Yifeng Huang

me@fyhuang.com
github.com/fyhuang

AI/ML Experience

Img2loc
Camera-only localization anywhere on Earth (no GPS) with deep learning.

- **In-browser inference demo!**
  - Source code: github.com/fyhuang/img2loc
- **Trained multi-label classifier over S2 cells.**
  - Multi-label classification head on top of fine-tuned Vision Transformer.
  - Positive weighted BCE loss to correct class imbalance.
  - Custom validation metrics including estimated Geoguessr score.
  - Repeatable training setup with cloud GPUs using Ansible, Conda.
- **Created 3 new image-geolocation datasets.**
  - Density-weighted sampling from OSM for Street View dataset.
  - Image filtering & retrieval from Flickr for large-scale geotagged image dataset.
  - Adaptive S2 cell splitting algorithm for selecting output labels.
- **Built web-based inference demo** using ONNX Web Runtime.
- PyTorch, Lightning, Pandas, TensorBoard, ONNX, OSM,

Personalized GenAI Image
Ongoing project to train generative AI model for creating localized/personalized graphics.

- Created input dataset for fine-tuning style transfer from existing assets.
- Model selection & fine-tuning.
- Targeted experiments and tuning for productionisation blockers.
- Batch scripts for large-scale generation of images.

Experience

YouTube Trust & Safety 2021-present
Engineering Manager, Staff SWE

- **Managed team of 8 engineers to launch 4 major projects in 2 years.**
  - New feedback product built for GenAI use cases, serving both high-volume user facing traffic and internal T&S use cases.
  - New user-facing appeals product built on DSA deadline.
  - Revamped flagging form built on modern YouTube x-plat stack.
  - Flag email notifications with low-latency online scheduling and batching.

- **Turned around low frontend productivity with codebase modernization project.**
    - Modernized the reporting form on a centralized, cross-platform tech stack.
    - Improved turnaround time from 2-3 quarters for minor changes, to launching new features and surfaces with just weeks of coding.

- **Increased team's ownership and scope across the org to encompass complete workflow.**
    - Added visibility on full flag lifecycle with real-time event listener service.
    - Asynchronous, low-latency architecture handles highly variable fanout up to 6 orders of magnitude.
    - Expanded collaboration with ML teams to increase coverage of ML flag filtering.
    - Designed flag signals service and started initiative to retrain models with new signals.

- **Responsible for full end-to-end stack.**
    - Backend C++ infrastructure processes 150M user flags per month.
    - Integration with abuse review systems and ML filtering models.
    - Frontend in TypeScript, Elements, Wiz Next.

2017-2020

rideOS – autonomous vehicle mapping & routing startup
Founding Engineer, Tech Lead

- **Tech lead for a team of 4 engineers.**
    - Launched company's first customer-facing product.

- **Worked across the stack on routing & dispatch projects.**
    - A* routing engine with no-downtime map updates from OSM.
    - Mapbox GL visualizations with pure pursuit vehicle simulation.
    - Ridehail API for autonomous vehicle dispatch with state management in Spanner.
    - Java microservices stack on Kubernetes + Helm + gRPC.
    - On-vehicle (ROS) and backend integrations with autonomous vehicle partners.

Other Work

- **Video stack lead at Google/Nest (2013-2017).** Led team of 3. C++ asynchronous event-driven firmware platform for 4k video encoding & live streaming. Cloud audio + video transcoding with Cgo and ffmpeg.
- **Computer Vision/Camera Consultant.** Real-time computer vision projects with C++ and CUDA.
- **Archive Box.** Personal file organizer. Python 3. Offline-first design using CRDTs.
- **World Budget.** Multi-currency budgeting webapp. Django + React + Postgres.

Education

Stanford University, B.S. Symbolic Systems. 2013