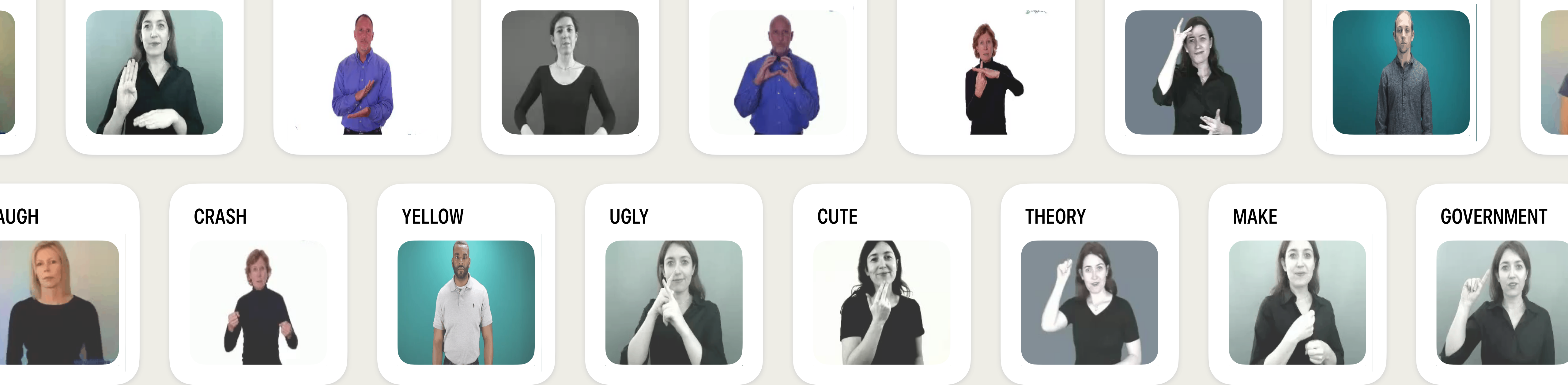




WEB4ALL 2025

Towards an AI-Driven Video-Based American Sign Language Dictionary

Exploring Design and Usage Experience with Learners

Saad Hassan¹, **Maty Bohacek**², Chaelin Kim¹, Denise Crochet¹

¹ Tulane University, ² Stanford University

[Video source: WLASL]

— Motivation

- Growing interest in ASL learning across various groups
 - Current challenges in ASL dictionary tools:
 - Feature-based dictionaries require linguistic knowledge
 - Text-based search ineffective for visual language
-

— Research Questions

- **RQ1.** How did ASL learners use a functional video-based ASL dictionary prototype to perform sign searches while answering comprehension questions about ASL videos?
 - **RQ2.** How do ASL learners believe their experience with the ASL dictionary can be improved?
 - **Contribution:** Fully automated dictionary code + insights
-

Contents

1. **Related Work**
 2. **Our Approach: AI Architecture**
 3. **Our Approach: User Interface**
 4. **Study: Methodology**
 5. **Study: Findings**
 6. **Concluding Remarks**
-

Related Work

Part 1/6

Related Work

- WoZ studies explored interface designs (Alonzo et al., 2019)
- Feature-based dictionaries, e.g. Handspeak (Lapiak, 2021)
- Video-based recognition advancements (Camgöz et al., 2020)
- **Gap:** Real-world usage with fully functional systems

Our Approach: AI Architecture

Part 2/6

— Our Approach: AI Architecture

- Video-based dictionary with AI recognition
 - State-of-the-art sign recognition model
 - Privacy-preserving video handling
 - Enhanced feedback on submissions
-

