

# Ariel N. Lee

@ ariellee@bu.edu |  LinkedIn |  GitHub |  Portfolio |  Boston, MA

## EDUCATION

---

### M.Sc., Boston University (BU)

Boston, MA

*Electrical & Computer Engineering, Subconcentration: Data Analytics; GPA: 3.71*

Sep 2020 – Jun 2023

**Activities:** Out in STEM; Graduate Women in Science & Engineering

**Relevant Coursework:** Machine Learning, Deep Learning, Artificial Intelligence, Computer Vision, Advanced Algorithms, Optimization Theory & Methods, Optimization for ML

### B.Sc., University of California, Los Angeles (UCLA)

Los Angeles, CA

*Microbiology, Immunology, & Molecular Genetics (MIMG); GPA: 3.45*

Sep 2011 – Jun 2015

## PUBLICATIONS & COMPETITIONS

---

[Paper, Arxiv 2023] - Under Submission

Ariel N. Lee, Sarah Adel Bargal, Janavi Kasera, Stan Sclaroff, Kate Saenko, Nataniel Ruiz. "Hardwiring ViT Patch Selectivity into CNNs using Patch Mixing" *arXiv preprint arXiv:2306.17848 (2023)*

[Competition, META AI 2023]

Meta AI Video Similarity Challenge: 8/196 overall, 1/42 in AI grad course

## GRADUATE & POST-GRADUATE RESEARCH EXPERIENCE

---

### Platypus-30B

Boston, MA

*Independent Researcher, Large Language Models*

Jun 2023

- Curated a dataset of 110,000+ question-answer pairs focused on STEM and logic for instruction fine-tuning Meta's LLaMa model using Low-Rank Adaptation on 4 A100 80GB GPUs.
- HuggingFace LLM Leaderboard evaluation pending, with preliminary benchmark tests showing significant improvement over competitive open-source models.

### Boston University

Boston, MA

*Graduate Researcher, Computer Vision*

Oct 2022 – May 2023

- Worked with **Dr. Nataniel Ruiz**, **Prof. Sarah Adel Bargal**, and **Prof. Kate Saenko** to research patch selectivity in modern convnets and ViTs. Released datasets for public use.
- Introduced new c-RISE explainability method to show that, by training CNNs with Patch Mixing, we simulate the natural ability of ViTs to ignore out-of-context information. NeurIPS submission pending.

### TeachForward & BU Wheelock Educational Policy Center

Boston, MA

*Data & Process Engineer, MLOps Dev Team*

Sep 2022 – Dec 2022

- Developed a feature extraction pipeline to analyze the use of teaching time based on 10,000+ videos of classroom observations.
- Created a simple user interface for client using gradio and Hugging Face spaces. User uploads a video and the pipeline returns mp4 files with object and activity detection annotations, among others.

### Boston University, AI4ALL

Boston, MA

*Research Intern, Computer Vision*

May 2022 – Aug 2022

- Worked with **Dr. Nataniel Ruiz** and **Prof. Sarah Adel Bargal** on counterfactual simulation and testing of modern convnets and ViTs.

## WORK EXPERIENCE

---

### Boston University

Boston, MA

*Deep Learning Course Grader*

Jul 2022 – May 2023

- Handled grading and answered student questions for the Deep Learning graduate course with **Prof. Sarah Adel Bargal** and **Prof. Brian Kulis** for multiple semesters.

## AI4ALL @ BU

Program Coordinator

Boston, MA

May 2022 – Aug 2022

- Organized and instructed a summer program for high school students dedicated to inclusion in AI.

## eMinutes

Corporate Paralegal (Remote)

Manager of Entity Management

Corporate Paralegal

Los Angeles, CA — Boston, MA

Aug 2019 – Mar 2021

Oct 2018 – May 2019

Apr 2017 – Oct 2018

- Evaluated technology deficits on main website, where all document production and communication is handled through a web-based system.
- Managed and trained two law school clerks throughout their clerkship, and maintained corporate governance for 25,000+ entities in 50 states.

## Law Offices of Sanford Jossen

Paralegal

Legal Assistant

Los Angeles, CA

Oct 2016 – Apr 2017

Oct 2015 – Oct 2016

- Researched and drafted legal documents, and summarized complex medical records.

## PROJECTS

---

### Ensemble Effect: Leveraging Fine-tuned Models for Prompt Prediction | [GitHub](#)

- Ensemble-based approach for predicting text prompts used to generate Stable Diffusion images.
- Surpassed the performance of traditional image captioning models by employing fine-tuned OpenAI CLIP and ViT models, along with the CLIP Interrogator (BLIP+CLIP), using a custom dataset of 105,000 image-prompt pairs

### Visual Odometry: Mapping Out the Camera Path | [GitHub](#)

- 3rd place in CS 585 Computer Vision class challenge, focused on estimating the camera path by recovering relative motion between successive frames.
- Implemented Random Sample Consensus (RANSAC) and linear triangulation from scratch for fundamental matrix and camera pose estimation, respectively.

### Crypto of the Future: Reinforcement Learning | [GitHub](#)

- DL reinforcement algorithm — proximal policy optimization — to devise an automatically generating strategy for Ethereum transactions.
- LSTM is used to make predictions for next day closing prices, which are in turn used to construct the automatic policy.

## UNDERGRADUATE RESEARCH EXPERIENCE

---

### UCLA Department of MIMG

Undergraduate Researcher, Characterization of Novel Bacteriophages

Los Angeles, CA

Sep 2014 – Jun 2015

- Worked with **Dr. Giorgia Pirino** to advance phage therapy research in the SEA-PHAGES project by isolating a novel bacteriophage: PH8s.
- Probed potential gene functions via electron microscopy and plaque assays, leading to a fully annotated genome added to the [NCBI GenBank database](#) and presented at the MIMG symposium.

### UCLA Department of Psychology

Undergraduate Researcher, Directed Research in Medicine

Los Angeles, CA

Jun 2014 – Aug 2015

- Conducted research with **Dr. Thomas Minor** for senior project by using learned helplessness to model symptoms of Post-Traumatic Stress Disorder.

## SKILLS

---

**Programming & Technologies:** Python (PyTorch, TensorFlow, NumPy, Pandas, scikit-learn), Java, MATLAB, OpenCV, GCP, Lambda Cloud, Git/GitHub, Hugging Face Hub (spaces, datasets, models), Docker, LaTeX

**ML Techniques:** LLM instruction fine-tuning, ViT training and fine-tuning, CNN training, novel data augmentation techniques, ML pipeline deployment, open-source models and datasets