

Recognizing American Sign Language Nonmanual Signal Grammar Errors in Continuous Videos

Elahe Vahdani¹, Longlong Jing¹, Yingli Tian¹, Matt Huenerfauth²,

¹Ph.D. Program in Computer Science, The City University of New York

²The Rochester Institute of Technology

Jan 12th, 2021

ICPR 2020

ASL Linguistics Background

- Sign language communication:
 - Hand Gestures, Facial Expressions, Head Movements



CANNOT



DON'T KNOW



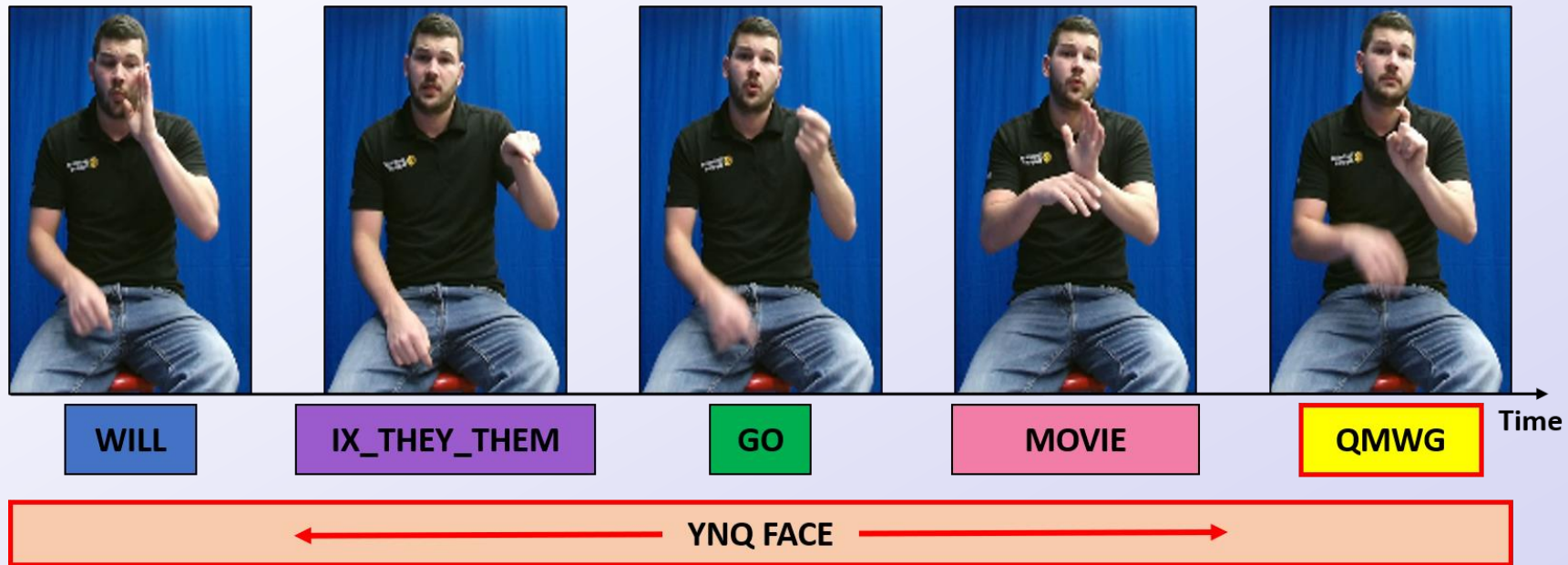
HOW



WHAT

Motivations and Goals

- Recognizing the **grammar** of ASL in continuous signing videos.




Will they go to a movie?

Motivations and Goals

- Technologies to support ASL education.
 - Immediate feedback regarding **grammar mistakes** in ASL videos.

Video Name



play 00:00 / 00:06

Feedback

Feedback

From time 0:00:02 to 0:00:02 : "When asking a WH question dont forget to lean forward (a little) and lower your eyebrows."

