Ziyao Zeng (Adonis)

Last Update: Dec. 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**
- Research towards <u>Multimodal Perception</u> <u>inspired by human perception</u>, especially **3D Perception** and **Vision-Language Models**

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
  ACM Multimedia 2022 (Brave New Idea), acceptance rate≤12.5% | 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, Jianbo Shi
  arXiv technical report, 2021, submitted to ICRA 2023 | Link: https://arxiv.org/abs/2111.10332
    - 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**
  Longtian Qiu, Renrui Zhang, Ziyu Guo, **Ziyao Zeng**, Yafeng Li, Guangnan Zhang
  arXiv technical report, 2021, submitted to a top-tier conference | Link: https://arxiv.org/abs/2112.02399
    - 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.
- **PointCLIP V2: Adapting CLIP for Powerful 3D Open-world Learning**
  Xiangyang Zhu*, Renrui Zhang*, Bowei He, **Ziyao Zeng**, Shanghang Zhang, Peng Gao
  arXiv technical report, 2022, submitted to a top-tier conference | Link: https://arxiv.org/abs/2211.11682
    - 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.
- **iQuery: Instruments as Queries for Audio-Visual Sound Separation**
  Jiaben Chen, Renrui Zhang, Dongze Lian, Jiaqi Yang, **Ziyao Zeng**, Jianbo Shi
  arXiv technical report, 2022, submitted to a top-tier conference | Link: https://arxiv.org/abs/2212.03814
    - 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.

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, iQuery, 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, PointCLIP V2
- **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
  -Intern 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 / Tutoring Experience

CVPR 2022 Reviewer (Emergency reviewer, reviewing two papers); Volunteer of WWF China & Greenpeace
Freshman Mentor (Instructed by Prof. Rui Fan, 2020 Fall – 2021 Spring); ACM SIGMM member;

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)