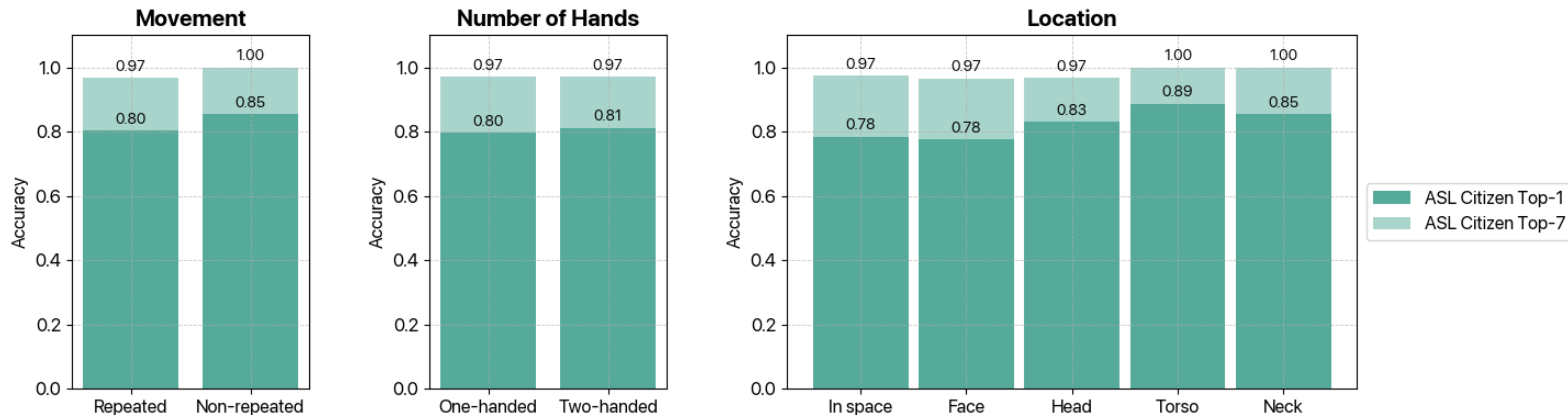
— Our Approach: AI Architecture

- Sign recognition model (SPOTER architecture), trained on the ASL Citizen dataset
 - Video quality assessment
 - Confidence communication
 - Result presentation
-

— Our Approach: AI Performance

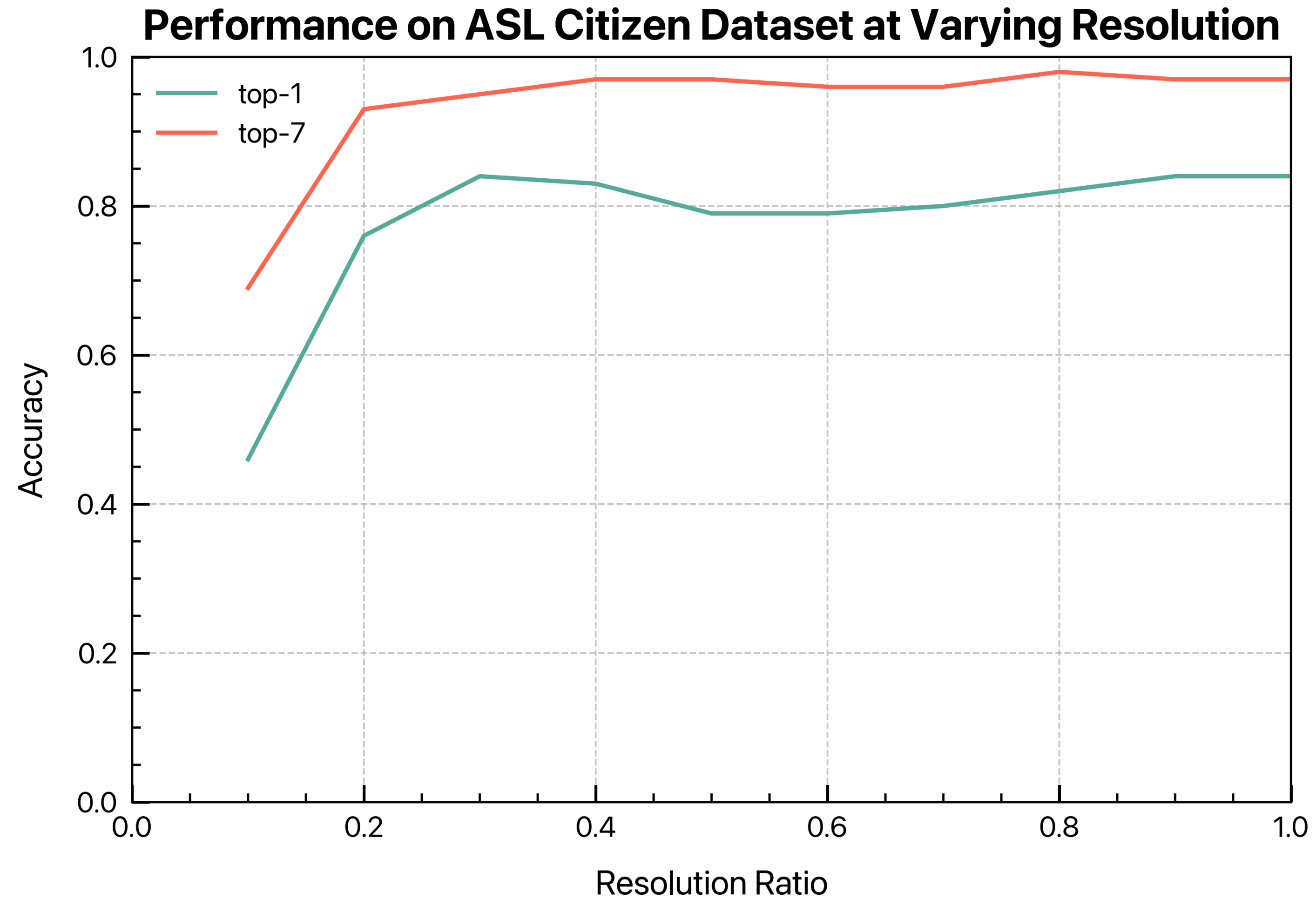
- Overall accuracy:
 - 81% top-1 accuracy
 - 97% top-7 accuracy
 - nDSG of 0.92
 - Performance varies by sign characteristics
 - Robust to lower resolution videos
-

