

Shijia Yang

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), 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

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 under review

- 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

INDUSTRY EXPERIENCE

Shanghai Baosight Software

ML Infrastructure Intern

December 2020 - February 2021

Shanghai, China

- Worked on building machine learning infrastructure from scratch to scale the service to the company level
- Created high-availability Kubernetes clusters and simplified the cluster management process by installing Rancher
- Deployed JupyterHub and Spark to provide individual coding environments with data processing tools

PERSONAL PROJECTS

Multi-modal Multi-task Learning for 3D Perceptive via Closed-loop Representation

- 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

- 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

Reinforcement Learning Gaming Strategy

- Simulated dice game with given rules and implemented Tabular Q-Learning algorithm
- Fully optimized C++ code to reduce training time up to 10^{10} epoch every 3 hours
- Ran experiments with self-designed reward function and reached the top win rate among all RL algorithms

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