RUTGERS
THE STATE UNIVERSITY OF NEW JERSEY

Hand ANalyzed Dynamics; HAND

Team Members: Samuel Marran James Artuso
Advisor: Sasan Haghani

TEAM
S24-42

Abstract

Studies have shown that the addition of a physical component in the process of teaching sign language can greatly enhance the retention rate [2]. This, coupled with an interactive experience, results in a superior learning environment. Our project is meant to fill that void by providing real-time communication with a robotic agent. By coupling a computer vision model with a robotic hand, our prototype has been able to participate in games of rock-paper-scissors to simulate the act of sign language teaching.

Rock-paper-scissors is the perfect game to demonstrate the viability of this method due to its call and response nature [1]. Our project interprets hand poses from a camera inputs. With low latency, it determines the pose quickly and responds with the robot. This ability is shown in its ability to win rock-paper-scissors 100% of the time. In addition, the robot can understand the American Sign Language alphabet. This is to show the software's scalability to a larger classification task.

Problem Formulation & Challenges

The main design challenges to develop a real time robotic hand with rapid response capabilities, are listed as follows:

- Designing realistic hand movement
- Reducing latency in both hardware and software
- Tracking the hand in real time
- Picking a neural network architecture

Robotic Solution

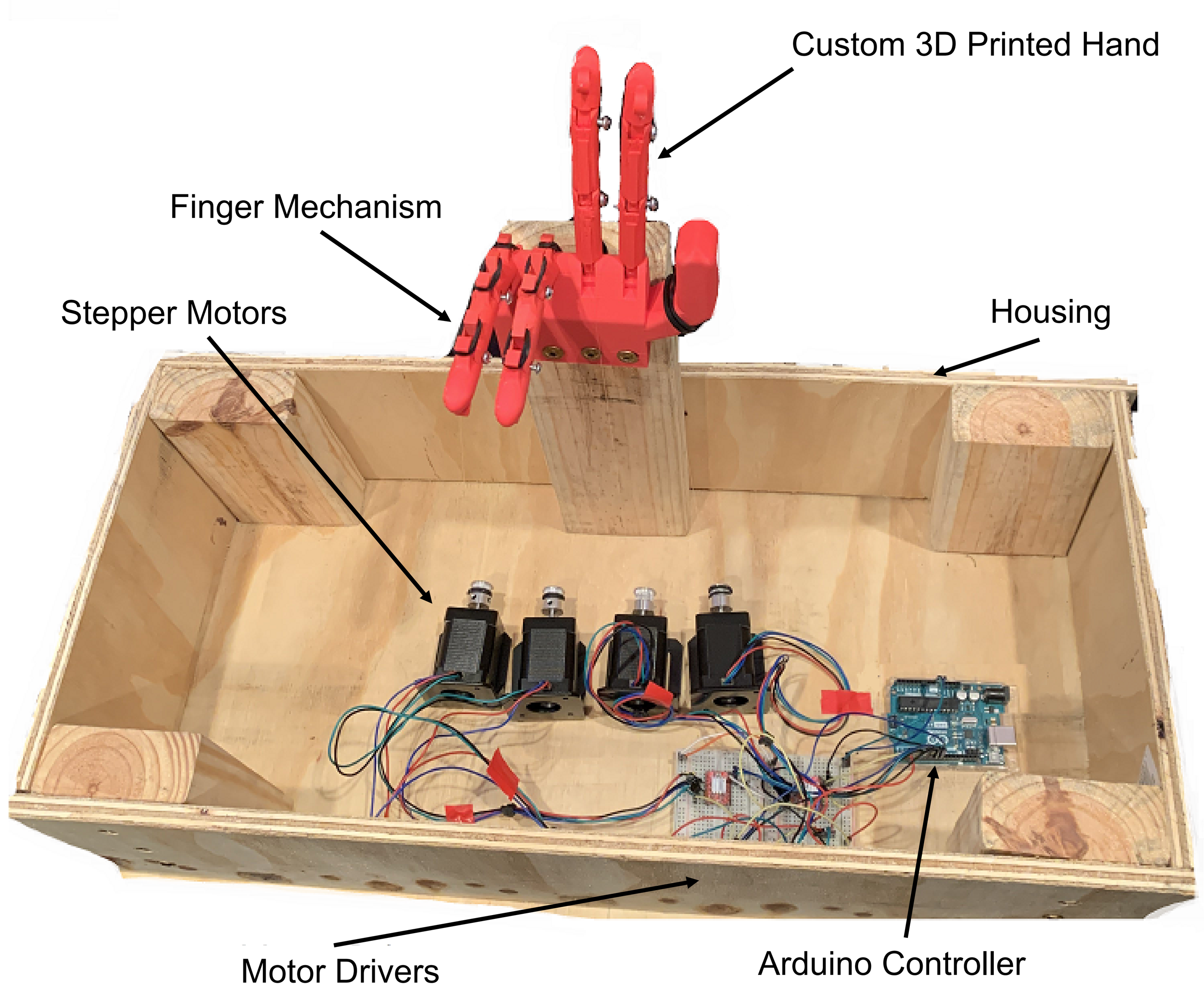

Fig. 1 Full HAND and Electrical Equipment Diagram

Main Results

Pose Recognition Software:

