Ziyao Zeng (Adonis)

Last Update: Nov. 2022

Tel: (+86) 173-2652-1629 | Email: zengzy@shanghaitech.edu.cn | Homepage: https://adonis-galaxy.github.io/homepage

Summary

- Intern with **Prof. Jianbo Shi** at UPenn GRASP Lab; **Prof. Xuming He** at ShanghaiTech PLUS Lab. **Reviewer of CVPR 2022**
- Broadly interest in **Computer Vision**, mainly in **Strong Perception** and **Cross-Modality**, particularly **3D Perception** and **Vision-Language Model**

Educations

ShanghaiTech University   **Sept. 2019 - (Expected) June 2023**
B.Eng. in Computer Science, Major | Innovation and Entrepreneurship, Minor
  -International Summer Program, Osaka University, June 2022 | Entrepreneurial Mindshift Program, Babson College, March 2021

Publications (*: equal contribution)

- **Can Language Understand Depth?**
  Renrui Zhang*, **Ziyao Zeng***, Ziyu Guo, Yafeng Li | **Mentor: Renrui Zhang**, Ph.D. candidate at MMLab, CUHK
  ACM Multimedia 2022 (Brave New Idea) | Link: https://arxiv.org/abs/2207.01077
    -We are **the first to** conduct **zero-shot training-free adaptation** from semantic language to quantified vision tasks **(monocular depth estimation)**.
    -**Our DepthCLIP surpasses existing unsupervised methods** and even **approaches the early fully-supervised networks**.

- **DSPoint: Dual-scale Point Cloud Recognition with High-frequency Fusion**
  Renrui Zhang*, **Ziyao Zeng***, Ziyu Guo*, Xinben Gao, Kexue Fu | **Mentor: Prof. Jianbo Shi**, GRASP Lab, University of Pennsylvania
  arXiv technical report, 2021, in submission to ICRA 2023 | Link: https://arxiv.org/pdf/2111.10332.pdf
    -We proposed DSPoint to **conduct dual-scale processing** with local point convolution and global voxel attention for efficient point cloud processing.
    -For better feature blending, we conduct inter-scale cross-modality interaction by **communicating high-frequency coordinates information**.

- **VT-CLIP: Enhancing Vision-Language Models with Visual-guided Texts**
  Renrui Zhang*, Longtian Qiu*, Wei Zhang, **Ziyao Zeng** | **Mentor: Renrui Zhang**, Ph.D. candidate at MMLab, CUHK
  arXiv technical report, 2021 | Link: https://arxiv.org/pdf/2112.02399.pdf
    -To adapt CLIP to few-shot classification, we enhance vision-language modeling **via visual-guided texts**.
    -Specifically, we **guide text feature to adaptively explore informative regions on the image** and aggregate the visual feature by cross-attention.

- **iQuery: Instruments as Queries for Audio-Visual Sound Separation**
  In submission to a top-tier conference as a co-author | **Mentor: Prof. Jianbo Shi**, GRASP Lab, University of Pennsylvania
    -We re-formulate visual-sound separation task and **propose Instrument as Query (iQuery) with a flexible query expansion mechanism**.
    -We demonstrate state-of-the-art performance, with **up to 44.2 % improvement** of SDR on MUSIC benchmark.

- **PointCLIP V2: Adapting CLIP for Powerful 3D Open-world Learning**
  In submission to a top-tier conference as a co-author | **Mentor: Prof. Peng Gao**, Shanghai AI Laboratory
    -We introduce a **realistic shape projection** module, and leverage large-scale language models to **automatically design 3D-semantic prompt**.
    -Our approach significantly surpasses PointCLIP by **+42.90%, +40.44%, and +28.75% accuracy** on three datasets for zero-shot 3D classification.

Course Project

- **Twitter Emotion Classification** (CS181 Artificial Intelligence)
  Mentor: Prof. Kewei Tu, ShanghaiTech University |Link: https://adonis-galaxy.github.io/homepage/research/AI_CS181_project.pdf
- **Generalized DUQ: Generalized Deterministic Uncertainty Quantification** (CS282 Machine Learning)
  Mentor: Prof. Xuming He, PLUS Lab, ShanghaiTech University | Link: https://adonis-galaxy.github.io/homepage/research/ML_CS282_project.pdf
- **Seek Common while Shelving Differences: A New Way for dealing with Noisy Labels** (CS280 Deep Learning)
  Mentor: Prof. Xuming He, PLUS Lab, ShanghaiTech University | Link: https://adonis-galaxy.github.io/homepage/research/DL_CS280_project.pdf

Research / Working Experience

- **Shanghai AI Laboratory**   Oct. 2022 - Present
  -**Mentor: Prof. Bin Zhao** | Project: Visual Grounding
- **GRASP Lab, University of Pennsylvania**   April 2021 - Present
  -**Mentor: Prof. Jianbo Shi** | Project: DSPoint, Audio-Visual Sound Separation, Reconstruction-based Image Segmentation
- **Research with Renrui Zhang, Ph.D. student at MMLab, CUHK**   Nov. 2021 - July 2022
  -**Mentor: Renrui Zhang** | Project: DepthCLIP, VT-CLIP
- **PLUS Lab, ShanghaiTech University**   July 2020 - June 2021
  -**Mentor: Prof. Xuming He** | Project: Mix-data Anomaly Segmentation, Noisy Labels Classification, Uncertainty Estimation
- **UISEE AI Lab, UISEE Technology**   July 2021 - April 2022
  -Algorithm Engineer | Develop point cloud recognition & monocular 3D object detection algorithm for self-driving cars
- **Shanghai Advanced Research Institute, CAS**   Aug. 2020 - June 2021
  -Intern Lecturer | Teaching advanced scientific topic to public, including deep learning and computer vision

Service

Reviewer: CVPR 2022 (Emergency reviewer, reviewing two papers); ACM SIGMM member

Tutoring Experience

Freshman Mentor | Instructed by Prof. Rui Fan   2020 Fall – 2021 Spring

Awards / Competitions

Merit Student of ShanghaiTech University (Top 15%)   2021
State Grid AI Innovation Competition in Power Dispatch and Control (Top 10%, 24/242)   2021

Skills

Python, C/C++, PyTorch, Unity/C#, Bilingual (Chinese/English: CET6-607, GRE-327+4, TOEFL-103)