Yu-Teng Li (Kevin)

✉ yutengli@berkeley.edu   in linkedin.com/in/yutengli   📞 415-418-0756   ⌧ thekevinli.github.io

EDUCATION

University of California, Berkeley   **Aug 2019 – May 2023**
B.S. Electrical Engineering & Computer Science   (GPA: 3.7 / 4)

EXPERIENCE

Applied Research Scientist, Firefly, *Adobe Inc.,* San Jose, CA.   **Aug 2023 – Present**

- Lead modeler for multimodal pretraining of **Firefly Image 5**, including workflows from text-to-image, instruction editing, character reference to layer generation. Handled distributed training over 1000+ GPUs with billion-scale data samples daily.
- Core member of foundation model research team for **Firefly Image 4** – developed the first effective recipe of aesthetics finetuning (SFT) that demoed at Adobe MAX and proposed a curriculum learning approach to counter GenAI smoothness artifacts during training.
- Proposed **UniFusion**, a novel architecture to fuse Diffusion Transformer with VLM as a universal encoder for text and image, which powers multiple OOD generalization behaviors such as multi-image reference generation with a model trained only with single-image reference.
- Led personalization for **Firefly Image 3** – improved Dreambooth recipe's stability with VLM-predicted concept and super-class, improved memory efficiency with AdaFactor, enabled ControlNet and Stylization, leading to multi-million dollar revenues on enterprise customers.

Undergraduate Student Instructor (Head of Discussion, CS182), *UC Berkeley*.   **Aug 2022 – May 2023**

- Leading the curriculum design of discussion sections in a course with 300+ students, CS182, Deep Learning, in Spring '23.
- Taught 20+ students in weekly discussions as a group tutor and designed homework material on **Transformers** in Fall '22.

Software Engineering Intern, *Adobe Inc.,* San Jose, CA.   **May 2022 – Aug 2022**

- Develop **Smart Grid for Face Stylization** – a project that enables users to stylize portraits by interpolating the latent space of **Generative Adversarial Networks** for any style combinations in a real-time image-editing app (*10 fps*).
- Implemented a full-stack web app from scratch, including the frontend UI, backend APIs and ML model deployment of GANs, using *TypeScript*, *Webpack*, *Node.js*, *Express*, *Flask, AWS* and *PyTorch*.
- Propose a **layout algorithm** to cluster and present 10k styles generated from GAN within a 16x16 grid (condensed **97.5% of all styles**) based on style encoding affinity scores, by experimenting with UMAP, K-means clustering and linear programming algorithm.
- Designed and trained a **CycleGAN**-based Face Translation between attributes to **mitigate warping artifacts** in Photoshop **Liquify Filter**.

Student Researcher, *Berkeley Artificial Intelligence Research,* Berkeley, CA.   **Feb 2022 – Oct 2023**

- Led a research project in **Active Learning on Image Segmentation** for autonomous driving, in Trevor Darrell's group. [lab website]
- Developed an **experiment baseline system** from scratch, which supports image-, pixel- and region-based active learning methods with multiple sampling heuristics. The system turned **days** of coding to set up each new experiment, into just **minutes** by modifying a config file.

Machine Learning Intern, *Dell Technologies Inc.,* Taiwan.   **Jul 2021 – Aug 2021**

- Developed 4 models to forecast Wi-Fi systems' throughput with Gradient Boosting Decision Tree, using 1.4M entries of noisy lab test data.
- My project attained **88% testing accuracy** in predicting Wi-Fi systems' throughput on **54 Dell laptops and workstations**, shortening the development of a Wi-Fi system **from 40 to 33 weeks** by providing instant forecasts instead of having to wait for prototyping and lab testing.

Student Researcher, *Vision and Learning Lab @ National Taiwan University*.   **Jan 2021 – Oct 2021**

- Implemented a research project to tackle Domain Generalization by synthesizing inter-domain styles by training a GAN in episodic learning.

AWARDS & PUBLICATIONS

UniFusion: VLM as Universal Encoder in Image Generation [Tech report coming soon] – First author for UniFusion, a novel architecture that leverages a single, frozen Vision-Language Model as a universal encoder for all input modalities. We proposed Layerwise Attention Pooling module to aggregate intermediate features in VLMs and demonstrated superior image generation that beats Flux.1 [dev], and image editing, character reference that rival Qwen-Image.

Hyperbolic Active Learning for Semantic Segmentation under Domain Shift [ArXiv] – ICML 2024. HALO usings hyperbolic neural network for pixel-level active learning in semantic segmentation, and is the first active learning approach that surpasses the performance of supervised domain adaptation with merely 1% small portion of labels.

Neighboring state-based RL Exploration [ArXiv] – First author for a research project on a model-agnostic, on-policy exploration method in deep reinforcement learning. Our method consistently outperforms Double DQN baseline by 49% in test time. Built in *PyTorch*, *MuJoCo*.

Winner of IoT Security Challenge & Best Encryption Solution, *Cal Hacks* [Devpost]   **Oct 22 - 24 2021**

- Built **CypherAI** – a Fully Homomorphic Encryption solution for ML systems on IoT devices to run CNNs over user data for face recognition.

Winner of Best Machine Learning Hack & Microsoft Hack for Social Good, *HackMIT* [GitHub]   **Sep 18 - 19 2021**

- Developed **ClassCaster** – a ML-powered course workload forecaster and study group recommendations using K-means clustering.