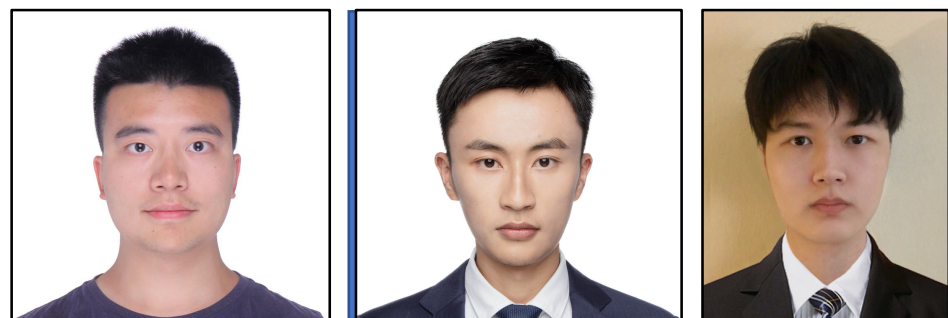


# Sign Language Video Anonymization

Zhaoyang Xia<sup>1</sup>, Yuxiao Chen<sup>1</sup>, Qilong Zhang<sup>1</sup>, Matt Huenerfauth<sup>2</sup>, Carol Neidle<sup>3</sup>, Dimitris Metaxas<sup>1</sup>

<sup>1</sup>Rutgers University; <sup>2</sup>Rochester Institute of Technology; <sup>3</sup>Boston University

