

Yi Li

PhD Candidate · UC San Diego

+1 (858) 281-3673 · [✉ yil898@ucsd.edu](mailto:yil898@ucsd.edu) · [🏠 svcl.ucsd.edu/people/yili](https://svcl.ucsd.edu/people/yili) · [📷 JerryYLi](#) · [🌐 yi-li-jerry](#)

Research Interests

- Video Understanding** Human action recognition · Spatiotemporal representation learning
- Multimodal Learning** Visual (images, videos) · Textual (natural language) · Audio (speech, music)
- Bias in ML/CV** Dataset bias · Algorithmic fairness · Bias mitigation

Education

University of California San Diego

La Jolla, CA, USA

PhD in Electrical & Computer Engineering

2017 - Now

- **Specialization:** Signal & image processing
- Joined Statistical Visual Computing Lab (SVCL) in 2017, conducting research in computer vision
- **Thesis:** Holistic Bias Mitigation of Datasets and Models in Computer Vision
- **Advisor:** Prof. Nuno Vasconcelos, **GPA:** 3.97/4.

The Chinese University of Hong Kong

Hong Kong SAR

BEng in Electronic Engineering (First Class Honors)

2013 - 2017

- **Specialization:** Digital signal processing (DSP) and multimedia
- Entered ELITE Stream of the Faculty of Engineering in 2015
- **Thesis:** Binary Image Denoising with BinURE-LET
- **Advisors:** Prof. Thierry Blu & Prof. Tan Lee, **Major GPA:** 4.00/4.

Experience

Intel Labs

San Diego, CA

Research Intern

Jun. - Nov. 2022

- **Collaborators:** Subarna Tripathi, Kyle Min
- Learning long-term spatiotemporal representations from video-language pretraining

MIT-IBM Watson AI Lab

Cambridge, MA

Research Intern

Jun. - Nov. 2021

- **Collaborators:** Rameswar Panda, Yoon Kim, Chun-Fu Chen, Rogerio Feris, David Cox
- Improving text-only machine translation by visual hallucination learned on image-text corpora

Statistical Visual Computing Lab, UC San Diego

La Jolla, CA

Graduate Student Researcher

Jan. 2018 - Now

- Quantitative measurement and systematic mitigation of data bias
- Principled approach for collecting large-scale, bias-controlled datasets
- Model debiasing for improved performance, transferability and robustness in video action recognition

Publications

SViTT: Temporal Learning of Sparse Video-Text Transformers

CVPR 2023

Yi Li, Kyle Min, Subarna Tripathi, Nuno Vasconcelos

Vancouver, Canada

- Proposed SViTT, a sparse video-text architecture with hybrid edge and node sparsity
- SViTT is trained with a temporal expansion curriculum with increasing clip length and model sparsity
- SViTT outperforms dense transformers on video retrieval and question answering, with reduced compute cost

VALHALLA: Visual Hallucination for Machine Translation

CVPR 2022

Yi Li, Rameswar Panda, Yoon Kim, Chun-Fu Chen, Rogerio Feris, David Cox, Nuno Vasconcelos

New Orleans, LA

- Proposed VALHALLA framework for improving natural language modeling from multimodal data
- Trained on image-text corpora, VALHALLA learns to hallucinate visual features to assist text-only inference
- Improved machine translation results on text-only & multimodal datasets over state-of-the-art

Improving Video Model Transfer with Dynamic Representation Learning

CVPR 2022

Yi Li, Nuno Vasconcelos

New Orleans, LA

- Proposed DRL, a framework for learning video representations that capture temporal dynamics
- By maximizing their dynamic score, DRL prevents networks from overfitting to spatial bias in video data
- Improved transferability and generalization of video models over conventional training

BEV-Net: Assessing Social Distancing Compliance by Joint People Localization and Geometric Reasoning

ICCV 2021

Zhirui Dai, Yuepeng Jiang, Yi Li, Bo Liu, Antoni B. Chan, Nuno Vasconcelos

Virtual conf.

- Proposed BEV-Net, a framework for localizing people in world frame from monocular cameras
- Developed a differentiable homography module to map image into BEV coordinates for end-to-end training
- Released a new benchmark on social distancing compliance assessment

Learning Representations from Audio-Visual Spatial Alignment

NeurIPS 2020

Yi Li^{*}, Pedro Morgado^{*}, Nuno Vasconcelos

Virtual conf.

- Proposed audio-visual spatial alignment (AVSA) task for self-supervised learning from 360° videos
- AVSA uses contrastive learning to match audio and visual signals from the same viewpoints
- Developed a transformer-based architecture for AVSA which models context across viewpoints
- Showed that training with AVSA improves representation for various downstream tasks

Background Data Resampling for Outlier-Aware Classification

CVPR 2020

Yi Li, Nuno Vasconcelos

Virtual conf.

- Proposed a resampling approach for out-of-distribution (OOD) detection with background data
- Formulated background dataset selection as adversarial reweighting problem
- Achieved superior OOD detection with resampled background data vs. uniform sampling

REPAIR: Removing Representation Bias by Dataset Resampling

CVPR 2019

Yi Li, Nuno Vasconcelos

Long Beach, CA

- Developed a differentiable data resampling approach (REPAIR) to reduce representation bias in datasets
- Applied the proposed method on action recognition datasets to remove static bias from videos
- Evaluated sensitivity of video recognition algorithms to the presence of bias in data
- Showed removing representation bias improves model generalization

RESOUND: Towards Action Recognition without Representation Bias

ECCV 2018

Yingwei Li, [Yi Li](#), Nuno Vasconcelos

Munich, Germany

- First study of representation bias in video action recognition datasets
- Quantitatively measured the object/scene/human bias of existing action recognition datasets
- Collected Diving48, a dataset of competitive diving for bias-free action recognition

Academic Service

Peer reviewer TPAMI, CVPR (**outstanding reviewer '23**), ICCV, ECCV, NeurIPS, ICLR

Student volunteer CVPR '19/20 AC meeting

Honors & Awards

2017 **Graduate Fellowship**, Dept. of Electrical & Computer Engineering, UCSD

Winbridge Scholarship, CW Chu College, CUHK

2016 **HKSAR Government Scholarship**, The Government of Hong Kong

EE Alumni Association Scholarship, Dept. of Electronic Engineering, CUHK

Teaching Experience

Teaching Assistant ECE 153 Probability & Random Processes for Engineers

ECE 271A Statistical Learning I

ECE 271B Statistical Learning II

ECE 271C Deep Learning and Applications

Skills

Programming Languages Python, MATLAB, C/C++, Java

Deep Learning Libraries PyTorch, JAX/Flax, TensorFlow, Caffe

Miscellaneous \LaTeX , HTML/CSS, Bash

Languages English, Chinese