Leveraging a Large Language Model for Movement Intention Recognition

Jared Comeau, Department of Electrical and Computer Engineering, University of New Hampshire
Advisor: Dr. Diliang Chen

INTRODUCTION

- Problem: Current systems rely on sensors such as electromyography (EMG) and inertial measurement units (IMU), to detect muscle or nerve activity.
- Background: Movement intention recognition is critical for devices like prosthetics. Traditional systems that use sensors can take about 109.67 ms [1]. Advancements in LLMs have the potential to analyze visual data and improve both the speed and accuracy of movement prediction systems.

Figure 1: Camera Glasses

RESEARCH QUESTION

- Can LLMs be effectively used to detect a user's movement intention using video from a pair of glasses?
- Discovering if a LLM can be effectively used to detect movement intention could improve the speed and accuracy of movement intention recognition to create a better or smoother experience for the user.

METHODOLOGY

- Take video using glasses of different movements.
  - Extract Frames From Video
  - Reduce Resolution of Frames
  - Stitch Frames into one image.
- Feed Stitched Image into LLM using its API.
  - Gemini 2.0 Flash – Fast and versatile
- Keep track of the time it takes to turn video into an image, send it to the LLM, receive an answer, and get a result.

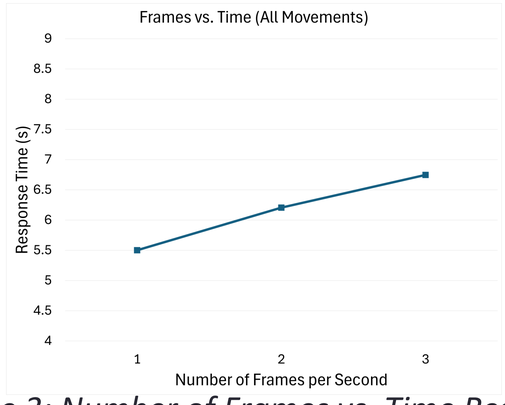

Figure 3: Number of Frames vs. Time Response

CHALLENGES

- Image Size
- LLM output type
- Figuring out best prompt
- Issues connecting with Gemini API
- Input video Difficulty

Figure 2: Stitched Input Image

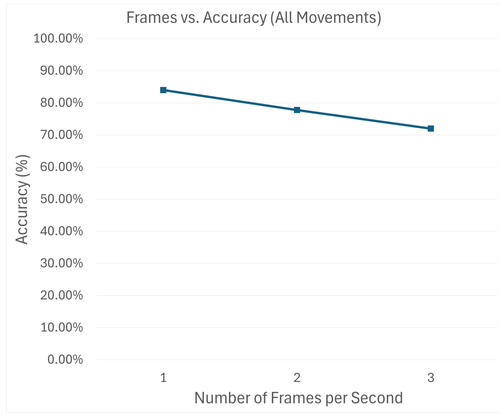

Figure 4: Number of Frames vs. Accuracy

CURRENT RESULTS

- Tested 5 Movements
- Tested 3 different FPS
  - Used same videos for all
- 10 Videos
  - 5-11 seconds
  - 3 videos for stairs, 1 for ramps, 2 for flat ground
- 1 FPS
  - Response times varied between 4 and 8 seconds
  - 42/50 Correct
- 2 FPS
  - Response time varied between 4 and 9 seconds
  - 35/50 correct
- 3 FPS
  - Response time varied between 4 and 10 seconds
  - 36/50 correct

FUTURE WORK

- Including more movements
- Larger Dataset
- More FPS
- Larger Number of Trials
- Improve Prompt

REFERENCES

[1] P. Zhang, J. Zhang and A. Elsabbagh, "Lower Limb Motion Intention Recognition Based on sEMG Fusion Features," in IEEE Sensors Journal, vol. 22, no. 7, April, pp. 7005-7014, 2022