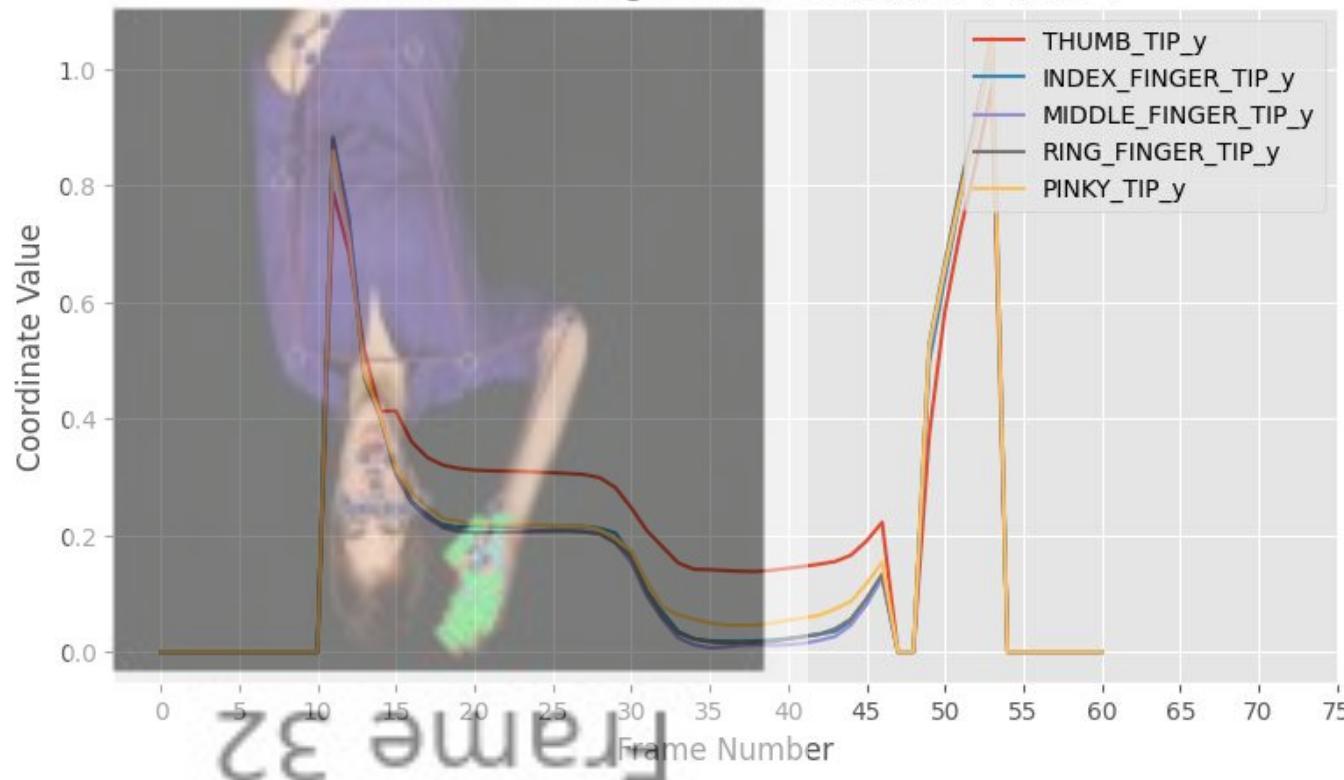
'Hello' There

Movement of Fingers Across Frames ("hello")



'Hello' There

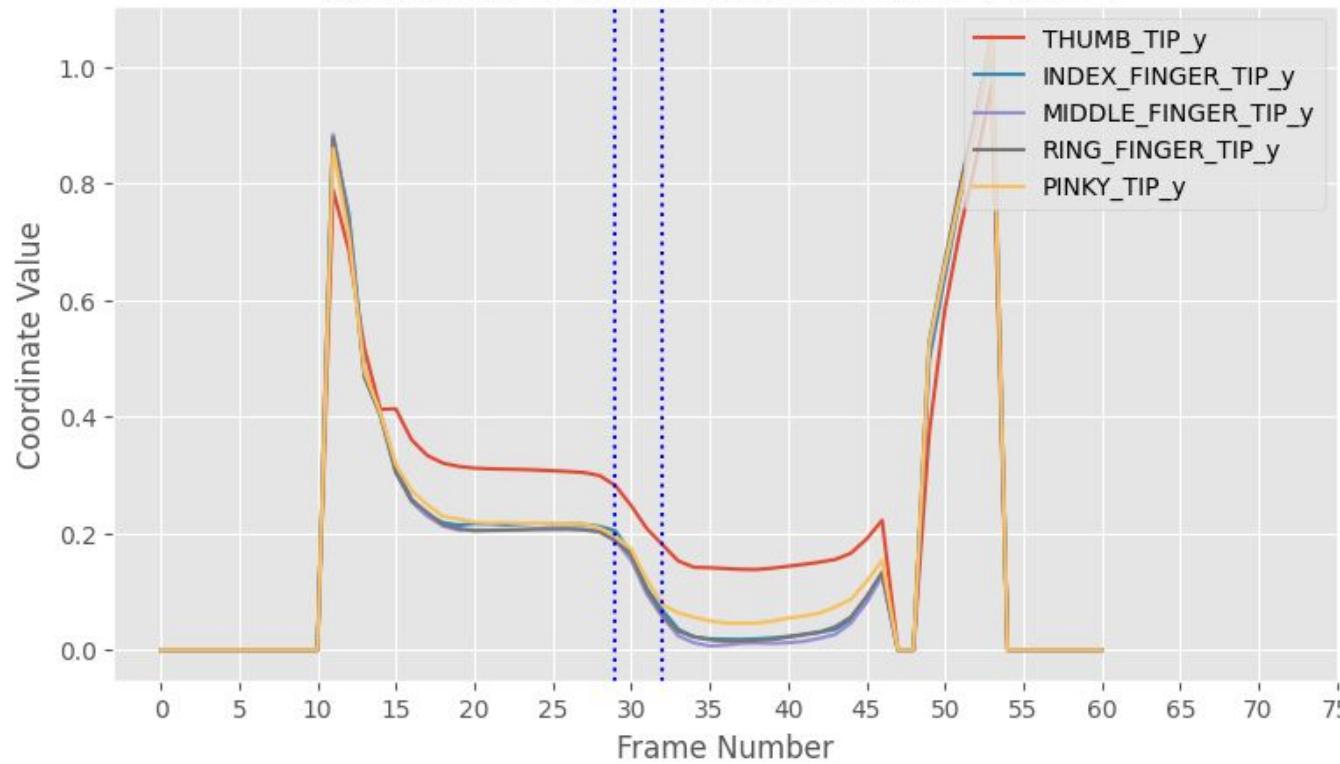
Movement of Fingers Across Frames ("hello")



Frame 32

'Hello' There

Movement of Fingers Across Frames ("hello")





