

# Shijia Yang

[shijiayang@berkeley.edu](mailto:shijiayang@berkeley.edu) • (626) 861-1312 • 6401 Shellmound St Apt 7208, Emeryville, CA

---

## EDUCATION

---

### University of California, Berkeley

September 2019 - May 2023

- Bachelor of Art in Computer Science
- GPA: 3.9/4.0
- Relevant Coursework:
  - Applied Theory - 3D Vision (Graduate) [A+], Machine Learning, DNN, Computer Graphics
  - Theory - Optimization, Linear Algebra, Probability, Discrete Mathematics, Multivariate Calculus

---

## RESEARCH EXPERIENCE

---

### Mechanical Systems Control Lab

March 2021 - Present

*Undergraduate Researcher, advised by Prof. Masayoshi Tomizuka*

Berkeley, CA

- Currently researching on multi-modal multi-task learning for autonomous vehicles
- Investigating implicit and explicit neural representation with long-term temporal images
- Focusing on improved 3D detection by enforcing model expandability and self-consistency

### Berkeley Artificial Intelligence Lab

March 2021 - Present

*Undergraduate Researcher, advised by Prof. Kurt Keutzer & Prof. Joseph E. Gonzalez*

Berkeley, CA

- Currently researching on data and training efficiency for 3D vision
- Conducting semi-supervised and self-supervised experiments for point-cloud classification
- Exploring efficient prompt tuning of language and vision models

---

## PUBLICATIONS

---

### [1] Image2Point: 3D Point-Cloud Understanding with 2D Image Pretrained Models

*Shijia Yang\**, Chenfeng Xu\*, Tomer Galanti, Bichen Wu, Xiangyu Yue, Bohan Zhai, Wei Zhan, Peter Vajda, Kurt Keutzer, Masayoshi Tomizuka

European Conference on Computer Vision (ECCV) 2022

- Proposed a simple pretraining scheme of using 2D model weights for 3D point-cloud understanding
- Brought consistent improvement on various baselines, including ViT and PointNet++, and 10.0% improvement for few-shot classification on the ModelNet40 dataset
- Explained feasibility of image-point-cloud transfer from the aspect of neural collapse

### [2] Time Will Tell: New Outlooks and a Baseline for Temporal Multi-View 3D Object Detection

*Jinhyung Park\**, Chenfeng Xu\*, *Shijia Yang*, Kurt Keutzer, Kris Kitani, Masayoshi Tomizuka, Wei Zhan

International Conference on Learning Representations (ICLR) 2023 Top-5%

- Formulated temporal camera-only 3D detection as multi-view stereo matching problem and observed problems of previous methods regarding the low matching resolution and limited history usage
- Proposed SOLOFusion which generate a cost volume with a long history of image observations with coarse but efficient matching resolution, then augment with short-term, fine-grained matching resolution
- SOLOFusion sets new state-of-the-art on nuScenes, achieving first place on the test set and outperforming previous best art by 5.2% mAP and 3.7% NDS on the validation set

### [3] Multitask Vision-Language Prompt Tuning

Shijia Yang\*, Sheng Shen\*, Tianjun Zhang\*, Bohan Zhai, Joseph E. Gonzalez, Kurt Keutzer, Trevor Darrell

Arxiv.

- Proposed the multitask vision-language prompt tuning (MVLPT) framework, including multitask prompt initialization and multitask prompt adaptation, and demonstrate the efficacy for each component
- Study the task transferability across 20 vision tasks with 400 combinations for each prompt tuning method to understand when MVLPT is most effective
- Evaluate MVLPT on the few-shot ELEVATER and cross-task generalization benchmarks, which sets the new state-of-the-art on 20-shot ELEVATER benchmark

---

## INDUSTRY EXPERIENCE

---

### Microsoft - Bing Turing Team

Research Intern

January 2023 - Present

Bellevue, Washington

- Work on efficient training and inference of the foundation generative models, such as DALL-E and Stable Diffusion
- Proposed a method to combine prompt tuning, a parameter efficient training method, with text-to-image models
- Take advantage of large language models like ChatGPT to generate input prompt to T2I models

---

## PERSONAL PROJECTS

---

### Multi-modal Multi-task Learning for 3D Perceptive via Closed-loop Representation *advised by Prof. Yi Ma*

- Proposed a paradigm, inspired by closed-loop feedback from control theory, for a multi-modal multi-task 3D model
- Performed five common tasks for point-cloud perception, including foreground detection and object detection
- Achieved at least 4% mAP improvement for object detection on Argoverse dataset compared with UNet3D baseline

### Snow Simulation *advised by Prof. Ren Ng*

- Used Taichi language designed for physical simulation and implemented the moving-least-square algorithm (MPM) to simulate realistic snow particles in 2D and 3D
- Studied snow particles' update rule for positions, affine velocity field, deformation gradient, etc.
- Implemented snowballs, frictional surface, and rigid body by adjusting particle distribution and parameters

### Data Free Bert Quantization

- Alleviating the problem that quantized models may not have original data to perform post-quantization finetuning
- Proposed token generation algorithm to generate pseudo data for low precision Bert model quantization
- Improved accuracy of quantized Bert by 12.2%-61.7% on GLUE tasks compared to quantized Bert baseline

### Quant Trading Platform

- Explored Interactive Brokers API in Java to access market and user account data
- Connected API to enable trading of selected stocks or market products and visualize market data
- Designed and wrote GUI using JavaFX for the platform, allowing users to manage accounts on different tabs

---

## LANGUAGES / SKILLS

---

**Tools:** NumPy, PyTorch, MMDetection3D, Pandas, Matplotlib, Scikit-Learn, SciPy, OpenCV, TensorFlow

**Computer Languages:** Python, Java, C++, SQL, HTML, C#, LaTeX, CUDA

**Languages:** English, Mandarin