Yifeng Huang

me@fyhuang.com
github.com/fyhuang

AI/ML Experience

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

- **Built in-browser inference demo using ONNX Web Runtime.**
  - 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.
- PyTorch, Lightning, Pandas, TensorBoard, ONNX, OSM, Ansible.

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.
- **Improved frontend productivity ~10x 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.

rideOS – autonomous vehicle mapping & routing startup 2017-2020
Founding Engineer, Tech Lead

- **Developed & launched company's first customer-facing SaaS product.**
    - Led a team of 4 to develop ride hailing API for autonomous vehicle dispatch.

- **Worked across the stack on routing & dispatch projects.**
    - A* routing engine with no-downtime map updates from OSM.
    - Java microservices stack on Kubernetes + Helm + gRPC.
    - Built on GCP tech stack: Spanner, GKE, Bigtable, BigQuery.
    - On-vehicle (ROS) and backend integrations with autonomous vehicle partners.

Google (Nest) 2013-2017
Senior Software Engineer

- **Nest Cam video stack tech lead.**
    - Led team of 3 on C++ firmware (Android/Brillo platform) for 4k video live streaming.
    - Multi-threaded, event-driven architecture based on epoll.
    - Optimized end-to-end playback latency from >30s to under 2s across the full stack: device firmware, cloud backend services, and in iOS/Android apps.

- **Launched real-time cloud transcoding service.**
    - On-demand, low-latency (# b-frames + 1) transcoding to for adaptive live streaming.
    - Go and C++ integrated with Cgo, backed by ffmpeg.

- C++, Android platform, h.264, Opus/Speex, ffmpeg, Scala.

Other Work

- **Dropcam (acquired by Google).** Firmware engineer on C/Lua codebase with async I/O and green threads. Embedded firmware in C for battery-powered Bluetooth LE product.
- **Computer Vision/Camera Consultant (2020-2021).** Real-time computer vision projects with C++ and CUDA.
- **Archive Box.** Personal file organizer. Python 3. Offline-first design using CRDTs.

Education

Stanford University, B.S. Symbolic Systems. 2013