GENERATION OF MOVEMENT EXPLANATIONS FOR TESTING GESTURE BASED CO-OPERATIVE LEARNING APPLICATIONS

Ayan Banerjee, Imane Lamrani, Prajwal Paudyal, and Sandeep K.S. Gupta.

iMPACT Lab

CISDE, Arizona State University.

sandeep.gupta@asu.edu

National Institutes of Health

INTRODUCTION

Trust is paramount in any AI enabled System

Lack of Trust

2 crashes in 5 months

Crash in Autonomous Cars

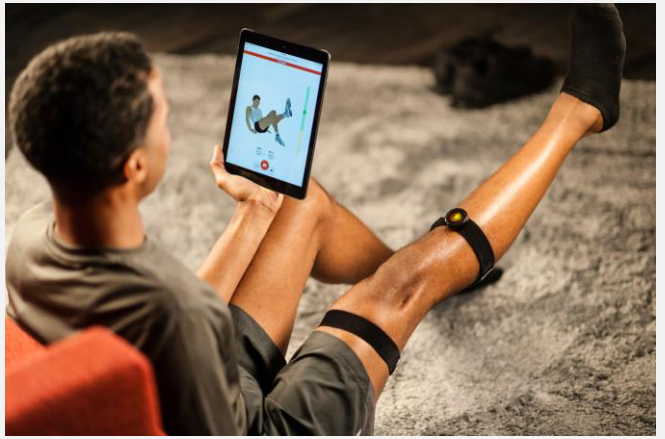
Hyperglycemia

Explanations are necessary for generating Trust in AI Systems

Why did the AI system fail?
How can the AI system be improved?

COOPERATIVE LEARNING

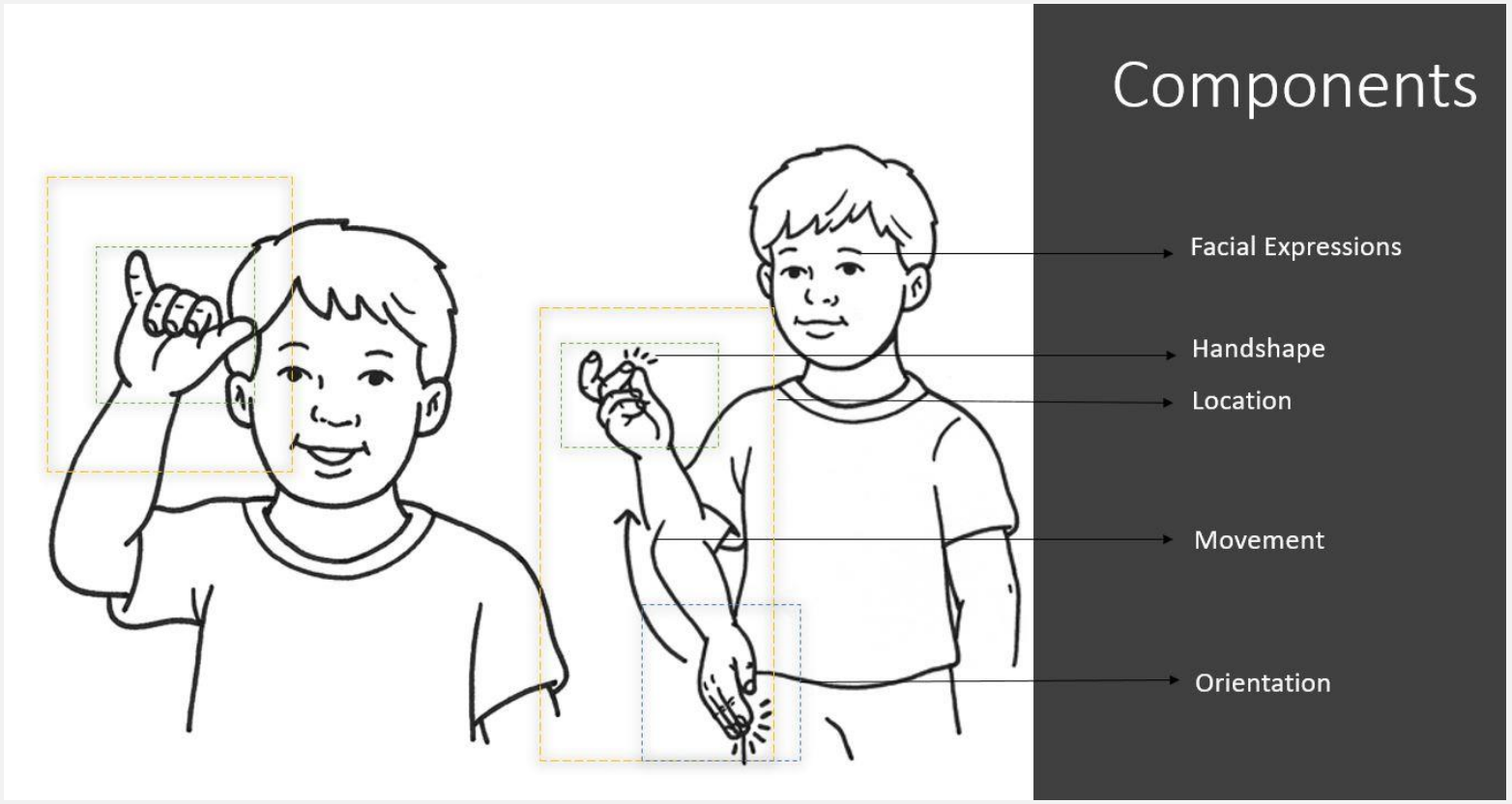
- **ML-based gesture CLA** include computer aided training of military personnel, rehabilitation for diseases such as or Parkinson, performance coaching in entertainment applications …
- **ML-based gesture CLA** recognize the correctness of the gesture performed by the learner.
- . **ML-based gesture CLA** should also provide feedback to the learner.



