Zhaoyang Xia

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

EDUCATION

Rutgers University   Sept. 2021 – Expected 2025

- Ph.D. in Computer Science, GPA: 4.0/4.0
- Advisor: Dimitris Metaxas

Rutgers University   Sept. 2019 – July. 2021

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

Fudan University   Sept. 2015 – July. 2019

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

PUBLICATIONS

1. **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]

2. 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]

3. 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]

4. 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]

5. Di Liu, Yunhe Gao, Qilong Zhangli, Ligong Han, Xiaoxiao He, **Xia, Zhaoyang**, Song Wen, Qi Chang, Zhennan Yan, Mu Zhou, et al. Transfusion: multi-view divergent fusion for medical image segmentation with transformers. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (**MICCAI**)*, pages 485–495. Springer, 2022 [PDF]

RESEARCH EXPERIENCE

Image/Video Style-Transfer/Editing with Diffusion Models
PhD Student, Rutgers University

- Currently working on image and video editing with ***Diffusion Models***.
- Experimented with DiffusionCLIP for image style transfer with text guidance.
- Utilized Stable Diffusion for image-to-image and depth-to-image transformation.
- Aimed to design new methods for video transformation/editing with improved frame consistency.

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 state-of-the-art FID with only 5% labeled data on CIFAR10 and CIFAR100. Achieved comparable FID with 5% labeled data on ImageNet100 data set.

Sign Language Video Anonymization   [Project 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   [Project 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**.

Knowledge Graph Integration
Researcher, Computer Science Department of Fudan University

- Applied text analysis and SQL to extract all the entities and relationships from different databases.
- Extracted mapping tables of keys from databases and constructed connections between knowledge bases. Building a new knowledge graph using Neo4j.

WORK HISTORY

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