Our Approach: AI Performance

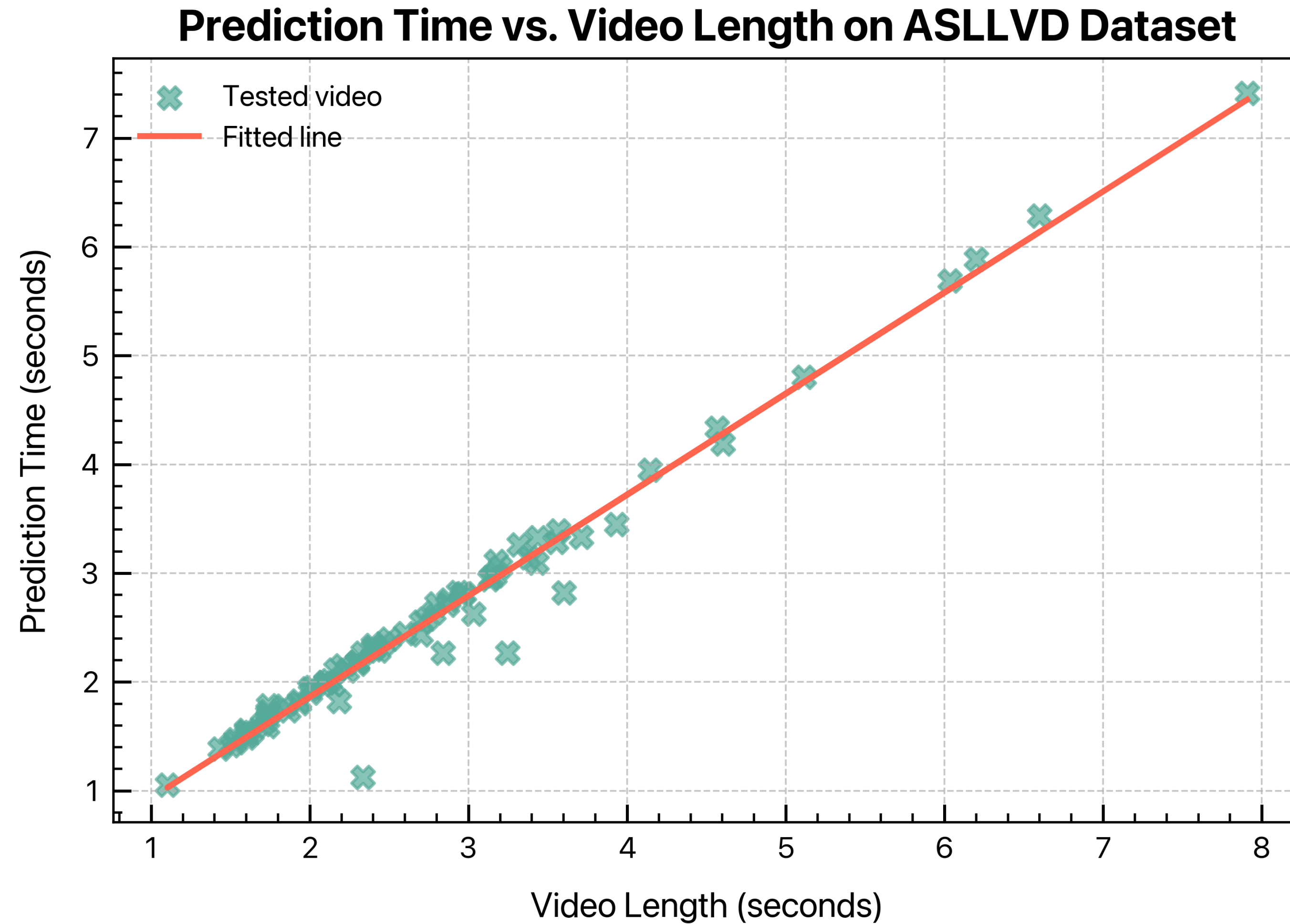


Performance grouped by movement (a), number of hands (b), and location (c).