Trust ➔ Explanation

Why was a gesture recognized positively or negatively?

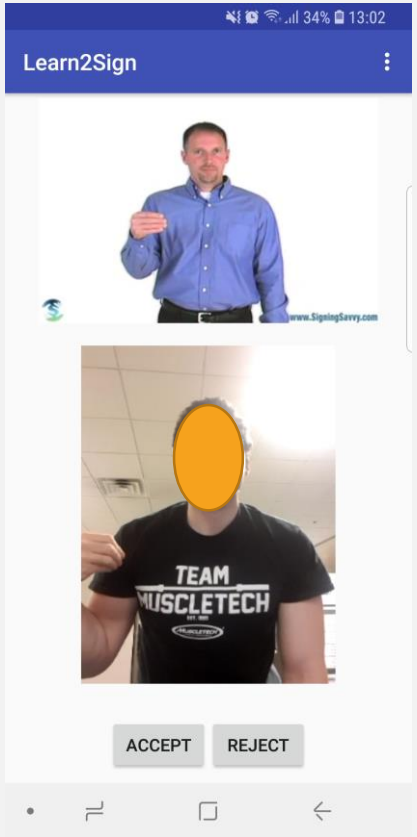
EXPLANATIONS FOR GESTURES: AMERICAN SIGN LANGUAGE EXAMPLE



Expert

Learner

Learn2Sign

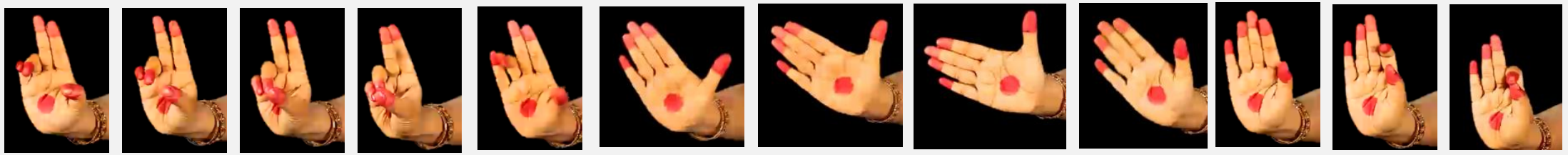


COMPONENTS OF EXPLANATION

- Handshape Explanation
  - Explanation generation: CNN, RNN, DNNs [Andreas, et al., B. Liu, et al., R, Hu, et al., A. H. Abdulnabi, et al.].
- Location Explanation
  - Visual explanations are intuitive.
- Movement Explanation
  - Very limited work.

Paper Focus: Movement Explanation and corrective feedback

CHALLENGE FOR EXPLAINING MOVEMENT