'Hello' There

Frame 29



Frame 30



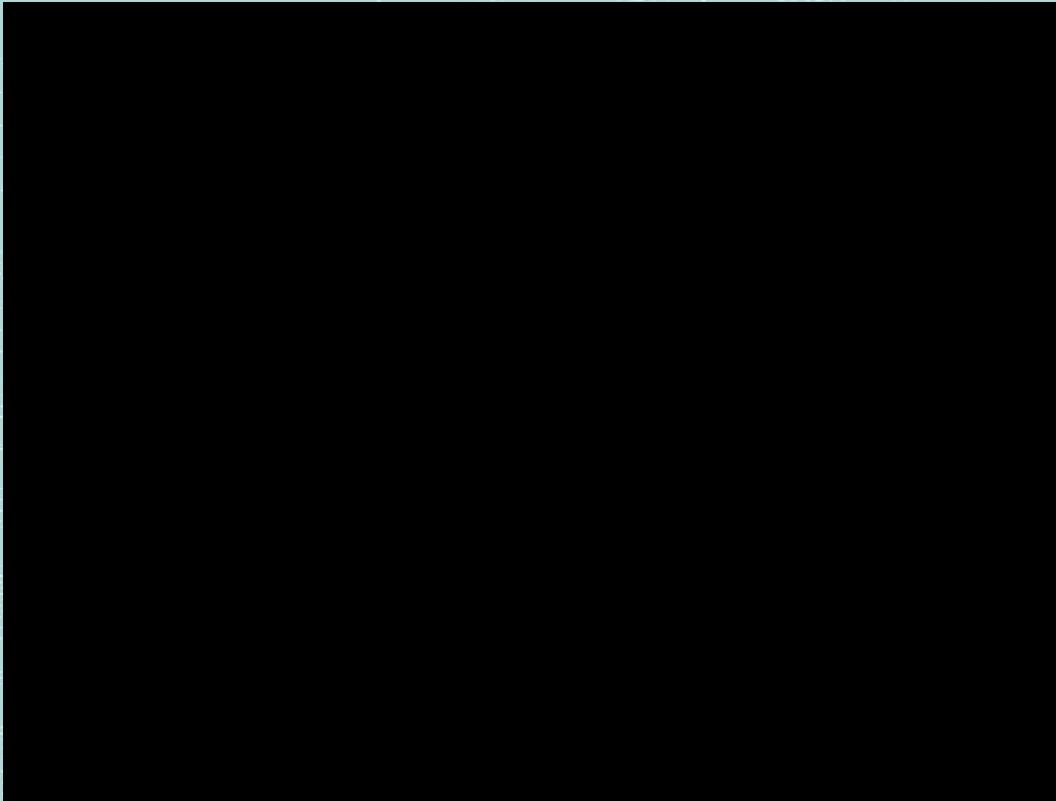
Frame 31



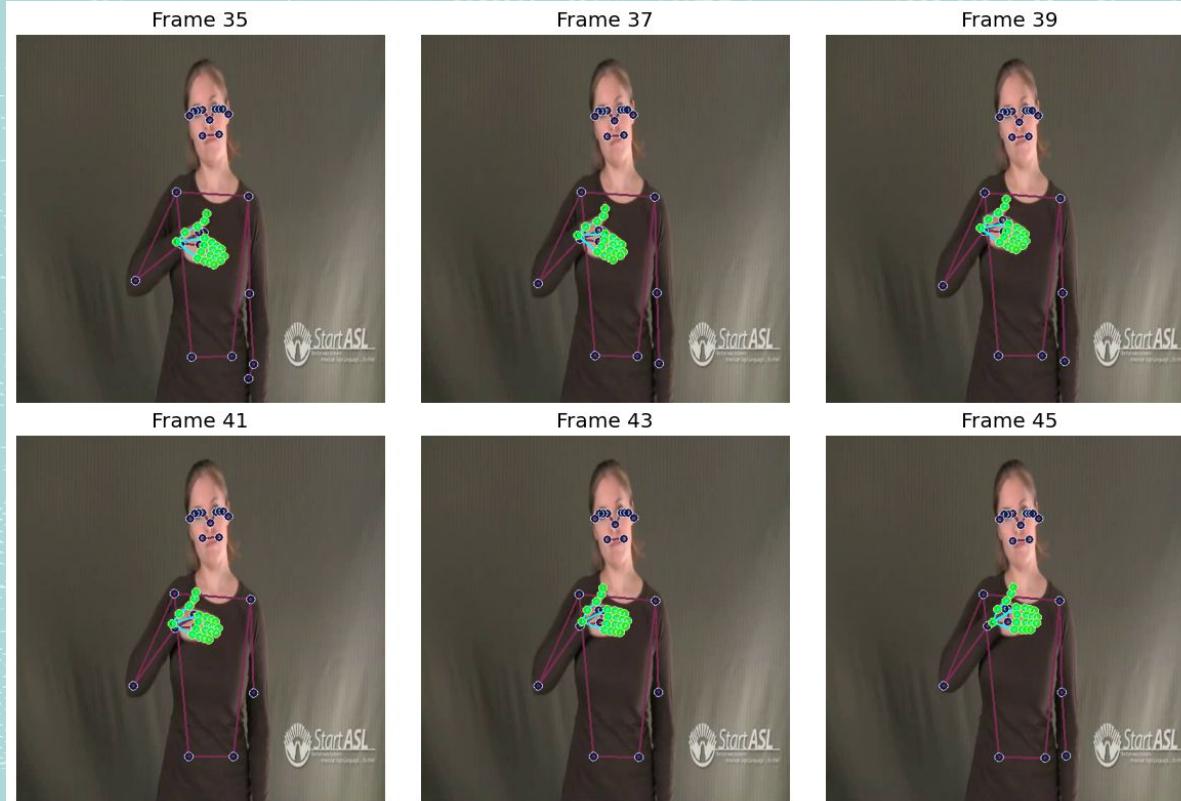
Frame 32



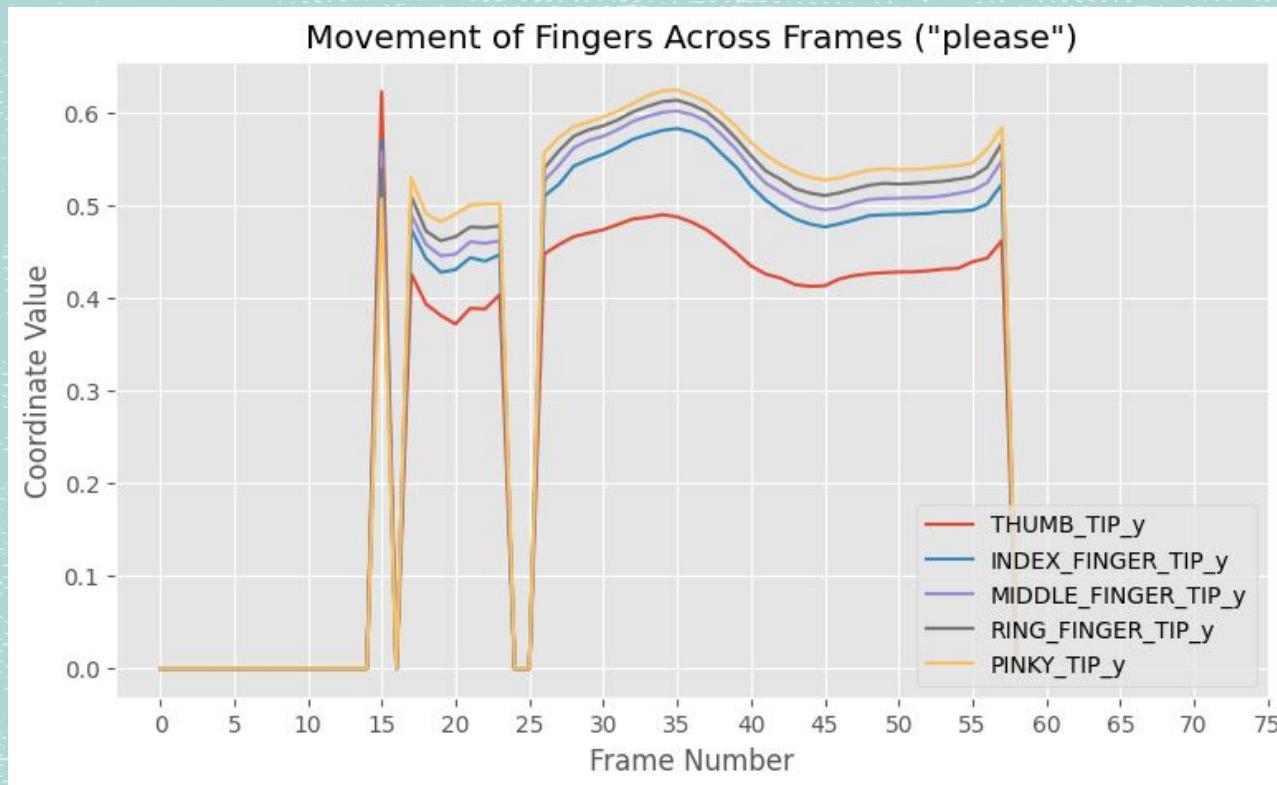
Pretty 'Please'