- Software is split to process the image in a pipeline
- Camera: Any generic off the shelf camera with suitable resolution
- Hand Landmark Identifier: Googles MediaPipe library provides a lightweight solution to creating 21 hand landmarks
  - This library is intended for mobile platforms, so it requires little computational cost
- Pose Identifier: LSTM Model
  - 1 Input layer
  - 2 stacked LSTM's
  - 2-layer decoding dense network
- Action Selection: Markov Chain based on pose estimation

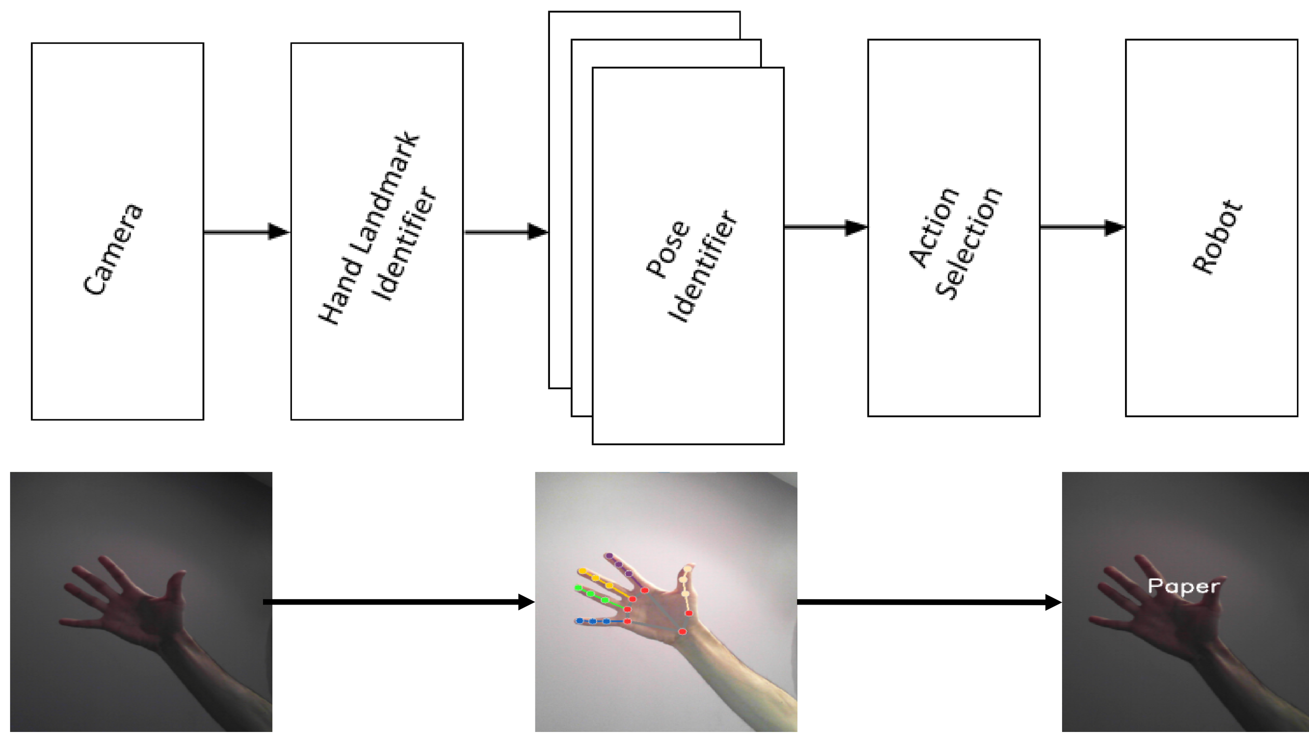

Fig. 3 Machine learning Flow Chart

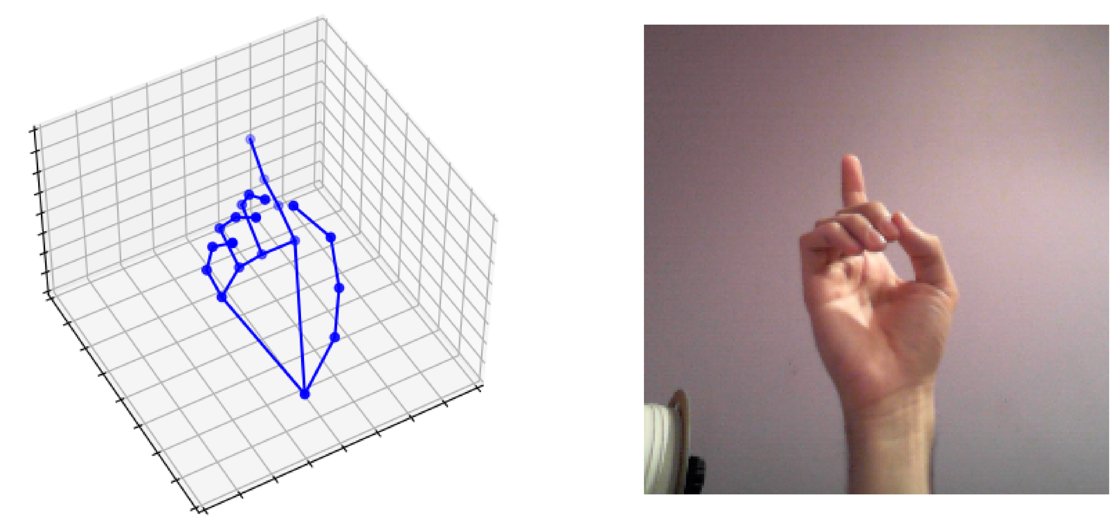

Fig. 2 Hand Landmark Indicator Diagram

Sign Language Dataset:

- 26 classes
- 40 samples per class
- 60 frames per sample
- Use rotations and noise to multiple dataset size by up to 6 times
- Similar dataset used for Rock-Paper-Scissors

Robot Hand:

- 4 independently actuated fingers
  - 100 degrees of movement
- Precises control via stepper motors
- High durability 30-35% 3D print infill
- Stable holding state with a 2 lb. pulling force

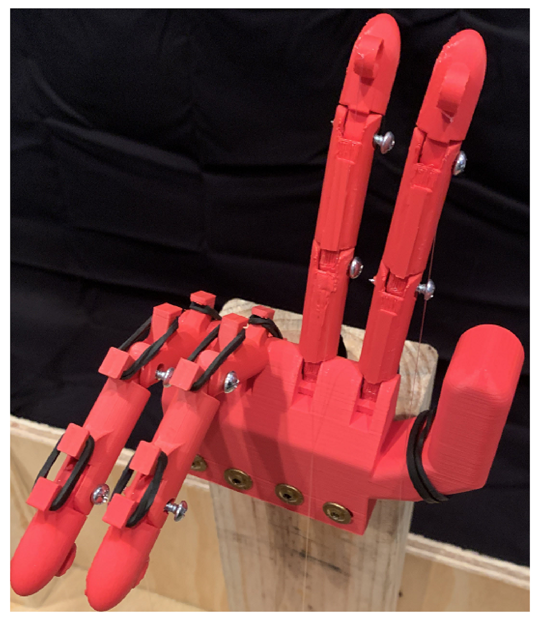

Fig. 4 HAND in Scissor State

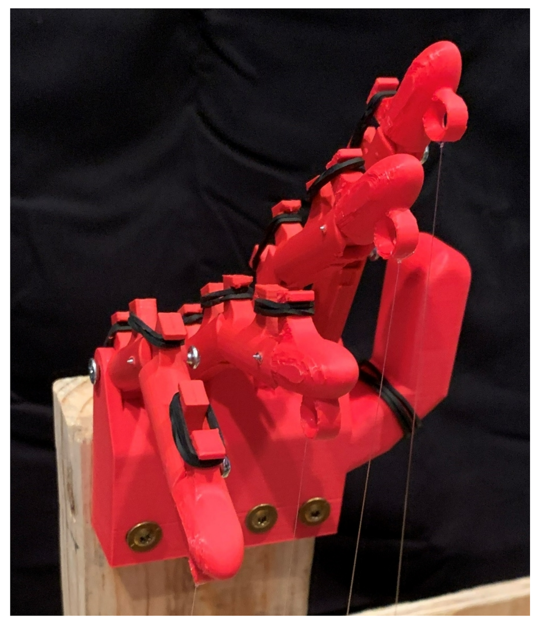

Fig. 5 HAND in Curled State

Future Works

- Fully actuated hand and wrist to enable more intricate interaction
- Full American Sign Language support
- Due to low computational cost mobile app support could increase the number of potential users
- Battery power and portable teaching assistant

References

[1] Ahn, H.S., Sa, IK., Lee, DW. et al. "*A playmate robot system for playing the rock-paper-scissors game with humans.*" Artif Life Robotics 16, 142 (2011)

[2] H. Kose and R. Yorganci, "*Tale of a robot: Humanoid robot assisted sign language tutoring,*" 2011 11th IEEE-RAS International Conference on Humanoid Robots, Bled, Slovenia, 2011, pp. 105-111, doi: 10.1109/Humanoids.2011.6100846

[3] N. Aphiratsakun, X. J. Blake, K. K. Tin and T. Ngwe, "*AI-based Rock-Paper-Scissors plug and play system,*" 2020 5th International STEM Education Conference (iSTEM-Ed), Hua Hin, Thailand, 2020, pp. 30-33, doi: 10.1109/iSTEM-Ed50324.2020.9332629.

[4] A. Sherstinsky, "Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) Network," arXiv.org, 2018.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," arXiv.org, Dec. 10, 2015.