SHAPE 1 $\frac{dp}{dt} = v, \frac{dv}{dt} = a, \frac{da}{dt} = f_1(a, v, x)$ Shape Transition Condition **SHAPE 2** $\frac{dp}{dt} = v, \frac{dv}{dt} = a, \frac{da}{dt} = f_2(a, v, x)$ Shape Transition Condition **SHAPE 3** $\frac{dp}{dt} = v, \frac{dv}{dt} = a, \frac{da}{dt} = f_3(a, v, x)$ Shape Transition Condition **SHAPE 4**

Unique shapes interspersed with differential dynamics

How do we explain differential dynamics?

a) What is the reason behind a given recognition result?

b) When does a learner succeed or fail a gesture execution?

c) When can a recognition result be trusted?

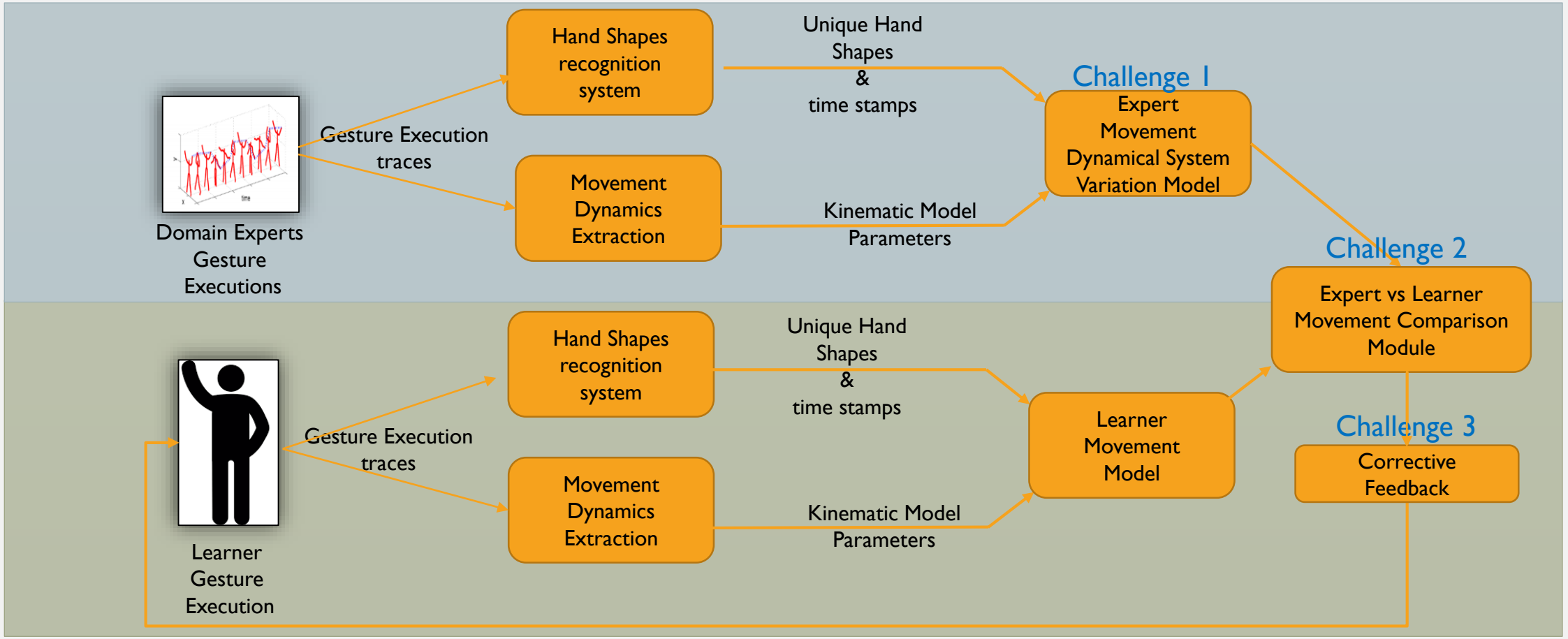
d) How does a learner correct the error in case of failure?