Pretty 'Please'



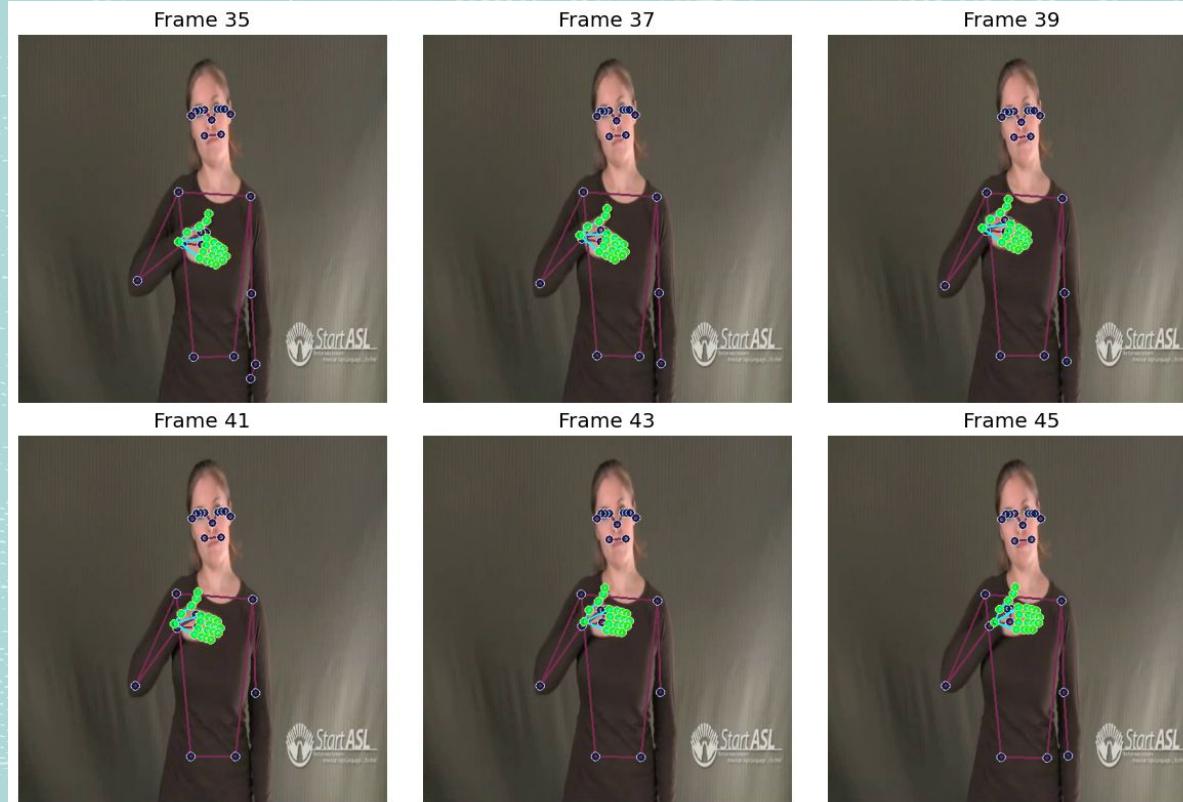
Pretty 'Please'



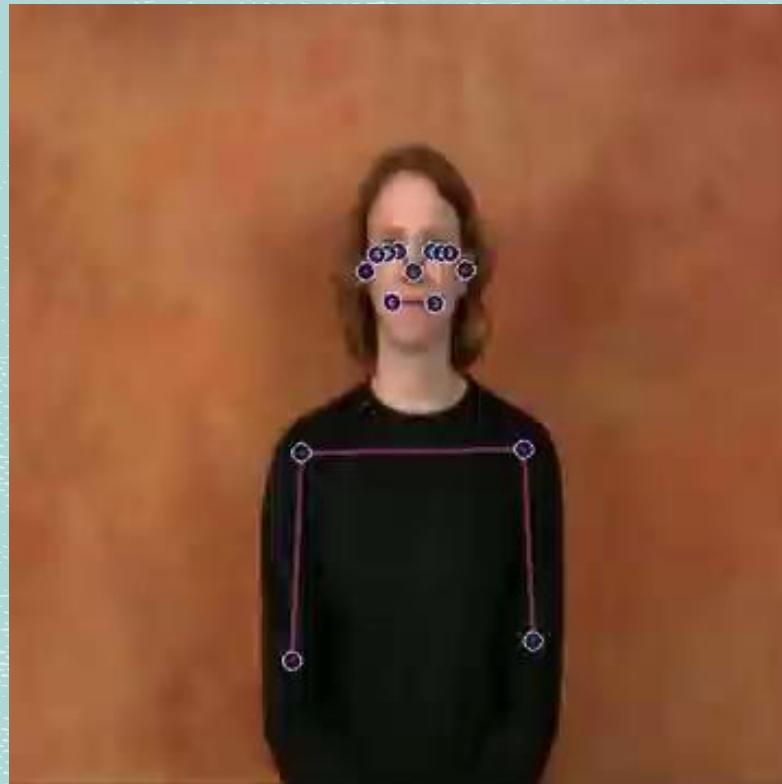
Pretty 'Please'