Our Approach: AI Performance



Our Approach: AI Performance



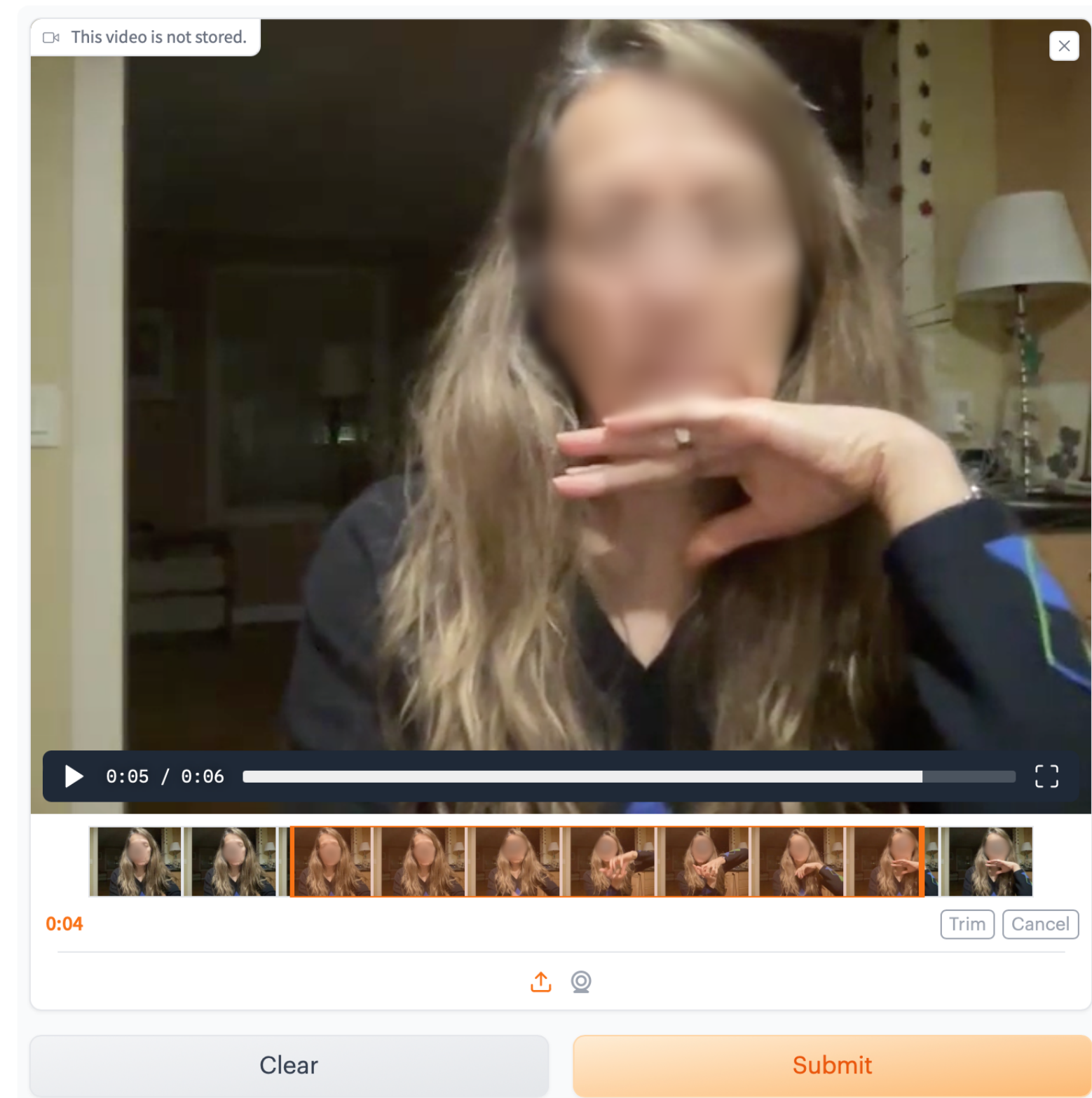