e) How do we tackle variations among experts?

CONTRIBUTIONS

>> Model Gestures	Hybrid Automata	Hybrid Model Mining Technique **HyMn**
>> Expert Variations	Reachability Analysis	Expert variation extraction from practical deployments
>> Learner Deviation	Reach set Comparison	Usage of Learn2Sign Mobile Application
>> Corrective Feedback	Parameter Variations in Kinematic Models	Multi-variate Parameter Estimation

SOLUTION ARCHITECTURE



DATA COLLECTION

- We collected Kinect data including video and bone movement data from **60 subjects** for **20 ASL gestures**.
  - About, After, Hello, Tiger, Cat, Cop, Help, Day, Deaf…
  - **30 users** performed every gesture nearly like an expert.
    - ➢Used to build expert hybrid model based movement recognition.
  - **10 users** performed every gesture nearly like an expert.
    - ➢Used to test the accuracy of the inferred expert hybrid model.
  - **20 users** made several errors in executing any gesture.
    - ➢Used to evaluate the movement explanation module.

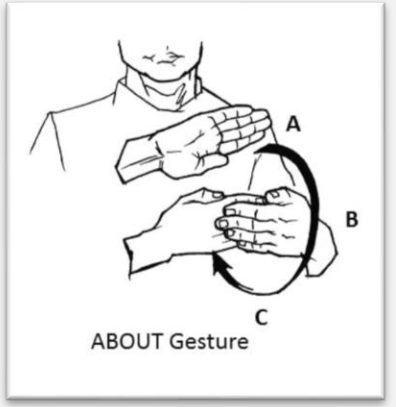
EXPERT VARIATION MODELING

Experts Hybrid Model Mining using HyMn Technique:

