Susan Liang

(+86) 13861292824 | liangsusan18@mails.ucas.ac.cn | liangsusan-git.github.io
Senior Undergraduate, Computer Science

Education

University of Chinese Academy of Sciences (UCAS) Beijing, China
Bachelor of Computer Science *Sept. 2018 – Now*

- General GPA: 3.89/4 Rank: 7/108

Research Interests

- Multi-modal Learning
- Self-supervised Learning

Research Experience

Semantic Correspondence Sept. 2021 – Now
Vision and Learning Lab, ***University of California, Merced***
Advisors: Prof. Ming-Hsuan Yang and Dr. Taihong Xiao

- Proposed a self-supervised deep learning approach for semantic correspondence.
- Exploited contrastive learning and cycle consistency to learn discriminative and consistent features.

Vector Graphics Learning and Generation Jun. 2021 – Aug. 2021
Institute for AI Industry Research, ***Tsinghua University***
Advisors: Dr. Yizhi Wang and Dr. Hao Xu

- Developed an encoder-decoder model to convert raster images to vector graphics.
- Used a differentiable rasterization pipeline to enable supervision by raster images.

Active Speaker Detection Oct. 2020 – Apr. 2021
VIPL, Institute of Computing Technology, ***Chinese Academy of Sciences***
Advisors: Prof. Shiguang Shan and Dr. Shuang Yang

- Developed an audio-visual multi-modal fusion scheme to detect when each visible speaker in the video is speaking.
- Proposed an permutation-equivariant layer with the capability of processing all speakers in the scene simultaneously.
- Exploited the skew-symmetry of inter-speaker relations which not only has reasonable interpretation but also reduces memory usage and FLOPS.
- Conducted extensive experiments on multiple datasets (AVA-ActiveSpeaker, Columbia and RealVAD) with outstanding performance.

Face Deformation Field Generation and Lip Reading Feb. 2020 – Sept. 2020
VIPL, Institute of Computing Technology, ***Chinese Academy of Sciences***
Advisors: Prof. Shiguang Shan and Dr. Shuang Yang

- Developed an encoder-decoder model to generate face deformation field (face-specific optical flow) which features the face motion.
- Trained the deformation field in a self-supervised manner with no annotations.
- Combined deformation field and gray-scale face images to recognize visual speech.

Publication

Yuanhang Zhang*, **Susan Liang***, Shuang Yang, Xiao Liu, Zhongqin Wu, Shiguang Shan, Xilin Chen
UniCon: Unified Context Network for Robust Active Speaker Detection
ACM International Conference on Multimedia (ACM MM), Oct. 2021 **(Oral)**

* indicates equal contribution.

Awards

ActivityNet CVPR 2021 Workshop AVA Active Speaker Detection Challenge **First Place** Jun. 2021
UCAS Overseas Graduate Studies Fellowship Aug. 2021

Technical Skills

Languages: Proficient in Python and C; Familiar with LaTeX
Frameworks: Proficient in PyTorch; Familiar with TensorFlow and PyTorch-Lightning
Developer Tools: Git, Docker, and Vim