SHUFAN LI

jacklishufan@berkeley.edu • https://github.com/jacklishufan • http://homepage.jackli.org

EDUCATION

University of California, Berkeley, Berkeley, CA Expected May 2023
Bachelor of Arts in Computer Science and Applied Mathematics GPA: 3.94/4.0
Honors: Edward Frank Kraft Award, MIT COVID Challenge 2020 hackathon winner, Dean's List, EECS Honor Program
Coursework: Linear Algebra, Data Structure, Algorithm, Computer Architecture, Deep Learning, Discrete Math, Probability, Convex Optimization, Analysis, Graduate course in algebra topology.

RESEARCH EXPERIENCE

Research Intern Berkeley Artificial Intelligence Research (BAIR) Spring 2022 – Present

- Worked on self-supervised learning in Vision and Learning Group led by Professor Trevor Darrell.
- Designed and implemented a self-supervised pretraining method that achieved state-of-the-art performance on various downstream tasks. Co-authored a paper.
- Explored ways of improving MAE pretraining on satellite images by conditioning on spatial resolution.
- Worked with industry partners to utilize multiscale satellite data for landcover segmentation.

Machine Learning Engineer Intern, Orka (hiorka.com) June 2021– Present

- Implemented a fall detection Convolutional Neural Network in Torch and achieved a 98% accuracy. The model was integrated into the company's next-generation hearing aid.
- Designed a perspective-rectification network based on Mask-RCNN. Improved the average precision of a line chart recognition model by 7%.
- Designed a line chart data extraction system that can be applied to real photos without any fine-tuning while trained only on synthetic data.

Research Intern, Environmental Systems Dynamics Laboratory - Berkeley Sep 2021 – June 2022

- Improved the consistency of an LSTM network that predicts U.S water streams by introducing physical constraints to the loss function and model architecture.
- Parallelized the legacy training pipeline to allow simultaneous hyper-parameter tuning on multiple nodes of Berkley Savio Computing Cluster.
- Extended an existing benchmark on HJ Andres reservoir by incorporating more model architectures and source locations.

Research Intern, Berkeley SETI Research Center Feb 2020 – June 2020

- Explored various approaches of detecting potential alien signals from Greenbank Telescope observation data using CNN.
- Maintained and updated a public dataset of radio telescopes observations.

ORIGINAL WORKS (http://homepage.jackli.org/)

Under Review

Co-first Author: Refine and Represent: Region-to-Object Representation Learning. (2022)
Akash Gokul, Konstantinos Kallidromitis*, Shufan Li*, Yusuke Kato, Kazuki Kozuka, Trevor Darrell, Colorado J Reed*

Co-first Author: Scale-MAE: A Scale-Aware Masked Autoencoder for Multiscale Geospatial Representation Learning. (2022)
Colorado Reed, Shufan Li*, Ritwik Gupta, * Sarah Brockman, Christopher Funk, Brian Clipp, Kurt Keutzer, Salvatore Candido, Matt Uyttendaele, Trevor Darrell*

First Author: Chart-RCNN: Efficient Line Chart Data Extraction from Camera Images (2022)
Shufan Li, Congxi Lu, Linkai Li, Haoshuai Zhou*

Technical Report

First Author: Interpreting Audiograms with Multi-stage Neural Networks (2021)
Shufan Li, Congxi Lu, Linkai Li, Jirong Duan, Xinping Fu, Haoshuai Zhou*

PERSONAL PROJECTS

Anomaly Detection with Verizon

- **Co-led** a client project of Student Association for Applied Statistics (SAAS) at Berkely which aimed to develop a classier that predicts network outage for Verizon.
- Implemented a variety of univariate and multivariate time series models from ARIMA to LSTM, and benchmarked their performance for the final report.

Deep Image Rectification with Woflow

- **Co-led** a client project of SAAS at Berkely which aimed to develop a deep network that can automatically rectify restaurant manual photos from customers of Woflow.
- Reimplemented a state-of-the-art solution based on literature and benchmarked its performance on proprietary dataset.

Echo Bot (re-treat.app)

- **Led** a team of three engineers in a campus-based startup that delivers an online mental health chatbot.
- Coordinated the implementation of a web-based chatbot design in Django and React.
- Implemented an NLP classifier that detects conversation topics and user sentiments utilizing the state-of-the-art distilled BERT model. Achieved an accuracy of 93% on public data collected from Reddit and Twitter.