Qilong Zhangli

qzhangli.github.io qilongzhangli@gmail.com qzhangli (949)880-6315

EDUCATION

Rutgers University Sept. 2021 – Present

- **Ph.D. in Computer Science**
- Advisor: Dimitris Metaxas

University of California, Irvine Sept. 2018 – June 2021

- **B.S. in Computer Science** GPA: 3.85/4.0
- SURP Fellowship
- Specialized in Intelligent Systems

RESEARCH EXPERIENCES

Meta Menlo Park, CA

Research Scientist Intern May 2023 – November 2023

- Scene Text Image Generation with Diffusion Models: Significantly enhanced the capability of existing text-to-image diffusion models to produce images with text that is visually appealing and contextually coherent, achieved without the need for spatial information or predefined layouts as input. (CVPR 2024)

NEC Labs America San Jose, CA

Research Intern March 2023 – May 2023

- Language-driven Multi-dataset Panoptic Segmentation: Developed advanced panoptic segmentation methods using high-capacity vision-language models. This work bridges the gap between user-specified linguistic cues & visual segmentation tasks, particularly in multi-dataset setting. (ECCV 2024 in Submission; First Author)

Rutgers University New Brunswick, NJ

Graduate Assistant Sept. 2021 – Present

- Currently working on using diffusion models for complex scene image understanding/generation.
- Top-Down Instance Segmentation: Developed Region Proposal Rectification(RPR) module, a Transformer-based module for robust detection and segmentation that out-performance the baseline methods on both anchor-based and anchor-free approaches.
- Multi-modality Segmentation with Transformers: Developed TransFusion, a Transformer-based architecture for rich cross-view context modeling and semantic dependency mining, addressing the critical issue of capturing long-range correlations between unaligned data from different image modalities.
- American Sign Language Recognition: Developed skeleton-based isolated sign recognition model using Graph Neural Network and Transfer Learning. Worked on self-attention mechanism in spatial and temporal domains for accurate hand gesture recognition.

California Plug Load Research Center (CalPlug) Irvine, CA

Student Researcher March 2020 – May 2021

- Harnessing Machine-Learning to Personalize Cleft Lip Markings: Developed model for generating optimal cleft markings and projecting them onto surgical sites with the application of the High-Resolution Neural Network (HR-Net).
- Plug Load Energy Usage Simulator: Developed a plug load simulator using Python, which precisely calculates the energy consumption, and worked on debugging/troubleshooting issues in the software before launching.
- Micro/nanobubble Machine: Designed the UI for the LCD touch screen and developed the functional breadboard program so that users could adjust the machine's running time and the medicine dosage.

SELECTED PUBLICATIONS

1. Qilong Zhangli, Jindong Jiang, Di Liu, Licheng Yu, Xiaoliang Dai, Ankit Ramchandani, Guan Pang, Dimitris N Metaxas, and Praveen Krishnan. Layout-agnostic scene text image synthesis with diffusion models. In *The IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024

2. Qilong Zhangli, Dimitris N. Metaxas, and Samuel Schulter. RESI: Resolving inconsistent semantics in multi-dataset panoptic segmentation. In *European Conference on Computer Vision(In Submission)*, 2024
3. Ligong Han, Song Wen, Qi Chen, Zhixing Zhang, et al. ProxEdit: Improving tuning-free real image editing with proximal guidance. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 4291–4301, 2024
4. Di Liu, Qilong Zhangli, Yunhe Gao, and Dimitris Metaxas. Lepard: Learning explicit part discovery for 3d articulated shape reconstruction. *Advances in Neural Information Processing Systems*, 36, 2024
5. Qilong Zhangli, Jingru Yi, Di Liu, Xiaoxiao He, Zhaoyang Xia, 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*, pages 129–139. Springer, 2022
6. Di Liu, Xiang Yu, Meng Ye, Qilong Zhangli, Zhuowei Li, Zhixing Zhang, and Dimitris N Metaxas. DeFormer: Integrating transformers with deformable models for 3d shape abstraction from a single image. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14236–14246, 2023
7. Qi Chang, Zhennan Yan, Mu Zhou, Di Liu, Khalid Sawalha, Meng Ye, Qilong Zhangli, Mikael Kanski, Subhi Al'Aref, Leon Axel, et al. Deeprecon: Joint 2d cardiac segmentation and 3d volume reconstruction via a structure-specific generative method. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 567–577. Springer, 2022
8. Zhaoyang Xia, Yuxiao Chen, Qilong Zhangli, Matt Huenerfauth, Carol Neidle, and Dimitris Metaxas. Sign language video anonymization. In *Proceedings of the LREC2022 10th Workshop on the Representation and Processing of Sign Languages: Multilingual Sign Language Resources, Marseille, France, 25 June 2022*, 2022
9. Di Liu, Yunhe Gao, Qilong Zhangli, Ligong Han, Xiaoxiao He, Zhaoyang Xia, 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*, pages 485–495. Springer, 2022
10. Lohrasb Ross Sayadi, Usama S Hamdan, Qilong Zhangli, and Raj M Vyas. Harnessing the power of artificial intelligence to teach cleft lip surgery. *Plastic and Reconstructive Surgery–Global Open*, 10(7):e4451, 2022

INDEPENDENT PROJECTS

Kaggle Competition on Rainfall Prediction May. 2020 – June 2020

- Implemented gradient-boosted classification models (e.g., random forest, SVM) on satellite infrared imaging data, accurately predicting rainfall occurrences with 80% accuracy, ranking in the top 10%.

ChatbotSteve NLP Oct. 2020 - Dec. 2020

- Developed an NLP algorithm for character control in Malmo, enabling execution of complex commands from human speech, inlcuding basic movements, intricate interactions like identifying and interacting with entities.

PRESENTATIONS

1. Combining AI and AR for knowledge and skill transfer in cleft surgery. *4th International Comprehensive Cleft Care Workshop (CCCW)*, Oct 7, 2021
2. AI/ML-based facial analytics for natural language. *National Science Foundation(NSF): Convergence Accelerator Expo 2021*, July 2021
3. Using machine learning to develop an artificial intelligence algorithm that guides nasolabial repair (oral). *78th Annual Meeting of the American Cleft Palate-Craniofacial Association (ACPA)*, Oct. 2021
4. Harnessing machine learning to place cleft lip markings (oral). *89th Annual Meeting of the American Society of Plastic Surgeons (ASPS)*, Oct. 2020

SKILLS

- **Tools**: PyCharm, AWS, Google Cloud Platform, PyTorch, TensorFlow, OpenCV, detectron2, diffusers.
- **Computer Vision**: Detection, Semantic/Instance/Panoptic Segmentation, Vision-Language Models, Latent Diffusion Models, Generative Models.
- **Academic Service**: Reviewer for CVPR, ECCV, MICCAI