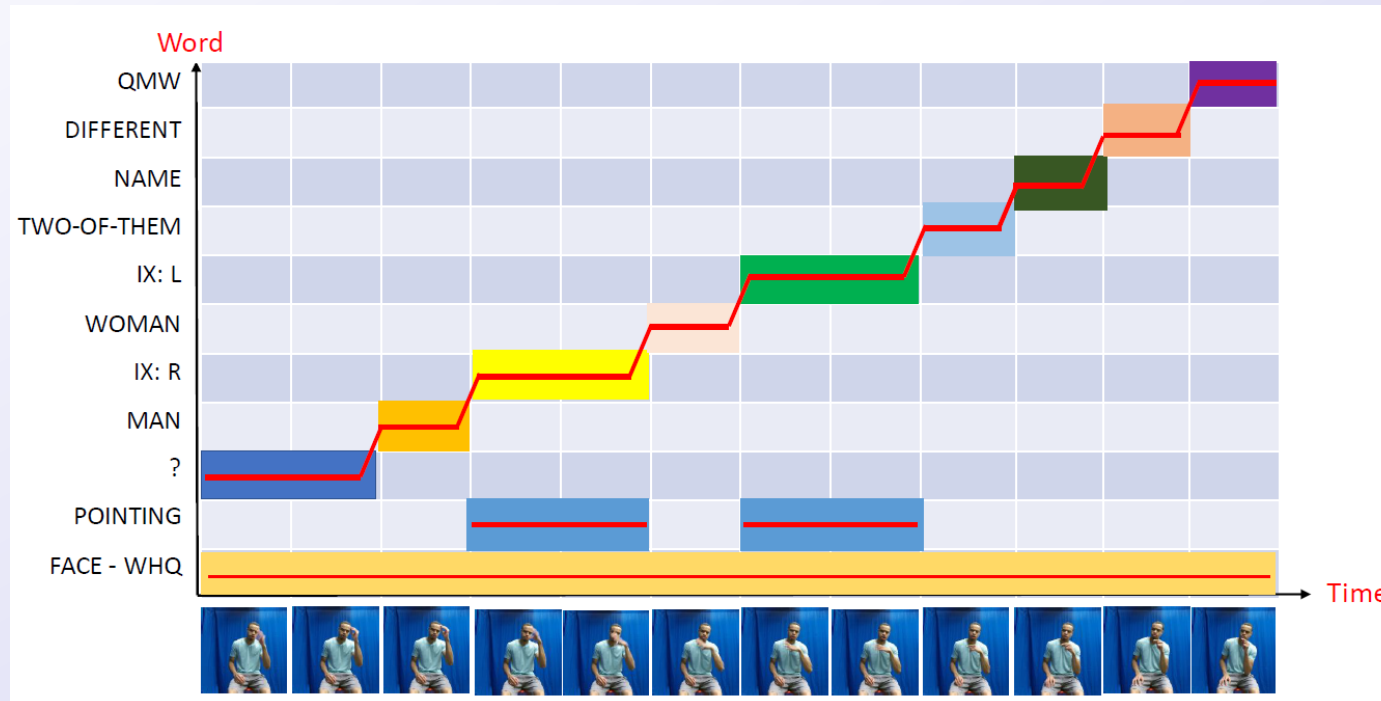
1

From time 0:00:01 to 0:00:02 : "When indicating something negative dont forget to furrow your eyebrows and shake your head (a little)."

