Haoyi Duan

haoyid@sdanford.edu · · page · haoyi-duan · +1 (650) 460-9909

Education Background

Stanford University, California, USA 09/2023 - Present
M.S. Eletrical Engineering
Zhejiang University, Zhejiang, China 09/2019 - 06/2023
B.Eng. Computer Science and Technology, **Chu Kochen Honors College**
· GPA: **3.97**/4.00 (**91.38**/100)
· Major GPA: **3.99/4.00** (**93.17**/100)
National University of Singapore, Singapore 05/2022 - 07/2022
School of Computing Summer Workshop

Research Interests

3D Vision; Multimodal Learning.

Publications

WonderJourney: Going from Anywhere to Everywhere *CVPR'2024*
Hong-Xing (Koven) Yu, **Haoyi Duan**, Junhwa Hur, Michael Rubinstein, William T. Freeman, Forrester Cole, Deqing Sun, Noah Snavely, Jiajun Wu, Charles Herrmann

Cross-modal Prompts: Adapting Large Pre-trained Models for Audio-Visual Downstream Tasks *NeurIPS'2023*
Haoyi Duan*, Yan Xia*, Mingze Zhou, Li Tang, Jieming Zhu, Zhou Zhao
Outstanding Graduation Thesis, Zhejiang University

LaPE: Layer-adaptive Position Embedding for Vision Transformers with Independent Layer Normalization *ICCV'2023*
Runyi Yu*, Zhennan Wang*, Yinhuai Wang*, Kehan Li, Chang Liu, **Haoyi Duan**, Xiangyang Ji, Jie Chen

Timestamps as Prompts for Geography-Aware Location Recommendation *CIKM'2023*
Yan Luo, **Haoyi Duan**, Ye Liu, Chung Fu-Lai

Beyond Two-Tower Matching: Learning Sparse Retrievable Interaction Models for Recommendation *SIGIR'2023*
Liangcai Su, Fan Yan, Jieming Zhu, Xi Xiao, **Haoyi Duan**, Zhou Zhao, Zhenhua Dong, Ruiming Tang

Selected Research Experience

WonderJourney: Going from Anywhere to Everywhere 9/2023 - 01/2024
Advisor: Prof. Jiajun Wu, Stanford Vision & Learning Lab
· Design a modularized framework, which starts at any user-provided location and generate a journey through a long sequence of diverse yet coherently connected 3D scenes.
· Leverage an LLM to generate textual descriptions of the scenes in this journey, a text-driven point cloud generation pipeline to make a compelling and coherent sequence of 3D scenes, and a large VLM to verify the generated scenes.

Adapting Large Pre-trained Models for Audio-Visual Downstream Tasks 12/2022 - 08/2023
Advisor: Prof. Zhou Zhao, Zhejiang University DCD Lab
· Proposed a novel Dual-Guided Spatial-Channel-Temporal attention mechanism, which leverages audio and visual modalities as soft prompts to dynamically adjust the parameters of pre-trained models based on the current multi-modal input features.
· Achieved state-of-the-art results across multiple downstream tasks; exhibited promising performance in challenging few-shot and zero-shot scenarios.

Honors & Awards

National Scholarship (top 0.2%)	2022
First-Class Scholarship of Zhejiang University (top 3%)	2020, 2021, 2022
Zhejiang Provincial Outstanding Graduate	06/2023
Outstanding Graduate of Zhejiang University	06/2023

Languages & Skills

· Languages: English (proficient); Chinese (native). TOEFL: 108(S23); GRE: (V155+Q169+3.5).
· Programming: C/C++, Python, CUDA, mini SQL, Java, Verilog, x86, Shell.
· Tools/Software: Pytorch, OpenGL, MindSpore, MySQL, LaTeX, Vivado, Adobe PS/Pr.