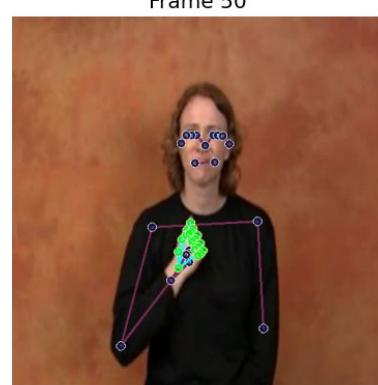
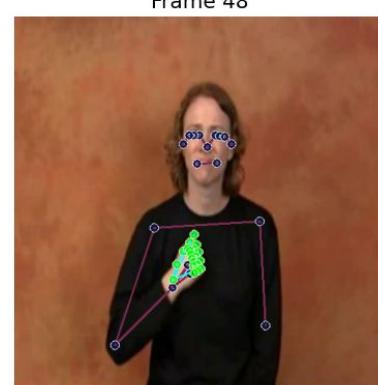
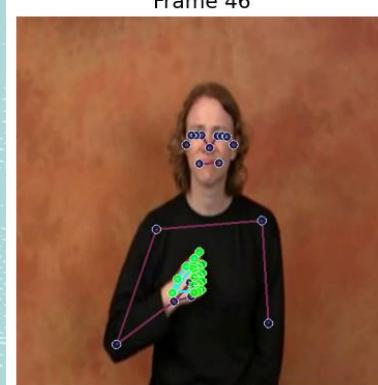
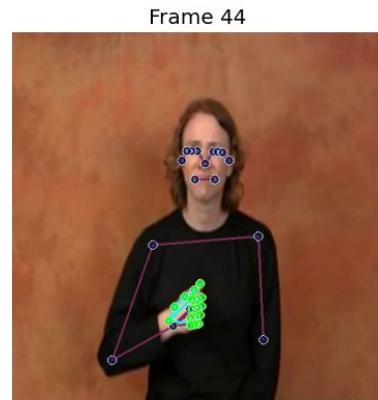
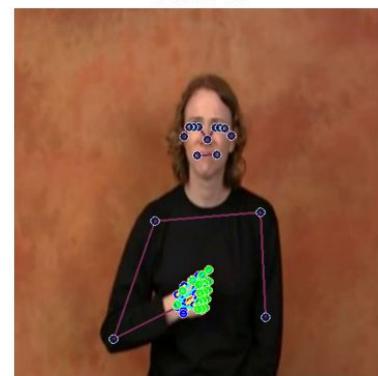
Pretty 'Please'



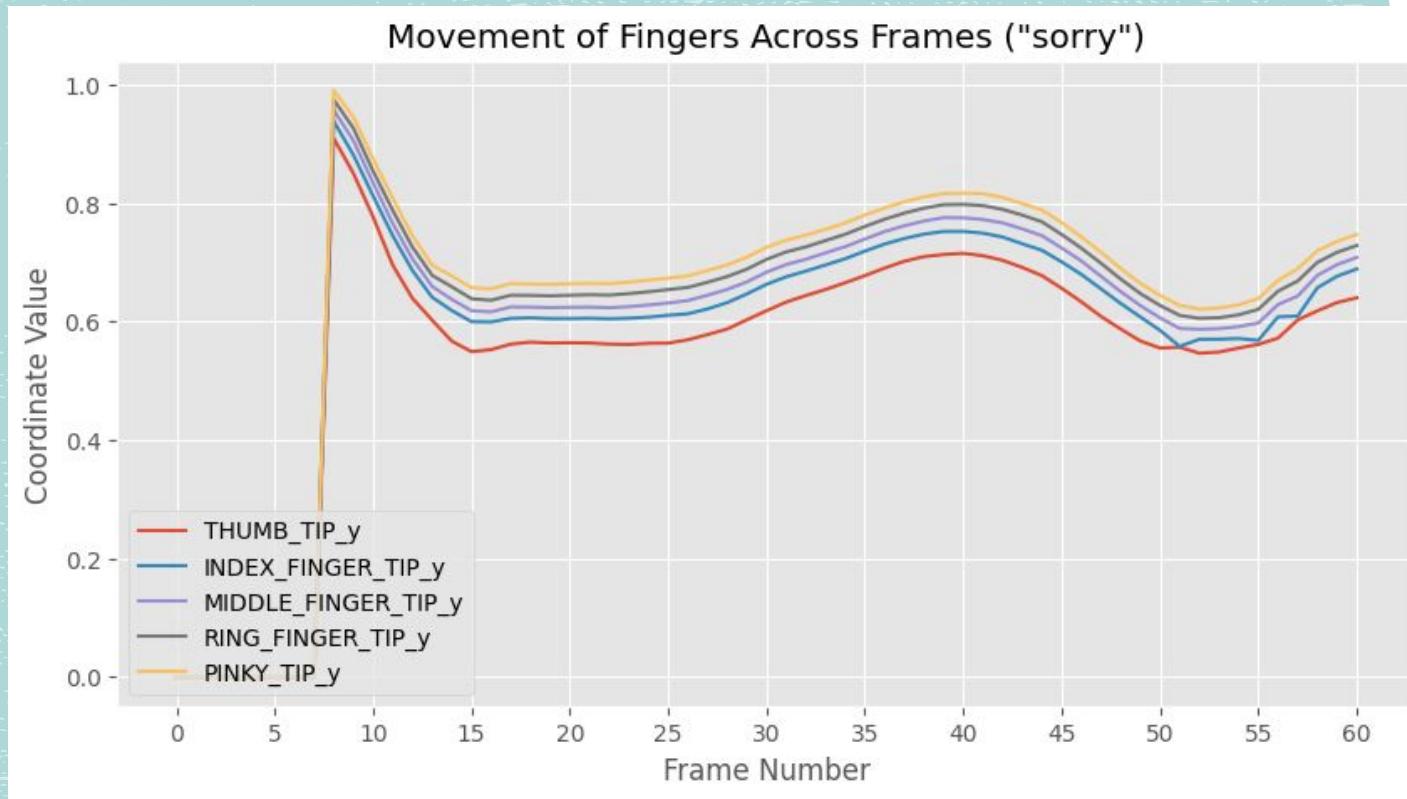
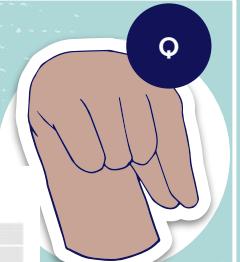
'Sorry'