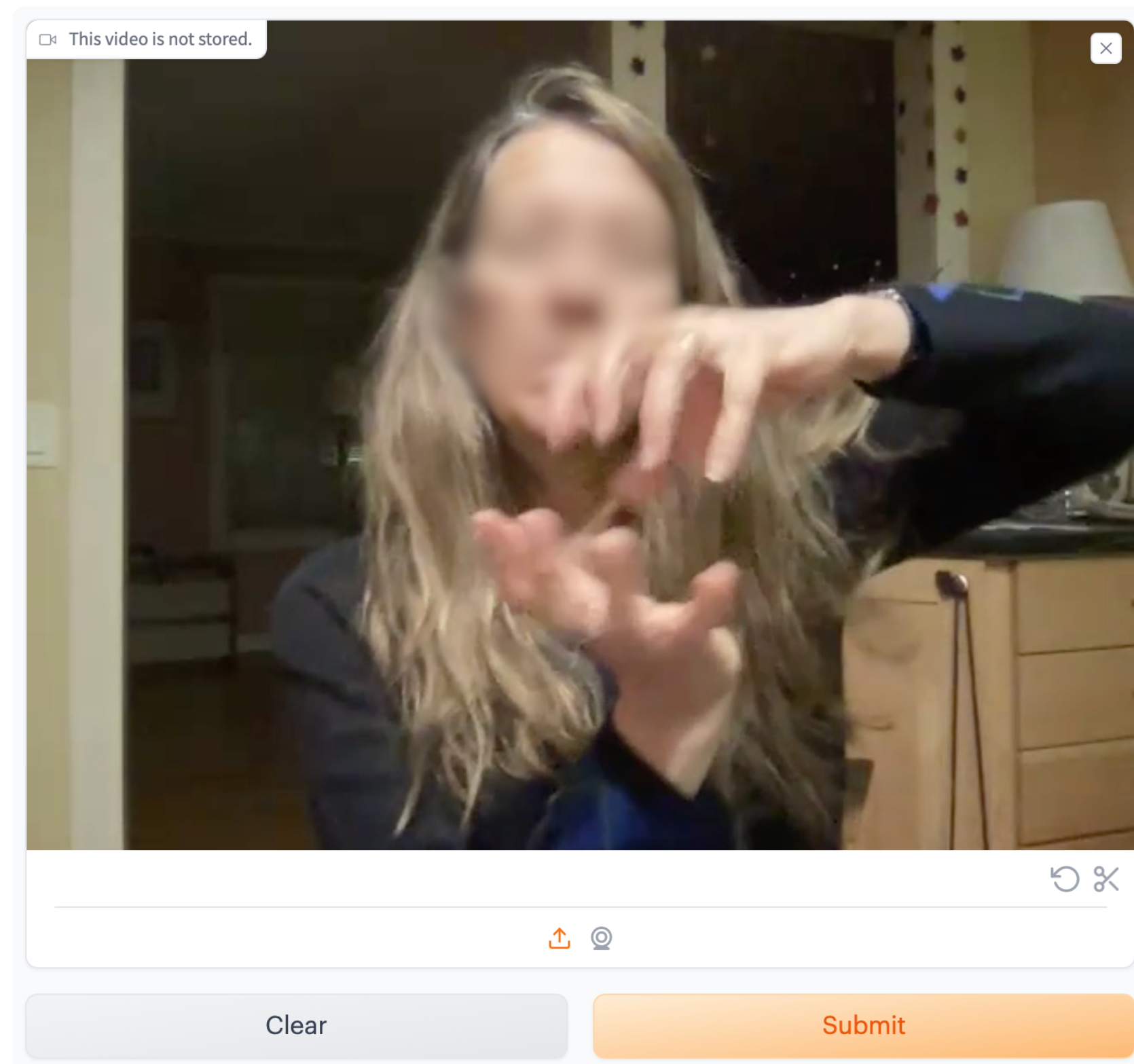
Our Approach: User Interface