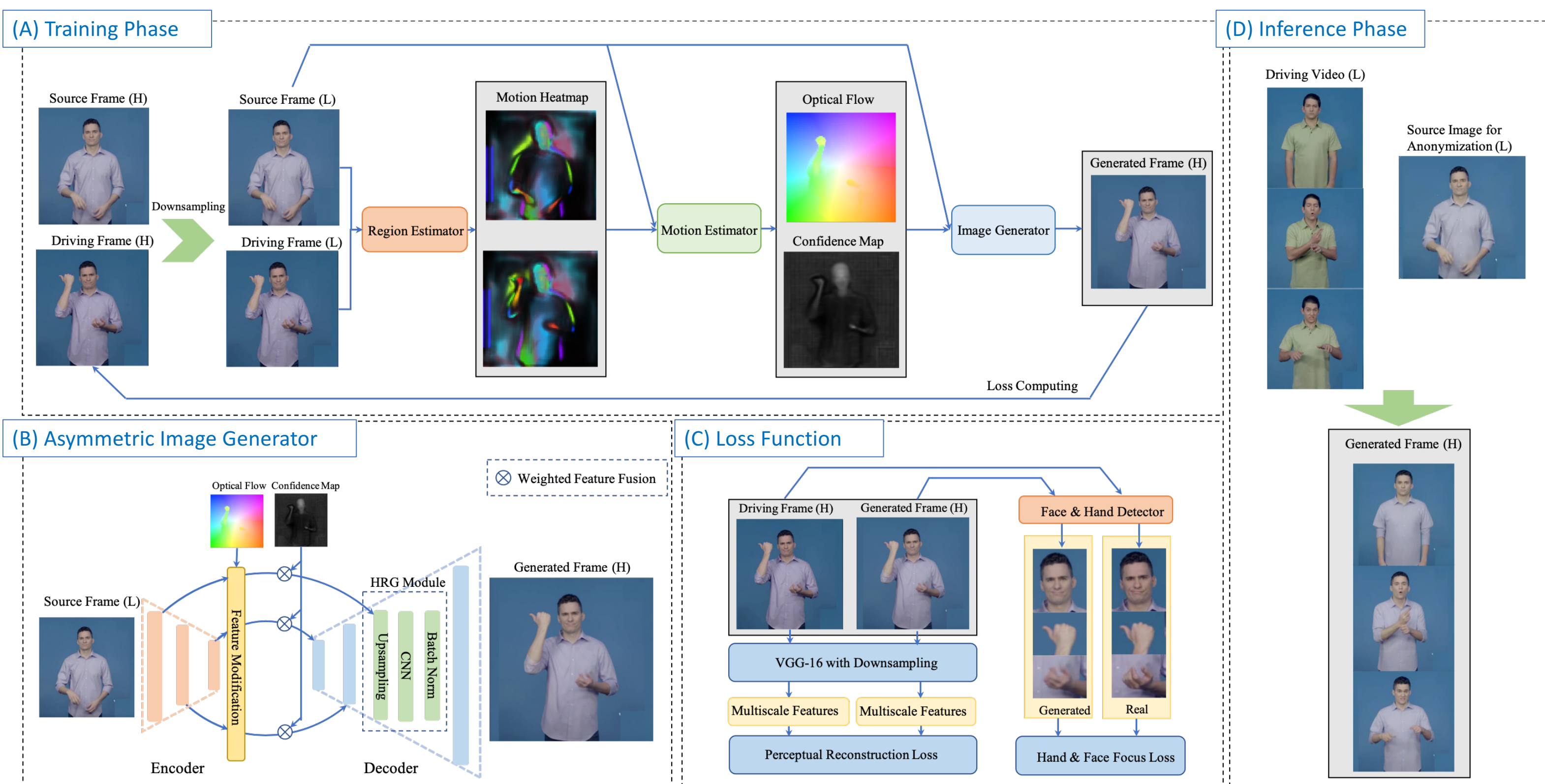


## Why Is Anonymization Needed?

- There is no standard written form for ASL, so, to communicate in their native language, signers frequently share videos.
- Critical linguistic information is expressed by face, arms & hands; these cannot be obscured (e.g., for sharing sensitive information on the Web).
- Video anonymization will have many applications, e.g. enabling anonymous peer review for academic submissions in ASL and neutrality for various ASL video resources; and increasing participation in video-based AI databases.

## Method Overview

- (A) Training Phase:** Model is trained to generate the driving frame from the source frame.
- (B) The Asymmetric Image Generator:** An encoder-decoder structured network with a High-Resolution Generation (HRG) module.
- (C) Multiscale perceptual loss** based on VGG-16 & **Hand & Face Focused Loss:** Computed between the high-resolution generated and driving frames, to improve face & hand generation.
- (D) Inference Phase:** The source image for anonymization and the extracted frames from the driving video are input to the model to generate the high-resolution anonymized videos.