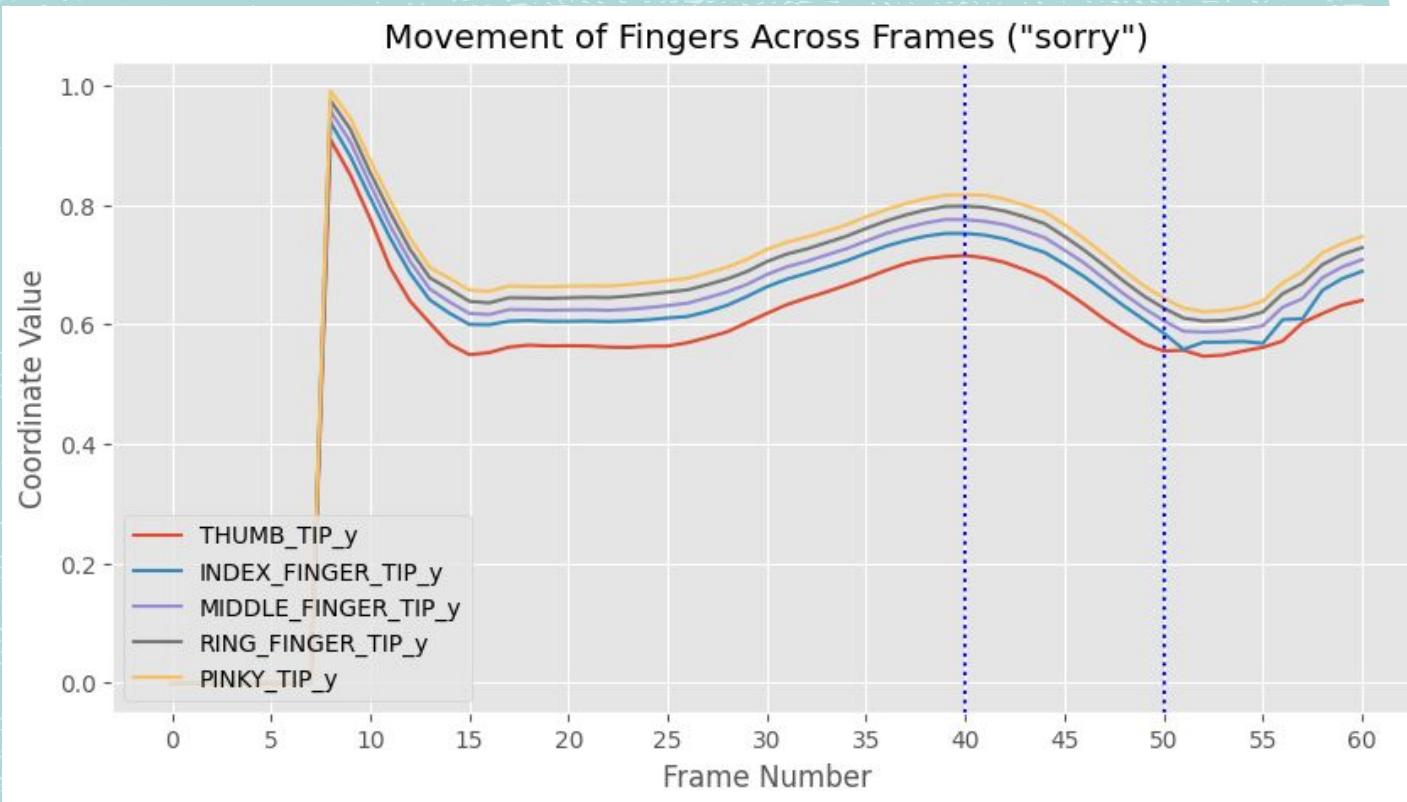
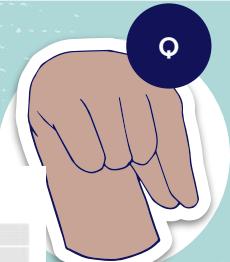
'Sorry'



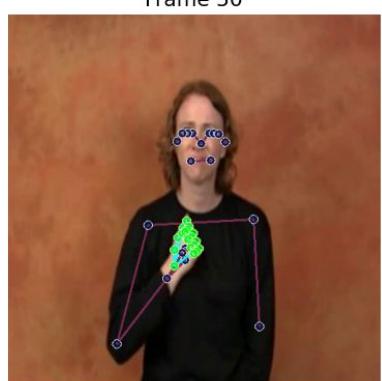
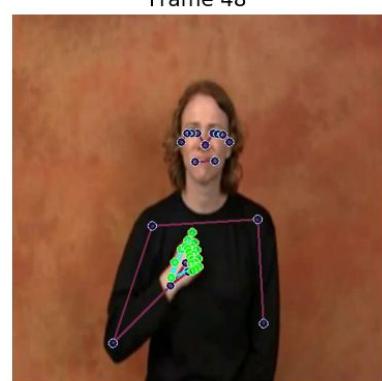
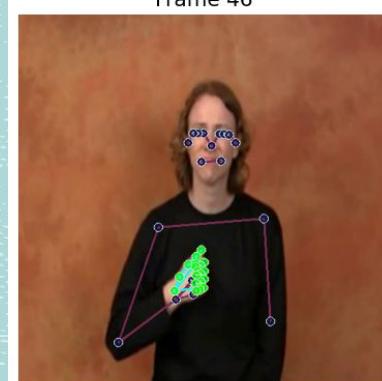
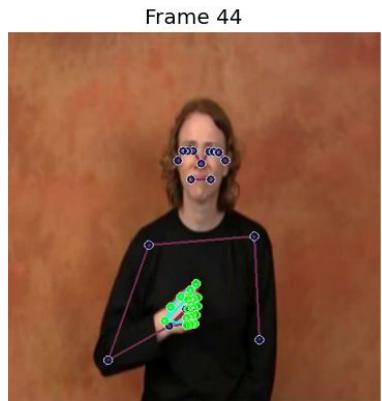
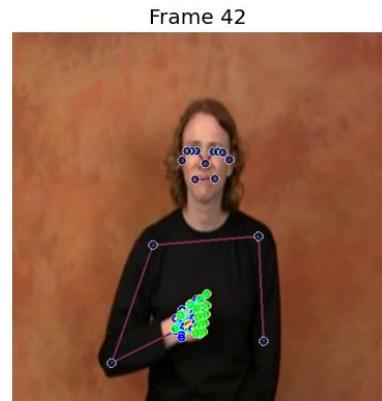
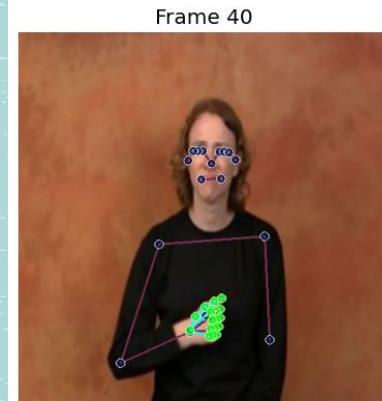
'Sorry'



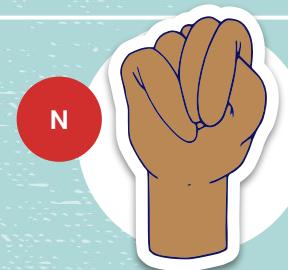
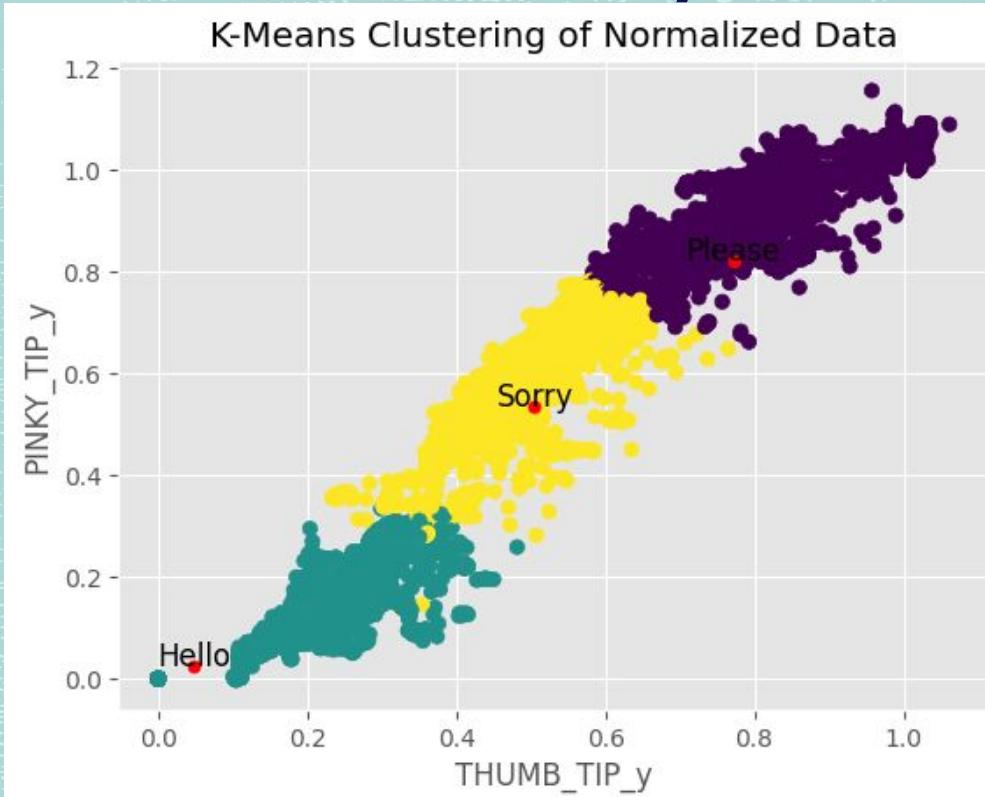
'Sorry'