Part 3/6

— Our Approach: User Interface

- Recording interface with trimming capability
 - Progress indicators for latency
 - Confidence labeling: Probably, Possibly, Unlikely
 - Filtering by linguistic properties
-

— Our Approach: User Interface



Recording Interface

Our Approach: User Interface

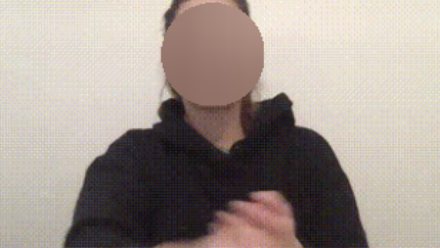
Top prediction

#1 Cookie Probably

Movement: repeated, Num. of hands: 2, Location: In space , Handshape: Closed 5 and clawed 5



Runner-up predictions

 #2 Cake Movement: repeated, Num. of hands: 2, Location: In space, Handshape: Closed 5 clawed 5	 #3 Toast Movement: repeated, Num. of hands: 2, Location: In space, Handshape: —
 #4 Squirrel Movement: repeated, Num. of hands: 2, Location: In space, Handshape: —	 #5 Crocodile Movement: repeated, Num. of hands: 2, Location: In space, Handshape: —
 #6 Sandwich Movement: repeated, Num. of hands: 2, Location: In space, Handshape: —	 #7 Rabbit Movement: repeated, Num. of hands: 2, Location: Side of head, Handshape: Closed 2/v

See full analysis for more details.

[More results](#)

Compact result view

Our Approach: User Interface

No video provided — please try again

- Record a video directly in the browser or upload a video from your computer.
- You can select the desired data source in the lower part of the left panel.
- To record a video or choose a file, click on the left panel.

Warning: Low video quality may affect the accuracy of sign language predictions.

- Ensure good lighting and minimal clutter in the background;
- Be in the center of the frame, with all your gestures visible;
- Allow maximum webcam quality.

Error and warning labels

— Our Approach: User Interface



Progress bar

Our Approach: User Interface

Movement ▾

Hands ▾

Locations ▾

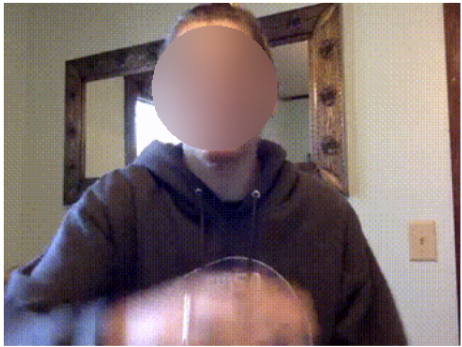
Handshape

Applied filter: None

Cookie

Probably

Repeated, Two-handed, In-space ,
Closed 5 and Clawed 5 handshape



Toast

Unlikely

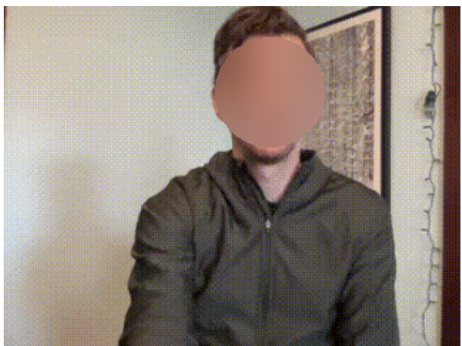
Repeated, Two-handed, In-space, —
handshape