2

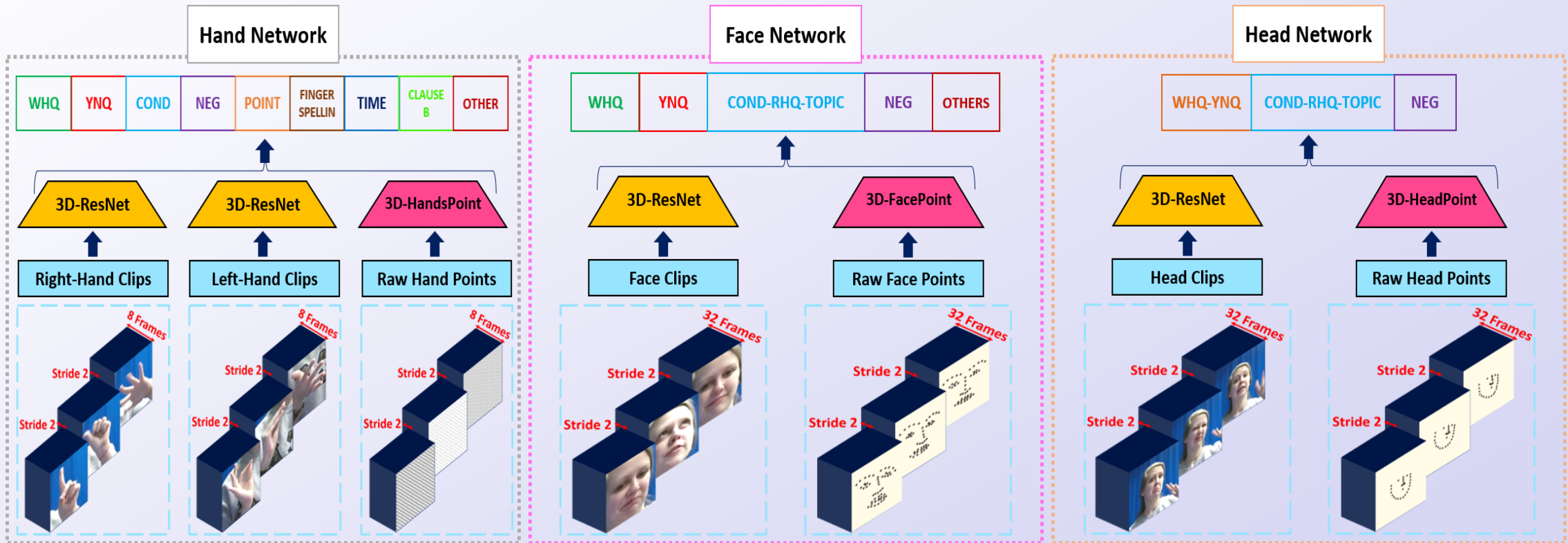
Challenges

- **Temporal Segmentation**
 - No **temporal boundaries** of sign.
 - Transition movements are subtle, diverse and hard to detect.