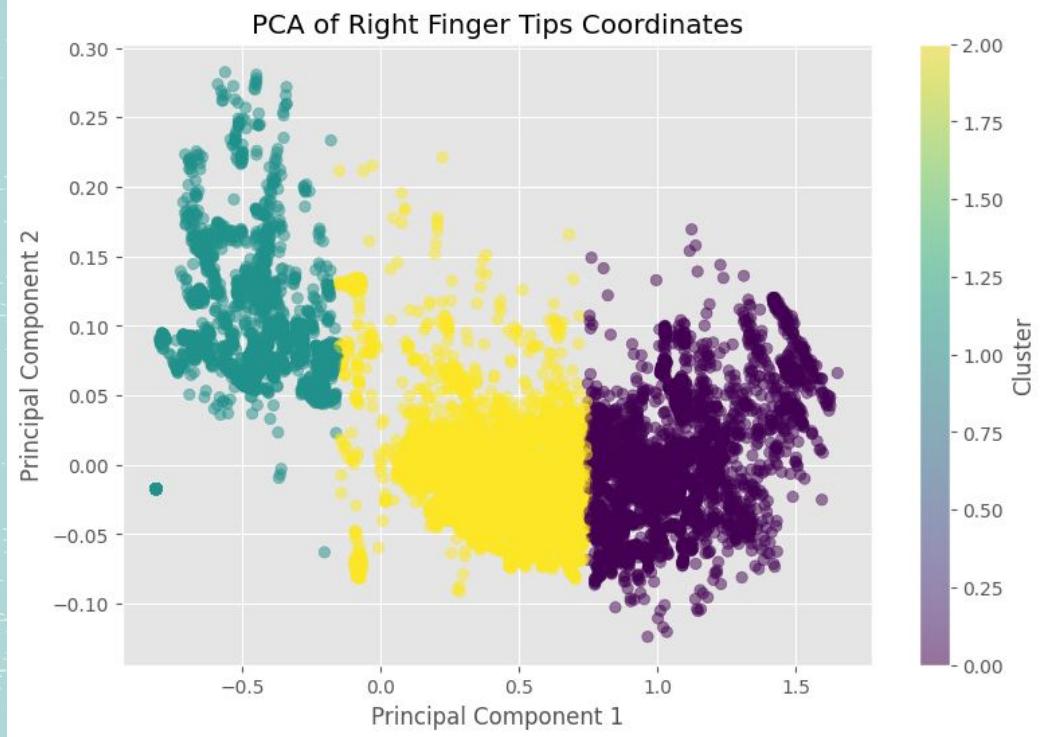
'Sorry'



Cluster Analysis



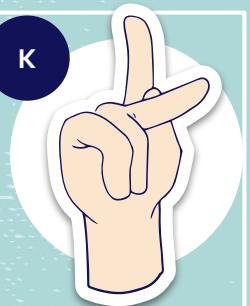
PCA

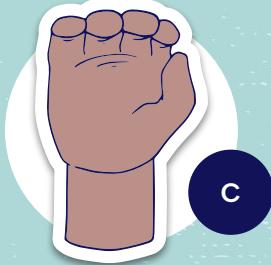


Explained variance:

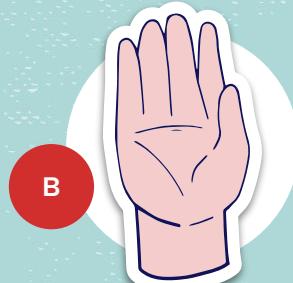
- PC1: 0.995
- PC2: 0.00388

K

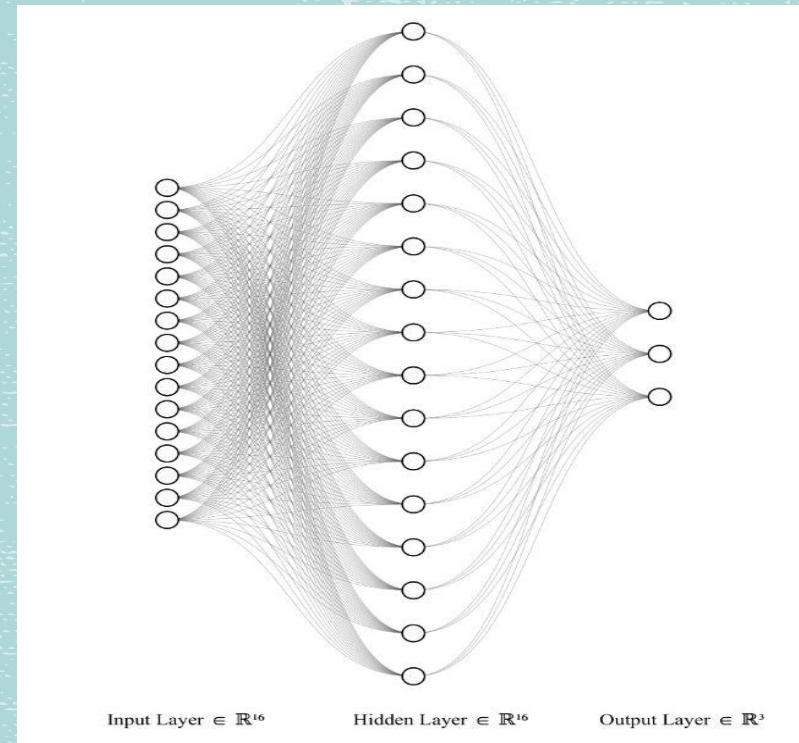




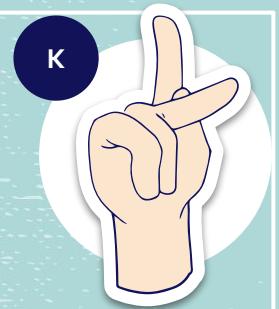
03. Modeling

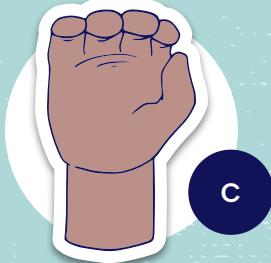


LSTM Model Architecture

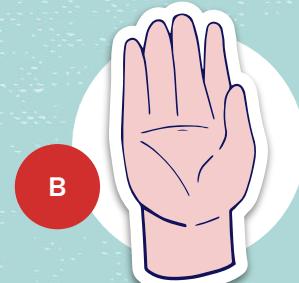


- Input layers: 64
- Hidden layers: 64
- Output layers: 3





04. Results



Results

	Accuracy	Precision	Recall	F1-Score
Comprehensive	0.69	0.73	0.69	0.68

Results

	Accuracy	Precision	Recall	F1-Score
Comprehensive	0.69	0.73	0.69	0.68
Pose + Hands	0.2	0.37	0.2	0.14