Cake

Unlikely

Repeated, Two-handed, In-space,
Closed 5 Clawed 5 handshape



Rabbit

Unlikely

Repeated, Two-handed, Headside,
Closed 2/v handshape



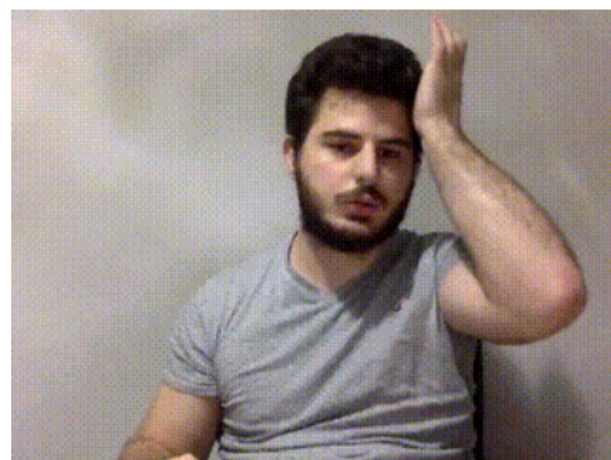
Detailed Analysis

Our Approach: User Interface

Lettuce

Probably

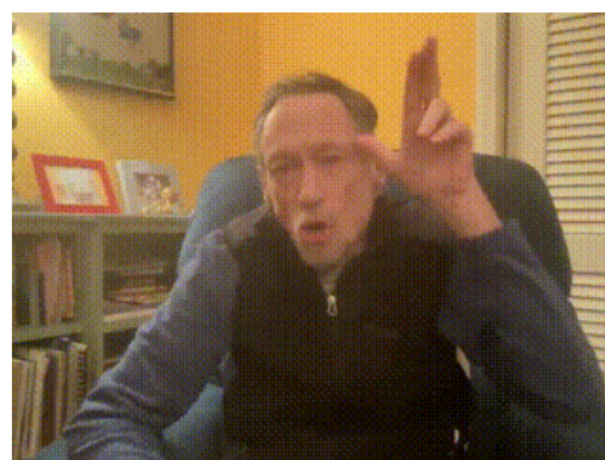
Uni-directional, One-handed,
Headside, — handshape



Horse

Possibly

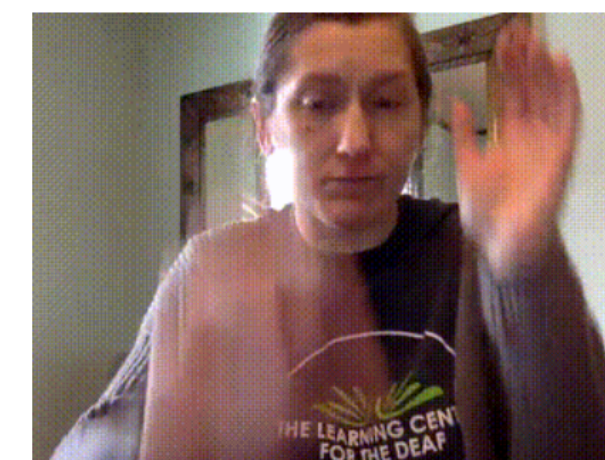
Repeated, Two-handed, Headside,
Closed 3 handshape



