

Zhaoyang Xia

🌐 jeffery9707.github.io ✉ zx149@rutgers.edu ☎ 908-217-3973

Intern Availability: Start: May 5th, End: September 9th

EDUCATION

Rutgers University

Sept. 2021 – Expected 2026

- Ph.D. in Computer Science, GPA: 4.0/4.0
- Advisor: [Dimitris Metaxas](#)

Rutgers University

Sept. 2019 – May. 2021

- M.S. in Computer Science (Data Science), GPA: 3.91/4.0

Fudan University

Sept. 2015 – Jun. 2019

- B.S. in Information and Computing Science (Data Science & Technology)

PUBLICATIONS

1. **Xia, Zhaoyang**, Carol Neidle, and Dimitris N Metaxas. DiffSLVA: Harnessing Diffusion Models for Sign Language Video Anonymization. *arXiv preprint arXiv:2311.16060*, 2023 [\[PDF\]](#) [\[Demo\]](#)
2. **Xia, Zhaoyang**, Yuxiao Chen, Qilong Zhangli, Matt Huenerfauth, Carol Neidle, and Dimitri Metaxas. Sign Language Video Anonymization. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources (LREC)*, pages 202–211, 2022 [\[PDF\]](#) [\[Demo\]](#)
3. Sooyeon Lee, Abraham Glasser, Becca Dingman, **Xia, Zhaoyang**, Dimitris Metaxas, Carol Neidle, and Matt Huenerfauth. American Sign Language Video Anonymization to Support Online Participation of Deaf and Hard of Hearing Users. In *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility (ASSET)*, pages 1–13, 2021 [\[PDF\]](#) [\[Demo\]](#)
4. Yuxiao Chen, Long Zhao, Jianbo Yuan, Yu Tian, **Xia, Zhaoyang**, Shijie Geng, Ligong Han, and Dimitris N Metaxas. Hierarchically Self-supervised Transformer for Human Skeleton Representation Learning. In *European Conference on Computer Vision (ECCV)*, pages 185–202. Springer, 2022 [\[PDF\]](#)
5. Qilong Zhangli, Jingru Yi, Di Liu, Xiaoxiao He, **Xia, Zhaoyang**, Qi Chang, Ligong Han, Yunhe Gao, Song Wen, Haiming Tang, et al. Region proposal rectification towards robust instance segmentation of biological images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 129–139. Springer, 2022 [\[PDF\]](#)
6. Ligong Han, Song Wen, Qi Chen, Zhixing Zhang, ... **Xia, Zhaoyang**, Akash Srivastava, and Dimitris N Metaxas. Improving Negative-Prompt Inversion via Proximal Guidance. *arXiv preprint arXiv:2306.05414*, 2023 [\[PDF\]](#)

RESEARCH EXPERIENCE

DiffSLVA: Harnessing Diffusion Models for Sign Language Video Anonymization [\[Demo\]](#)

PhD Student, Rutgers University

- Proposed zero-shot text-guided sign language anonymization. Designed methods for accurate gestures and facial expression transferring for ASL videos with *Stable Diffusion* and the *Image Animation* module.
- Applied cross-frame attention mechanism and optical flow guided latent fusion method with ControlNet for consistent video editing.

Improving Diffusion Models with Human Preference

Research Scientist/Engineer Intern, Adobe Inc.

- Designed algorithm for improving diffusion model with *Human Preference* data.
- Enhanced image generation quality by utilizing representations from the UNet for guidance during inference stage, focusing on aligning the output with human aesthetic preferences.

Open Set Semi-Supervised Image Generation with GANs

PhD Student, Rutgers University

- Constructed open set data set on ImageNet100, CIFAR10, and CIFAR100. Generated pseudo labels for open set instances by various methods, such as SimCLR with k-means, and ORCA model.
- Experimented Open Set Semi-Supervised Image Generation with BigGAN and StyleGAN-XL model.
- Achieved comparable FID with only 5% labeled data on CIFAR10 and CIFAR100.

Sign Language Video Anonymization [\[Demo\]](#)

PhD Student, Rutgers University

- Proposed a motion-based **Image Animation** model for the sign language video anonymization task to generate high-resolution videos with the signer identity changed, but with the preservation of linguistically significant motions and facial expressions.
- Proposed an asymmetric encoder-decoder structured image generator for high-resolution image generation. Explicitly guided the model to attain a clear generation of hands and faces by using bounding boxes to improve the loss computation.

American Sign Language Video Anonymization to Support Online Participation of Deaf and Hard-of-Hearing Users [\[Demo\]](#)

PhD Student, Rutgers University

- Applied First Order Motion Model for **Face Swap** to automatically disguise the face in sign language videos while preserving essential facial expressions and natural human appearance.
- Incorporated deep learning model and color-based model for **Skin Segmentation** to improve anonymization results.

Non-manual Grammatical Marker Detection for American Sign Language

PhD Student, Rutgers University

- **Action Detection** in videos containing continuing sentences of ASL.
- Applied Inflated Inception-v1 model with sliding window method to predict the probability of start/end point for facial actions.
- Successfully generated action proposals based on the predicted probability and provided a relatively balanced data set to prove the performance of action classification.

Explainable Recommendation System for Movies

Researcher, Computer Science Department of Fudan University

- Extracted features as tag preference and tag relevance from movie data.
- Utilized Explicit Factor Model based on features and rates of movies to do **Explainable Recommendations**.

WORK HISTORY

Research Scientist/Engineer Intern— Adobe Inc.

May.2023 – Dec.2023

- Designed algorithms for improving diffusion models with human preference data

Intern— MISUMI (China) precision machinery trading co., LTD

Nov.2018 – Feb.2019

- Predicted the loss of customers by applying LSTM neural network on customers' behavior data.
- Analysed the customers' chat context by applying the textrank model, word2vec embedding, and k-means clustering.

SKILLS

Programming Languages: Python, SQL, R

Frameworks: PyTorch, OpenCV

Academic Service: Reviewer for ECCV, CVPR