

# Yu-Teng Li ( Kevin )

 [yutengli@berkeley.edu](mailto:yutengli@berkeley.edu)  [linkedin.com/in/yutengli](https://linkedin.com/in/yutengli)  415-418-0756  [thekevinli.github.io](https://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

- Co-lead of **Firefly Image 5** multimodal pretraining, which supports 2k resolution text-to-image, instruction editing, character reference and **layer generation**. Planned all data ablation studies, handling distributed training over 1000+ GPUs with billion-scale data samples daily.
- Proposed **UniFusion**, a novel architecture for Diffusion Transformers to leverage VLMs as unified semantic encoders, which achieves generation/editing quality that rivals Flux.1[dev] and Bagel, as well as emergent behaviors e.g. **zero-shot multi-image reference generation**.
- Core member of **Firefly Image 4** foundation model team - 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.
- 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.
- Proposed **Inversion-based style transfer with weighted cross-attention**, a lightweight method to enables high-quality stylization data curation internally for instruction editing model training.

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

- Led research projects on **Active Learning on Image Segmentation** for autonomous driving, in Trevor Darrell's group. [ [lab website](#) ]
- Developed a config-driven experiment package which supports image-, pixel-, and region-based active learning (AL) with multiple sampling heuristics (e.g. entropy, mask autoencoding), eventually leading to the creation of **HALO**, the first AL method to surpass fully supervised baseline in segmentation benchmarks (ICML 2024).

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

- Led the discussion curriculum design for UC Berkeley's Deep Learning course (300+ graduate & undergraduate students, Spring 2023).
- Taught 30+ students in weekly discussions and designed homework material on topics including **Denoising Diffusion** and **Transformers**.

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

- Designed and trained a **CycleGAN**-based Face Translation between expressions to **mitigate warping artifacts** in Photoshop Liquify Filter.
- Develop **Smart Grid for Face Stylization** – a real-time image editing web app (**10 fps**) that enables users to stylize portraits by interpolating the latent space of **Generative Adversarial Networks** for any artistic style combinations.
- Propose an auto **layout algorithm** to efficiently present 10,000 styles generated from StylesGAN2 within a 16x16 grid (condensed **97.5% of all styles**) based on learnable style affinity scores, clustered with UMAP, K-means and linear programming.

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

- Developed models to forecast Wi-Fi systems' throughput with ensembled ResNet and Gradient Boosting Decision Tree. My models had 90% accuracy in predicting Wi-Fi throughput on **54 Dell laptop products**, shortening the development of Wi-Fi systems **from 40 to 33 weeks**.

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

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

## AWARDS & PUBLICATIONS

---

**UniFusion: VLM as Unified Encoder in Image Generation** [ [ArXiv](#) ] [ [Project Website](#) ] - The first architecture to use only a Vision-Language Model as condition encoder, without auxiliary signals from VAE to do image editing with quality rivaling Bagel and Flux. UniFusion enables emergent capabilities such as zero-shot multi-reference generation, and capability transfer where training on Editing helps T2I quantitatively.

**Hyperbolic Active Learning for Semantic Segmentation under Domain Shift** [ [ArXiv](#) ] - ICML 2024. HALO uses 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 on synthetic-to-real benchmarks such as GTAV → Cityscapes.

**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

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