Donkey

Unlikely

Repeated, One-handed, Headside, —
handshape



Confidence labels

Study: Methodology

Part 4/6

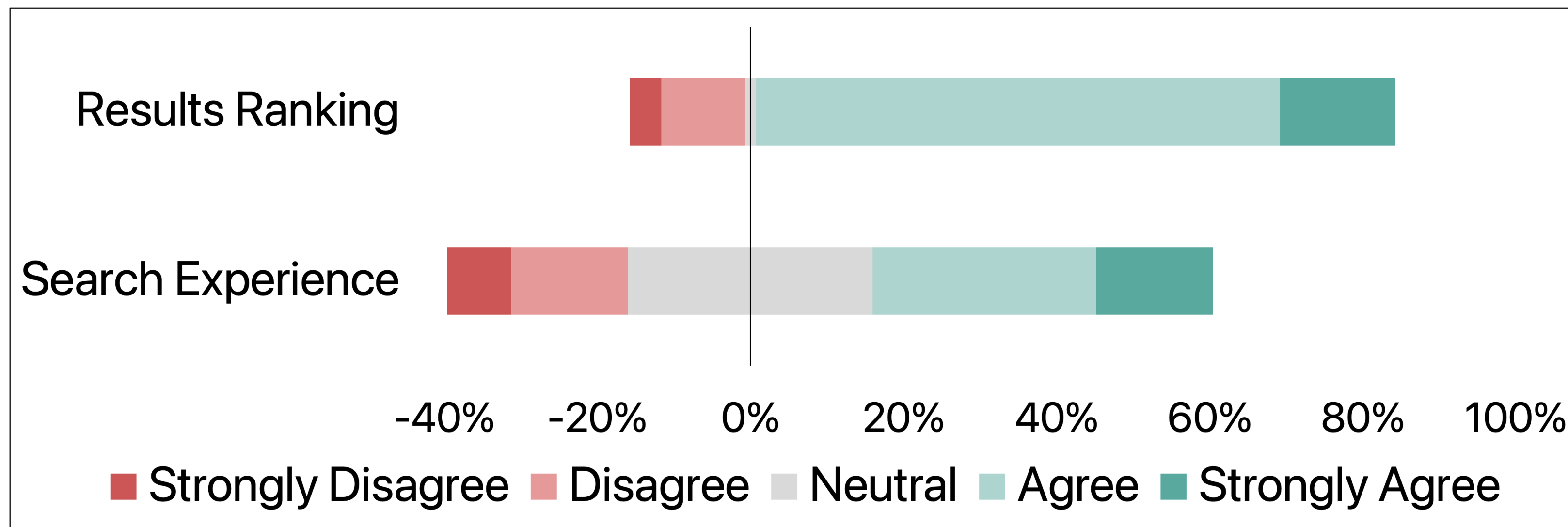
— Study: Methodology

- 12 ASL II students (novice learners)
 - Comprehension tasks with ASL videos
 - Latin square protocol for video assignment
 - Semi-structured interviews, screen recording and observation
-

Study: Findings

Part 5/6

Findings: Overall Impressions



Overall Usability Survey Responses

— Findings: Initial Impressions

- I never like using the other dictionaries... I did not know that we're on the way to, like, get the right signs that I want
 - Positive reception despite imperfections
 - Compared to "reverse image search" for signs
-

Findings: Learning Benefits

- "I didn't realize how similar these two signs were to each other. I'm gonna be careful when I sign with somebody"
- Beyond sign lookup:
 - Confirming sign understanding
 - Exposure to varied signing styles
 - Learning distinctions between similar signs

— Findings: Recording Challenges

- Difficulty isolating unfamiliar signs
 - Strategies for optimizing recordings:
 - Camera repositioning
 - Adjusting distance and signing space
 - Background considerations
-

— Findings: Navigating Unpredictability

- Self-attribution for failed searches
 - Appreciation for confidence indicators
 - Reattempt behaviors (3-5 tries typical)
 - User resilience to system limitations
-

— Findings: Latency

- "If it didn't have that loading bar, I'd probably think that the website was glitching out"
 - Progress indicators well-received
 - Tolerance for processing time
-

— Findings: Miscellaneous

- "For whatever reason, MILK would always come up as the sign"
 - Privacy considerations
 - Detection of biases in results
 - Some signs appeared repeatedly for certain users
-

Concluding Remarks

Part 6/6

— Design Implications

- Guidelines for video recording
 - Confidence presentation & latency status options
 - Privacy and data handling transparency
 - Post-query refinement approaches
-

— Limitations & Future Work

- Limited participant demographic
 - Further research opportunities: Broader user populations, Longitudinal studies, Larger vocabularies, Different sign languages
-

Conclusion

- First study of fully functional video-based ASL dictionary
 - Identified real-world usage patterns and challenges
 - Open-source prototype for future research
 - Guidance for designing and deploying video-based ASL dictionaries
-

Thank you.

Towards an AI-Driven Video-Based American Sign Language Dictionary: Exploring Design and Usage Experience with Learners



Saad Hassan



Maty Bohacek



Chaelin Kim



Denise Crochet