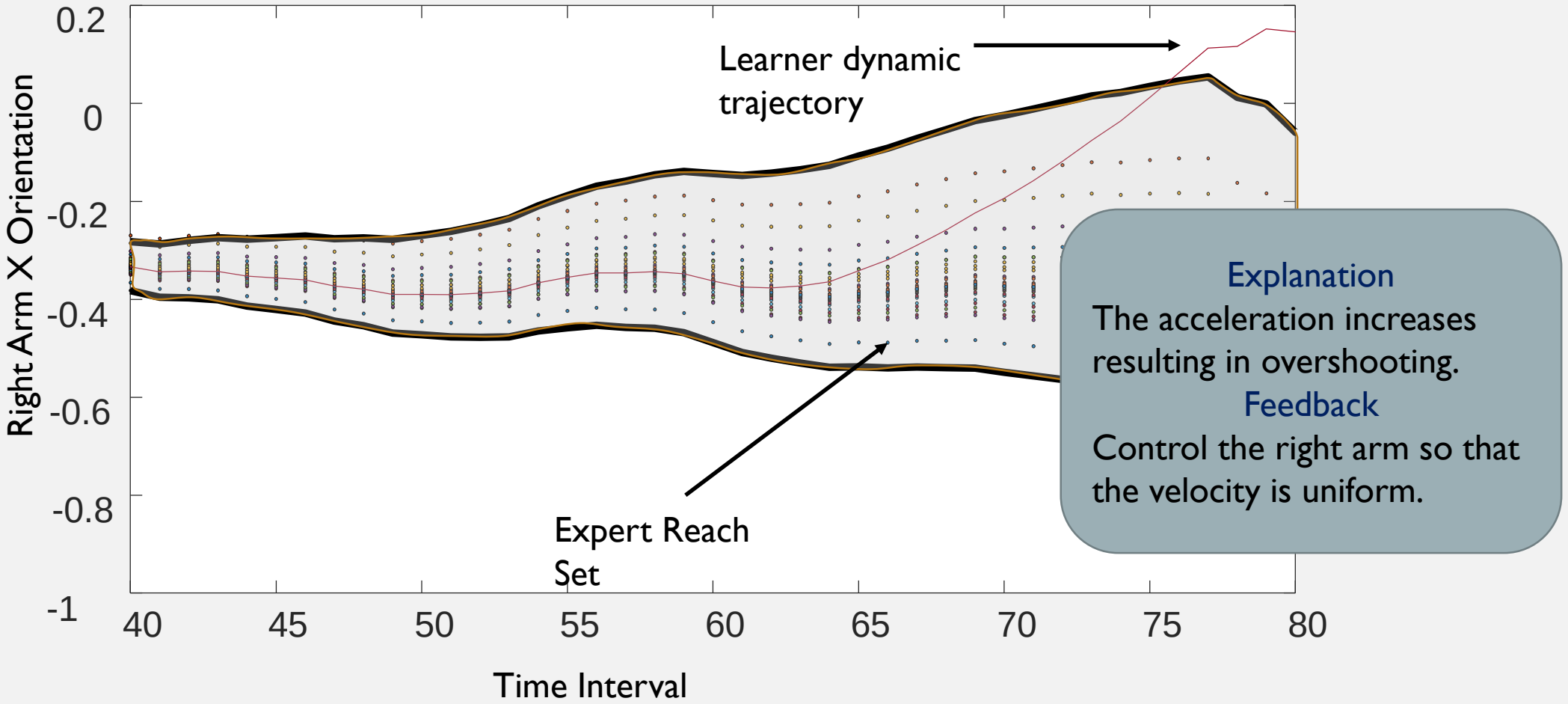
1. *Change-point detection using RuLSIF:*
   - ***Position A:*** Position of the back right hand on the top of the left hand.
   - ***Position B:*** Position of the left hand in front of the right hand.
   - ***Position C:*** Position of the left hand in front of the back of the left
2. *Multivariate Non-Linear Polynomial Regression Analysis:*

$$\frac{d\overleftarrow{p}}{dt} = \overleftarrow{v}\,, \frac{d\overleftarrow{v}}{dt} = \overleftarrow{a}, \frac{d\overleftarrow{a}}{dt} = x_1\,\overleftarrow{a} + x_2\,\overleftarrow{v} + x_3\,\overleftarrow{p} + x_4$$

$\overleftarrow{\boldsymbol{p}}$ is the position vector for the right arm
$\overleftarrow{\boldsymbol{v}}$ is the velocity vector
$\overleftarrow{\boldsymbol{a}}$ is the acceleration vector
$\boldsymbol{x_i}$ are parameters of the hand motion.



RESULTS



Explanation

The acceleration increases resulting in overshooting.

Feedback

Control the right arm so that the velocity is uniform.

CONCLUSION

- Applied software testing mechanism to generate explainable interfaces for an ML based gesture recognition system
- We propose a hybrid system model inference technique for a ML based gesture recognition system.
  - Modes are used to represent movement components.
  - Kinematic equations to describe the continuous spatio-temporal movement dynamics.
- Reachability analysis from the execution of an inferred expert model is used for corrective feedback.