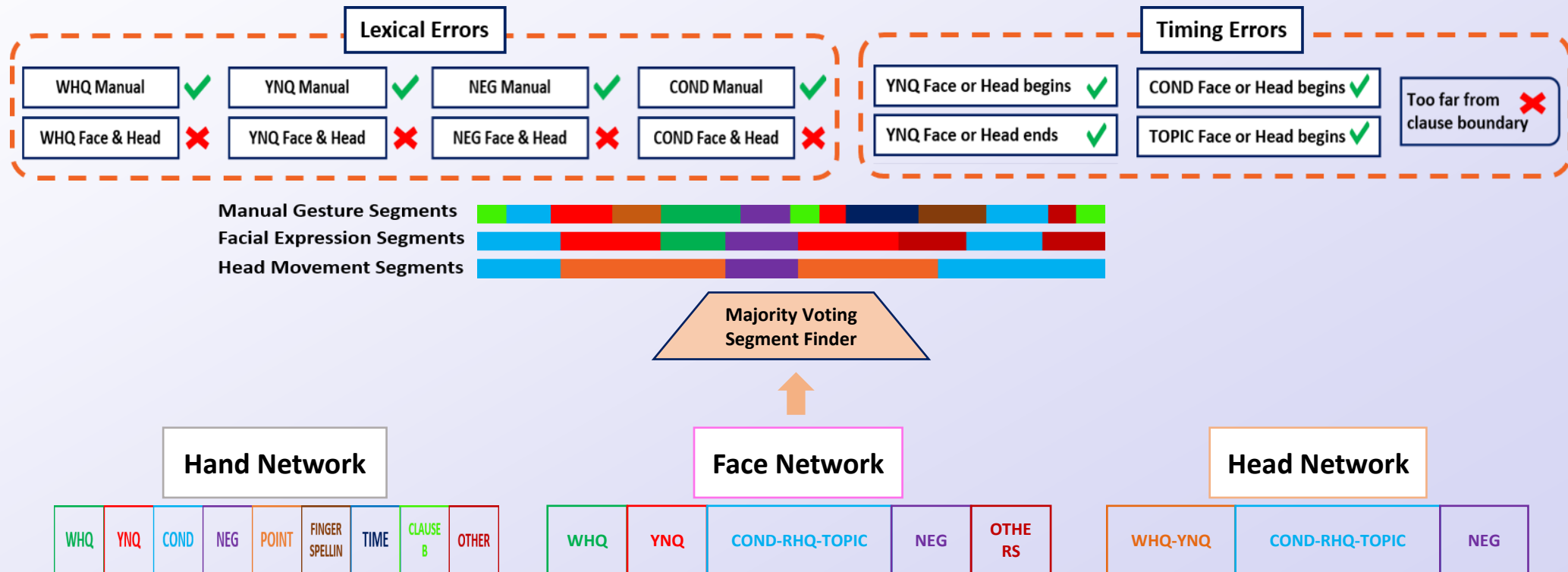
Framework

- 3D-ResNet for Spatiotemporal information.



Framework

- Correspondence between hand, face, and head network.

