ZHUOWAN LI

https://lizw14.github.io
410-900-2393 ⋄ zli110@jhu.edu

RESEARCH INTERESTS

My research interests lie in computer vision and natural language processing, including multimodal models, vision-and-language, compositional reasoning, model robustness, etc. I have experience in pretrained LLMs/VLMs. I am always excited to explore topics like generative AI, embodied AI and foundation models.

EDUCATION

Johns Hopkins University 2018 - present
Ph.D. in Computer Science
Advisors: Alan Yuille, Benjamin Van Durme

Tsinghua University 2014 - 2018
B.E. in Electronic Engineering
Second major: Journalism and Communication

RESEARCH EXPERIENCE

Amazon AWS, Santa Clara, CA May 2023 - present
Research intern. Mentors: Bhavan Jasani, Peng Tang, Shabnam Ghadar

· Compositional reasoning for document visual question answering using LLMs.

Facebook AI Research, Menlo Park, CA May 2021 - August 2021
Research intern. Mentors: Dhruv Mahajan

· Unsupervised domain adaptation with guidance of pretrained models.

Adobe Research, San Jose, CA May 2019 - Nov 2019
Research intern. Mentors: Quan Tran, Long Mai, Zhe Lin

· Propose new datasets and methods for a new task: context-aware group captioning.
· Published in CVPR 2020.

Sensetime, Beijing, China Oct 2017 - March 2018
Part-time research intern. Mentor: Shuai Yi

· Learn discriminative human representations by reconstructing multi-view images using GAN.
· Published in NeurIPS 2018.

Tsinghua University, Beijing, China Nov 2016 - May 2017
Research Assistant. Advisor: Prof. Shengjin Wang

· Improve person re-identification with fine-grained attributes and pose information.

PUBLICATIONS

ExoViP: Step-by-step Verification and Exploration with Exoskeleton Modules for Compositional Visual Reasoning
Yuxuan Wang, Alan Yuille, **Zhuowan Li***, Zilong Zheng*.
Under submission (2023)

Localization vs. Semantics: How Can Language Benefit Visual Representations Learning?
Zhuowan Li, Cihang Xie, Benjamin Van Durme, Alan Yuille.
Under submission (2023)

3D-Aware Visual Question Answering about Parts, Poses and Occlusions.
Xingrui Wang, Wufei Ma, **Zhuowan Li**, Adam Kortylewski, Alan Yuille.
NeurIPS 2023

Super-CLEVR: A Virtual Benchmark to Diagnose Domain Robustness in Visual Reasoning.
Zhuowan Li, Xingrui Wang, Elias Stengel-Eskin, Adam Kortylewski, Wufei Ma, Benjamin Van Durme, Alan Yuille.
CVPR 2023 ***Highlight***

Visual Commonsense in Pretrained Unimodal and Multimodal Models.
Chenyu Zhang, Benjamin Van Durme, **Zhuowan Li***, Elias Stengel-Eskin*.
NAACL 2022 ***Oral***

SwapMix: Diagnosing and Regularizing the Over-Reliance on Visual Context in Visual Question Answering.
Vipul Gupta, **Zhuowan Li**, Adam Kortylewski, Chenyu Zhang, Yingwei Li, Alan Yuille.
CVPR 2022

Calibrating Concepts and Operations: Towards Symbolic Reasoning on Real Images.
Zhuowan Li, Elias Stengel-Eskin, Yixiao Zhang, Cihang Xie, Quan Tran, Benjamin Van Durme, Alan Yuille.
ICCV 2021

Context-Aware Group Captioning via Self-Attention and Contrastive Features.
Zhuowan Li, Quan Tran, Long Mai, Zhe Lin, Alan Yuille.
CVPR 2020.

FD-GAN: Pose-guided Feature Distilling GAN for Robust Person Re-identification.
Yixiao Ge*, **Zhuowan Li***, Haiyu Zhao, Guojun Yin, Shuai Yi, Xiaogang Wang, Hongsheng Li.
NeurIPS 2018

PATENTS

Contrastive Captioning for Image Groups. US Patent App. 16/998,876, 2022.
Quan Tran, Long Mai, Zhe Lin, **Zhuowan Li**.

TALKS

Towards Generalizable Visual Reasoning
MIT: Computational Cognitive Science Group (Josh Tenenbaum's group) May 2023

MENTORING

Vipul Gupta	visiting intern, now PhD at Penn State Univeristy
Chenyu Heidi Zhang	JHU undergrad, now master at Stanford University
Chenyu Zhang	JHU master, now PhD at University of Trento
Xingrui Wang	USC master, now PhD at JHU
Varun Iyer	JHU master
Shitian Zhao	visiting undergrad from ECNU
Yijiang Li	JHU master student
Yuxuan Wang	visiting master student from Peking University

TEACHING

EN.601.461/661. Computer Vision — Johns Hopkins University
Role: Teaching Assistant — Spring 2022
Instructor: Kapil Katyal

SERVICE

Reviewer for CVPR, ICCV, ECCV, NeurIPS, ICML, ICLR.

SKILLS

Programming Languages	Python, MATLAB, C++, C, LaTeX
Deep Learning Tools	Pytorch, Tensorflow, Torch, Caffe