A. Network Architectures

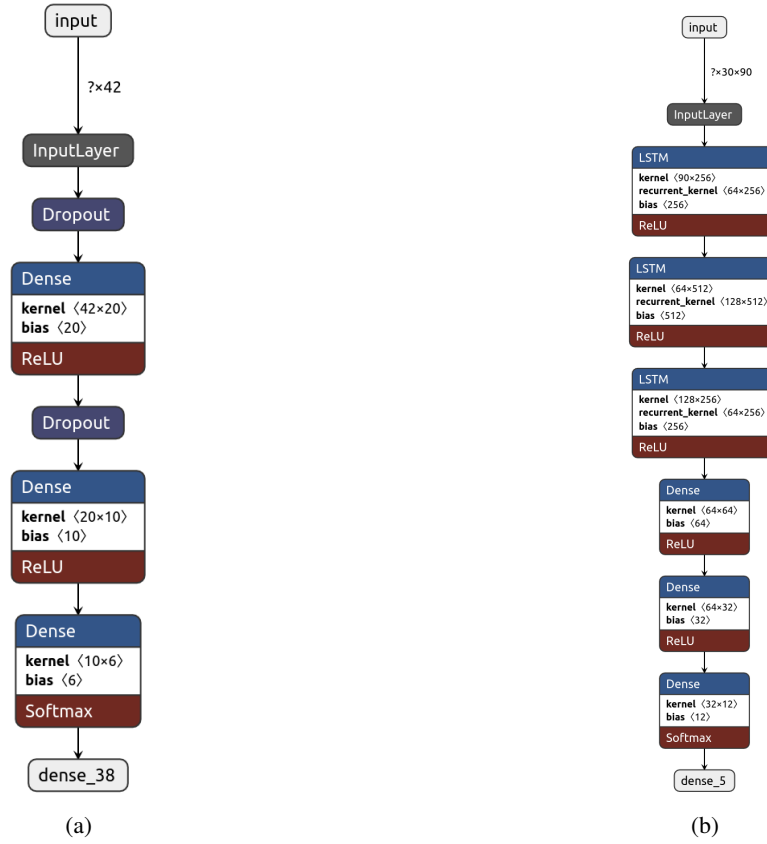Fig. 6: Classifier architecture used for static hand gestures (left) and sign language classification (right).

B. Robot modes and feedback options

Mode	Description	ASL	Speech
Imitate	Right fist movement is imitated (default mode)	-	-
Record	Record trajectory	record	'record'
Feedback	Trajectory modification occurs in this mode	feedback	'feedback'
Replay	Trajectory will be replayed in this mode	play	'replay'
Move	Move interface will be shown on the camera feed	move (from one location to another)	'move

TABLE I: List of modes.

Modification	Description
Up	Trajectory is translated up
Down	Trajectory is translated down
Left	Trajectory is translated left
Right	Trajectory is translated right
Pitch	Y-axis rotation performed on trajectory
Roll	X-axis rotation performed on trajectory
Yaw	Z-axis rotation performed on trajectory
Faster	Movement is sped up
Slower	Movement is slowed down
Bigger	Trajectory is scaled up (larger)
Smaller	Trajectory is scaled down (smaller)

TABLE II: List of trajectory modifications available in feedback mode.

C. GUI
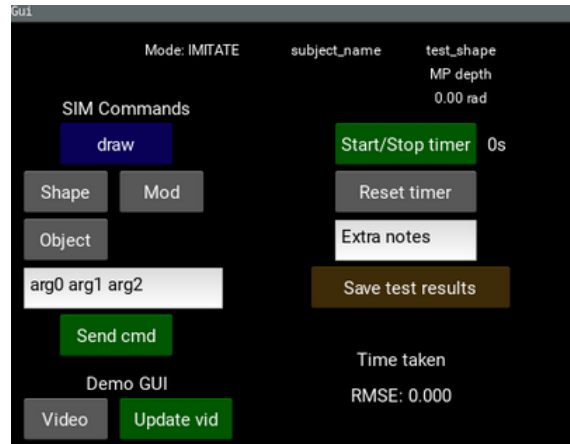
Here are some complementary screenshots of the GUI.

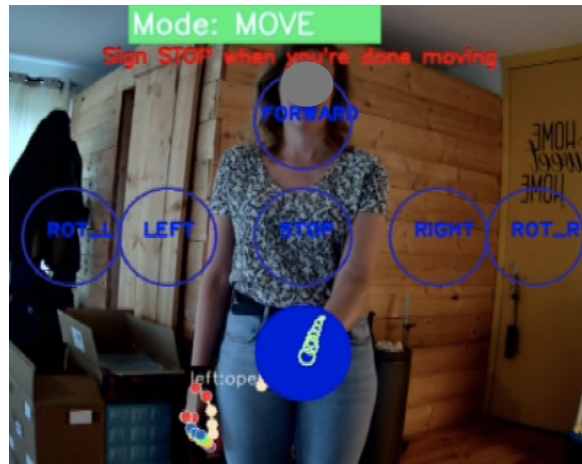Fig. 7: The graphical User Interface created for the user study.

Fig. 8: Move interface overlay on the webcam feed.

D. Setup

1) Hardware:
- Motherboard: ASRock B450M Pro4
- CPU: AMD Ryzen 3700x 8-Core Processor
- GPU: AMD RX-580 8GB
- Memory: 16 GB

2) Experiment Setup:
- Two monitors: The first monitor displayed the robot simulation (PyBullet), the camera feedback and the instructions (controlled by GUI) for the sign language commands. The second monitor was used to control the GUI.
- Webcam (Maccy 4MP 30fps) pointed at the participants.