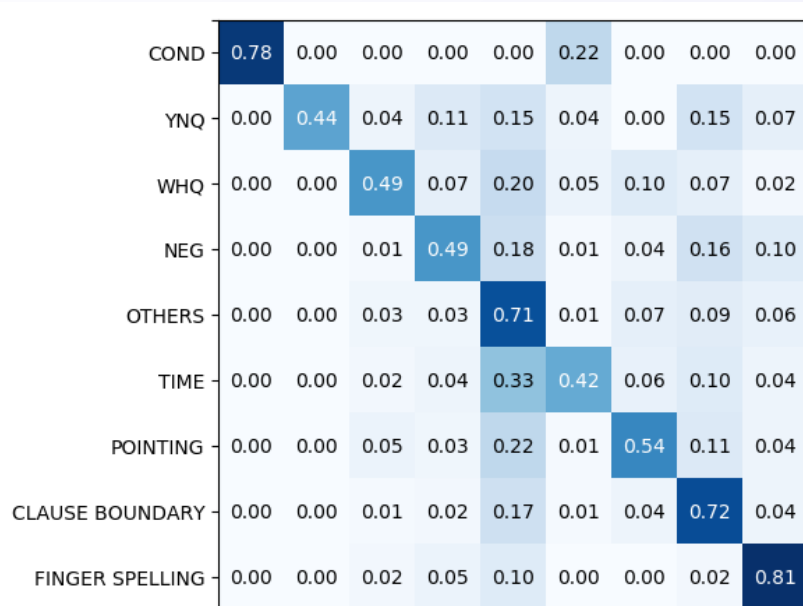


Dataset

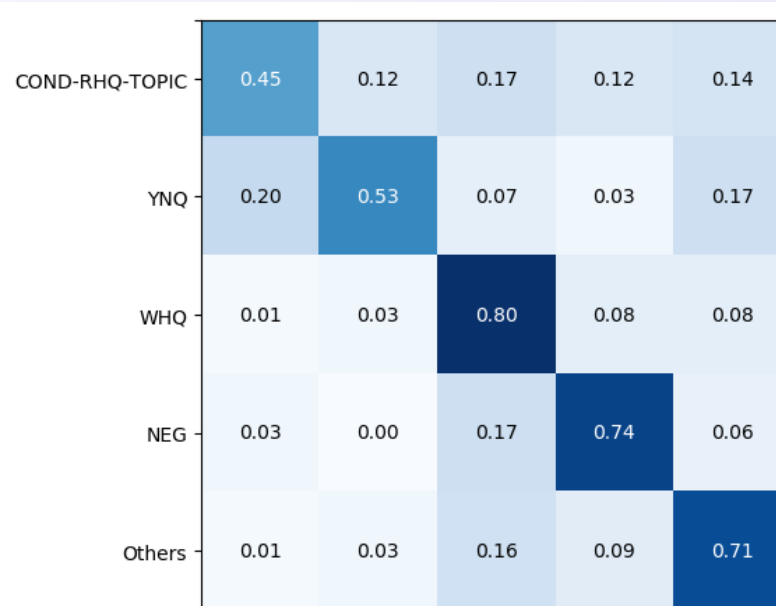
- **Continuous sign language dataset**
 - **ASL-HW-RGBD**
 - 1026 continuous videos
 - Signing by fluent and student signers
- **Different aspects of ASL grammar**
 - Conditional Sentences, WH Questions, YN Questions, .. and more.

Results

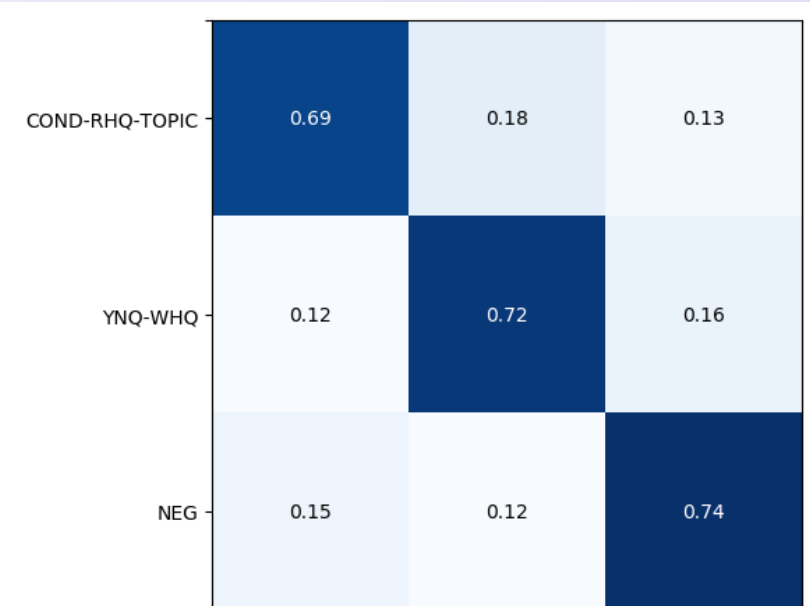
- Prediction results of the hand, face, head networks.



Hand Gesture Confusion Matrix



Facial Expression Confusion Matrix



Head Movements Confusion Matrix

Results

- **ASL nonmanual grammatical error recognition results.**

Results on ASL-HW-RGBD dataset

Error Type	Ground Truth	Recognized	TP Rate
Error-YNQ-Lexical	12	7	58.3%
Error-NEG-Lexical	36	14	38.8%
Error-WHQ-Lexical	40	23	57.5%
Error-COND-Lexical	24	18	75.0%
Error-TOPIC-Beginning	23	19	82.6%
Total	135	81	60.0%

Error Types

Error Type	Manual Sign	Nonmanual Signal
WHQ-Lexical	WHQ word	Not (WHQ or RHQ)
YNQ-Lexical	YNQ word	Not YNQ
NEG-Lexical	Negative word	Not Negative
COND-Lexical	Conditional word	Not Conditional
TOPIC-Beginning	Clause Boundary	Topic begins

Summary

- **Continuous sign language dataset: ASL-HW-RGBD**
- **Multimodal 3DCNN framework for hands, face, and head movements.**
- **Detecting ASL grammar mistakes from videos.**
- **Provide feedback for ASL students whether their signing is fluent or not.**