## Results

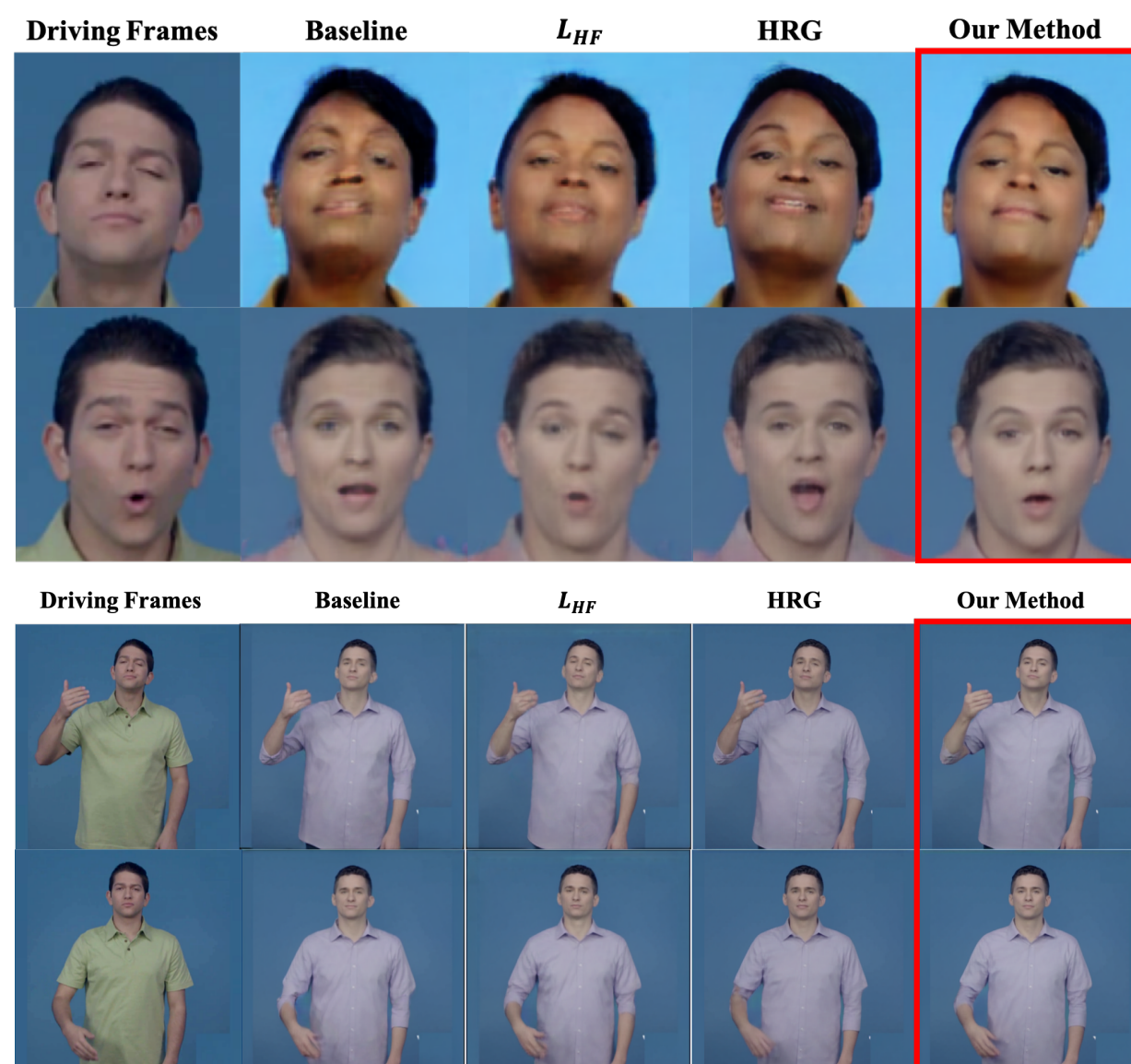
### Anonymized Video Examples



**Data Set** We trained our model on the American Sign Language Linguistic Research Project (ASLLRP) Continuous Signing Corpora\*. We selected 1,054 videos.

### Evaluation

Our method improves the video generation quality. The  $L_{HF}$  module improves generation of hands and face. The HRG module improves the resolution of the generated videos.



See the paper for KID and FID scores demonstrating the positive effects on realism from the enhanced High-Resolution Image Generation and the Hand & Face Focused Loss Function.

## Conclusion

We have been developing methods for anonymizing ASL videos with input from Deaf signers. Lee *et al.* (2021)\* reports on user studies with Deaf signers who evaluated our earlier results from using facial expression transfer for ASL video anonymization. Here, we extend the anonymization to the whole body. Preliminary user interviews indicate that our new method is extremely promising. We will continue refining the model for more accurate hand and face generation.

\*Data from <https://dai.cs.rutgers.edu/dai/s/dai>. For information about our previous work on facial expression transfer for video anonymization, see Lee *et al.* (2021) American Sign Language Video Anonymization to Support Online Participation of Deaf and Hard of Hearing Users. *ASSETS '21*. Article 22, 1–13.

\*The work reported here has been supported in part by grants from NSF: no. 2040638, 1763486, 1763523, and 1763569.