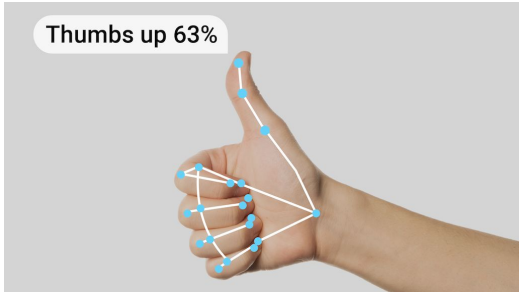
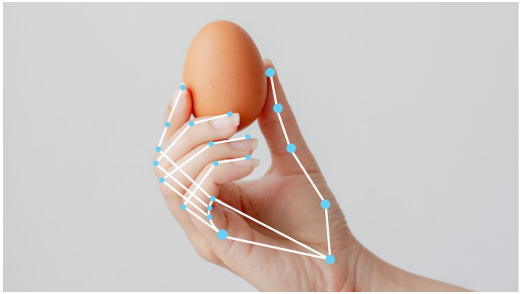
Spring 2025
EE 3530 - Digital Signal Processing
Final Project

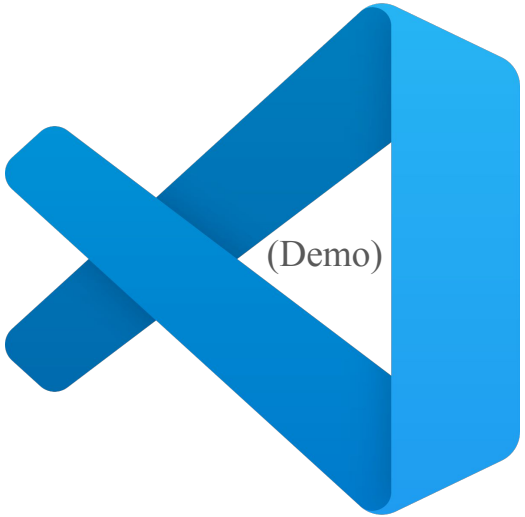
Hand Gesture-Controlled Volume



By: Spencer Hart



Motivation



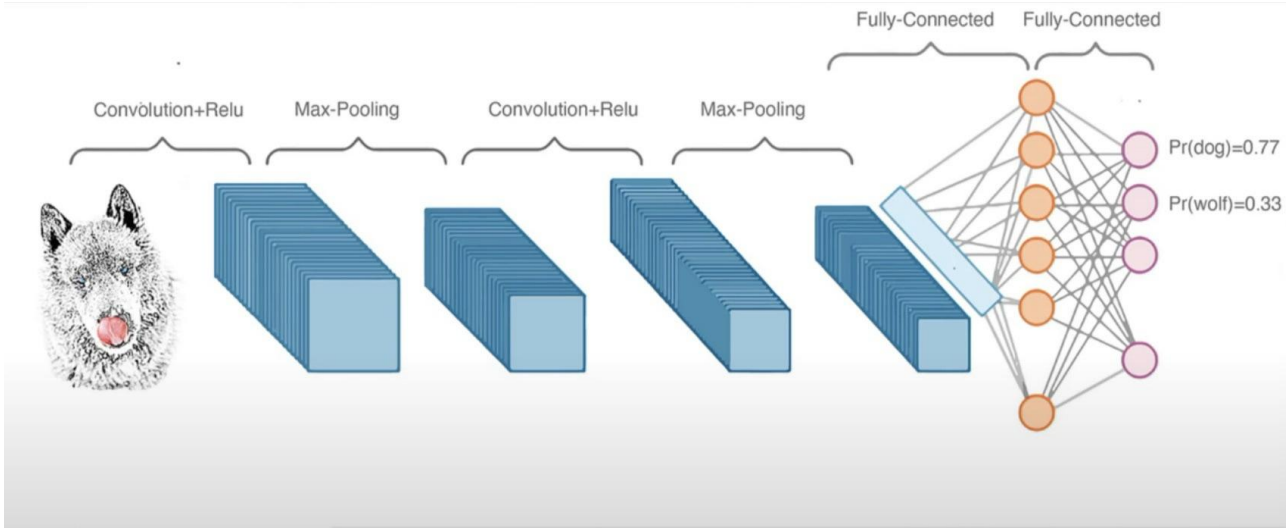
Problem Statement

“How can real-time video footage of hands be captured and processed to allow for hand recognition, tracking, and remote volume control?”

How is DSP involved in this project?

- Images (video) are digital signals
- Manipulation (recognition and tracking) requires processing
- Image processing achieved through machine learning
- Machine learning achieved through CNNs

Convolutional Neural Networks (CNNs)



Why should you care?

- It’s pretty cool
- Could be expanded to more gesture recognition applications (ASL, robotic control, video games)
- You don’t have to, but I’d be pretty sad if you didn’t

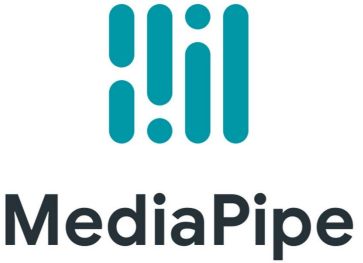
Approach: OpenCV & MediaPipe

OpenCV: **Open**-Source **C**omputer **V**ision

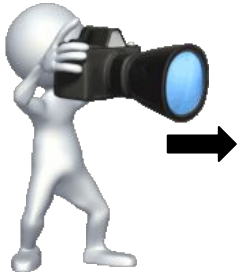
- Extremely convenient library for real-time computer vision

Google's MediaPipe

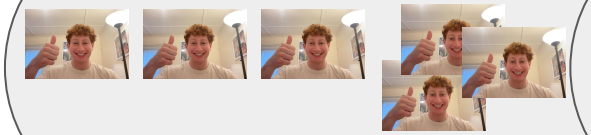
- A framework for creating data pipelines
- Well-packaged
- Not super-well documented
- Still convenient



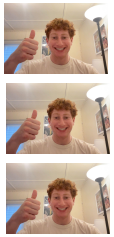
MediaPipe and Approach (continued)



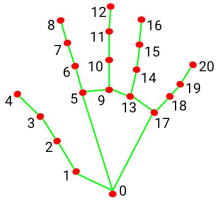
System Pipeline



Processing
(CNN)
→ Slow

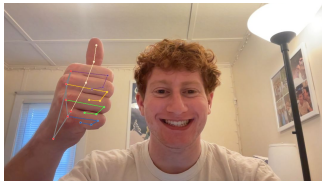


Packet Loss



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

(X, Y, Z)



Challenges

- Understanding OpenCV and MediaPipe syntax
- Functionality from multiple distances

Solution

- Finger-digit distance ratio

Conclusions & Future

- Demonstrated robust hand recognition and tracking system to control volume

- Further research on increased gesture recognition to perform more tasks (e.g. “thumbs-up”, “OK” sign, “STOP”) and increase applicability