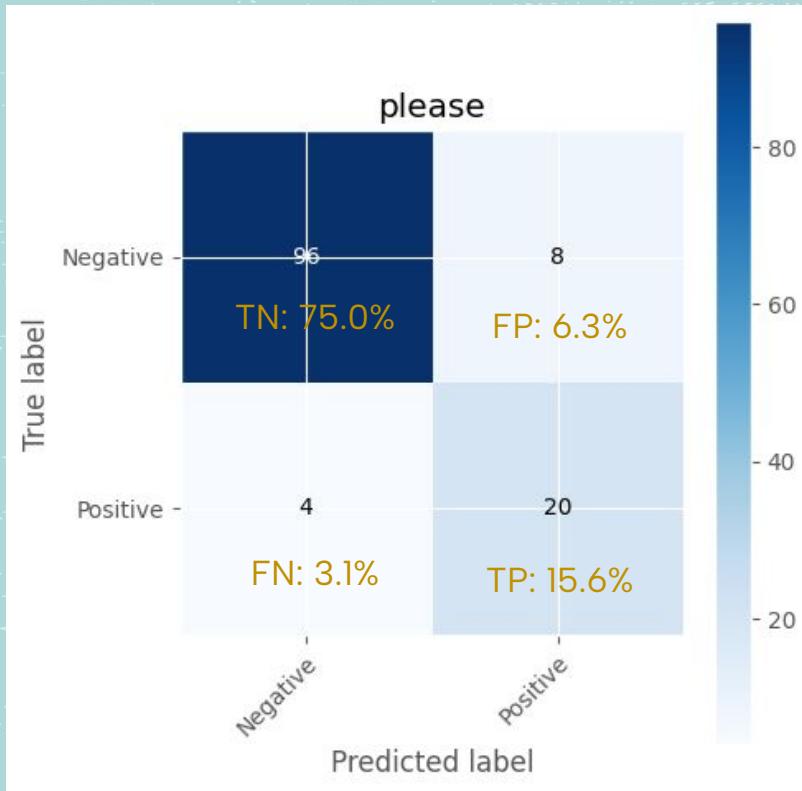
Results

	Accuracy	Precision	Recall	F1-Score
Comprehensive	0.69	0.73	0.69	0.68
Pose + Hands	0.2	0.37	0.2	0.14
Hands only	0.87	0.88	0.87	0.84

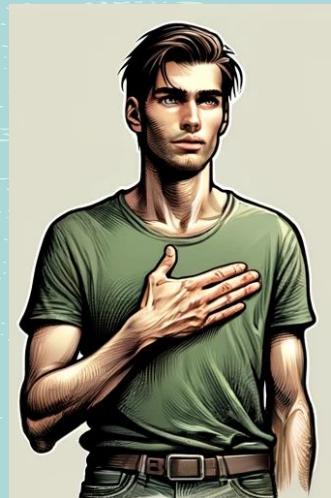


P

Confusion Matrix ('Please')



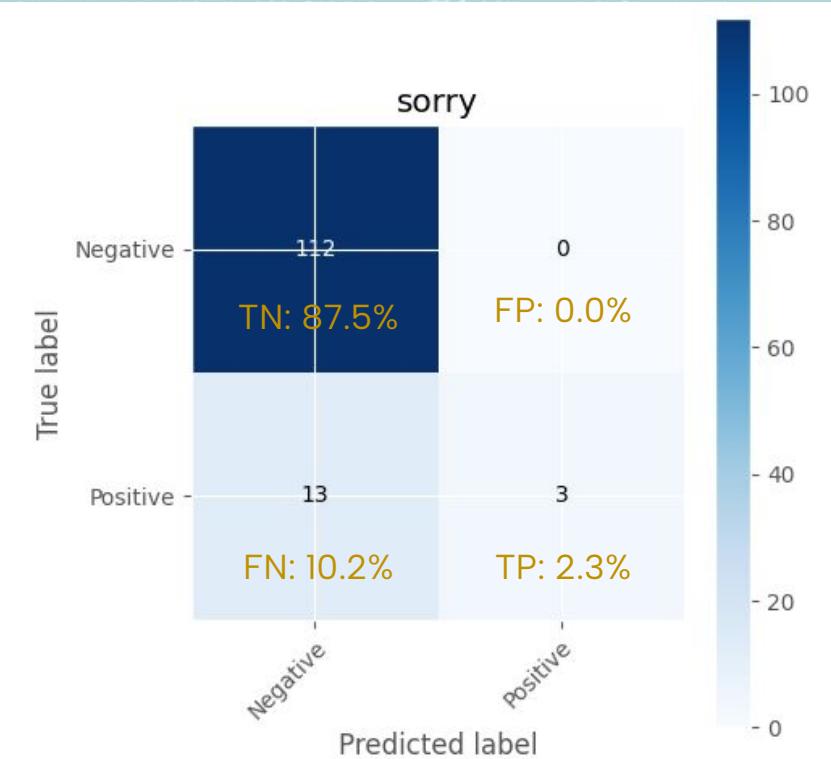
- Total samples: 128
- 28 videos of 'please'
- Accuracy: 90.6%
- Misclassified: 9.4%





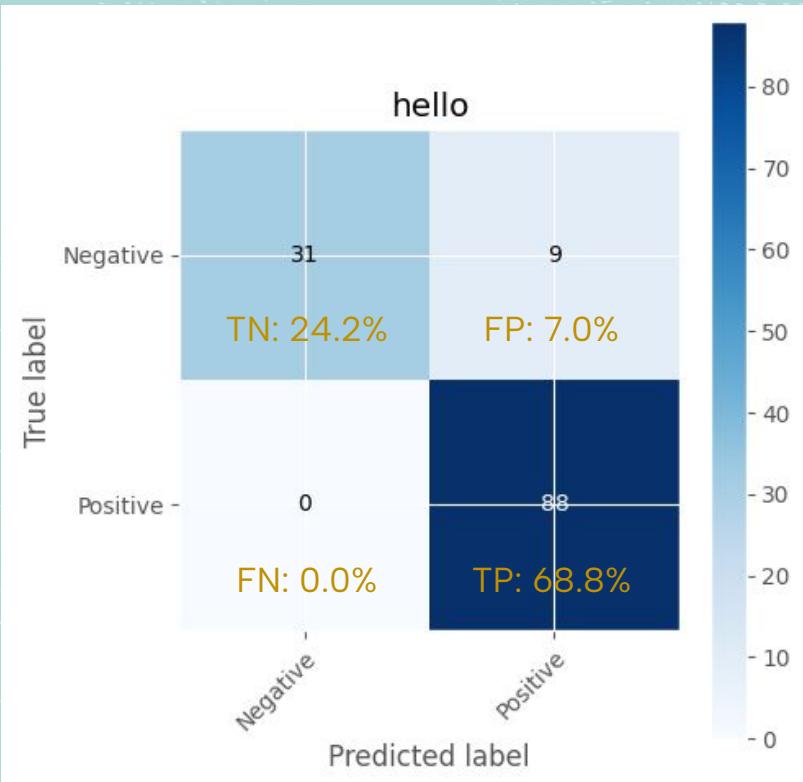
Confusion Matrix ('Sorry')

- Total samples: 128
- 16 videos of 'sorry'
- Accuracy: 89.8%
- Misclassified: 10.2%





Confusion Matrix ('Hello')

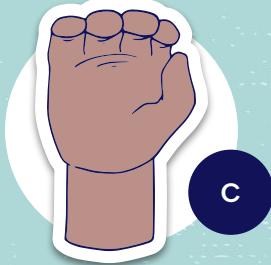


- Total samples: 128
- 88 videos of 'hello'
- Accuracy: 93.0%
- Misclassified: 10.2%



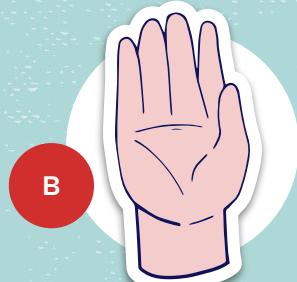
Live Demo



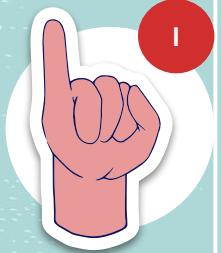


05.

Limitations



Challenges



Insufficient samples

Hard to capture nuances of different people signing

Time constraint in cleaning each of the video files

Current model is a 2 second delay between each detection



Future Work

Improve Word Classes

- Expand the number of samples of each word class
- Include more word class – ‘thank you’, ‘yes’, ‘no’

Improve Latency & Deployment

- Improve on the latency of the real-time prediction
- Integrate into a mobile device



Future Work

Unifying communication in Singapore:

localize the model for sign language



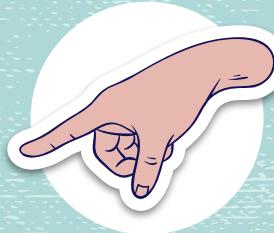
o

Conclusion

- MediaPipe's movement capture holds significant potential for the model
- Initial testing demonstrates encouraging accuracy for the model
- The model is adaptable for local needs
- We are confident in addressing any remaining challenges

thank you

@nyledimarco



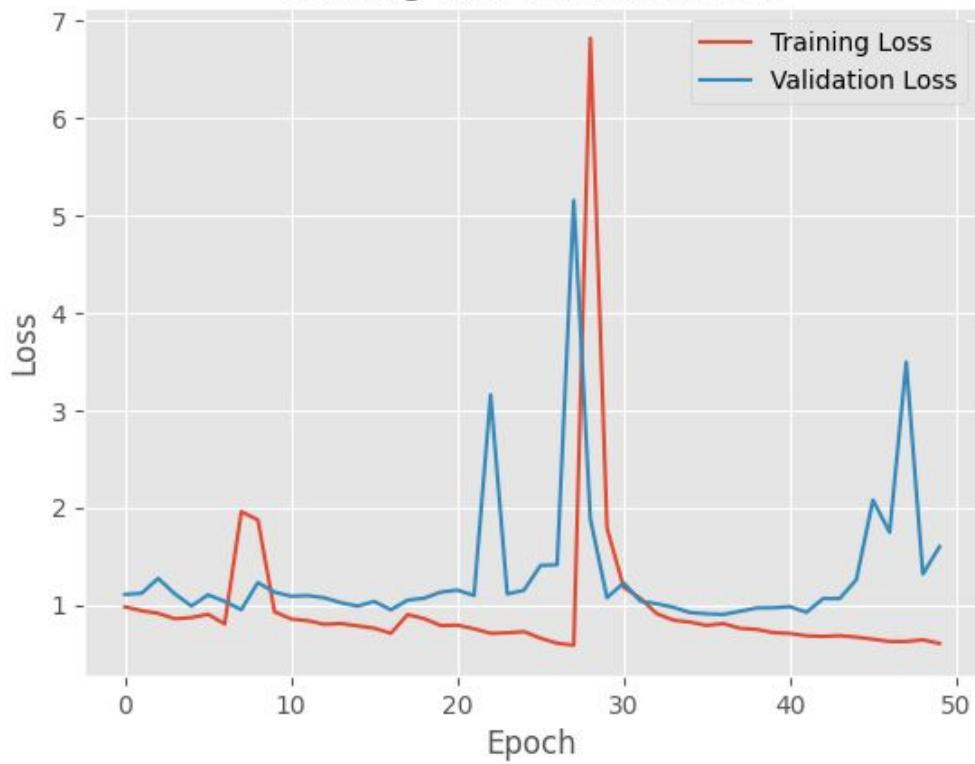


Appendix



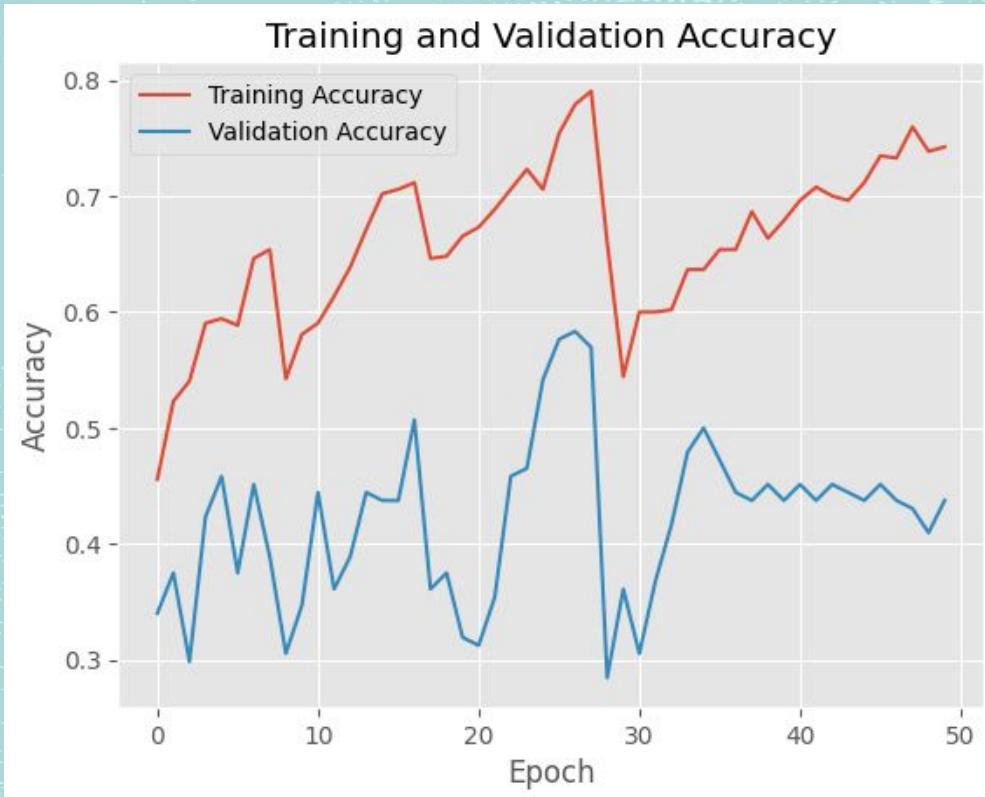
Model Evaluation

Training and Validation Loss



Test Loss: 0.468

Model Evaluation



Test Accuracy: 0.867



Results

	Accuracy	Precision	Recall	F1-Score
Comprehensive	0.69	0.73	0.69	0.68
Pose + Hands	0.2	0.37	0.2	0.14
Hands only	0.87	0.88	0.87	0.84