Vincent (Chenzhun) Huang

Email : chenzhuh@andrew.cmu.edu
Linkedin: https://www.linkedin.com/in/vincent-huang-6b6405149/
Mobile : +1-412-626-4852
Github: https://github.com/vincent-chenzhun-huang
Website: https://vincent-chenzhun-huang.github.io

Education

- **Carnegie Mellon University** — Pittsburgh, Pennsylvania
  *Masters of Science in Intelligent Information Systems; GPA: **3.88*** — *Sept 2022 - Dec 2023*
  Courses: Search Engines, Machine Learning, Deep Learning, Natural Language Processing, On-Device Machine Learning
- **McGill University** — Montreal, Quebec
  *Bachelor in Honours Computer Science; GPA: **3.88** (First Class Honours)* — *Sept 2017 - May 2022*
  Scholarships: Sarvard and Lee's Scholarship, Bourses d'excellence en sciences de l'informatique

  *Courses: **Parallel Computing, Operating Systems, Computer Networking**, Applied Machine Learning, Music&Audio Computing, Digital Signal Processing*

Experience

- **Apple Inc (Spring WebFlux, WebRTC, Docker, Kubernetes, Java, JavaScript)** — Austin, Texas
  Software Engineer (Intern) — *May 2023 - Present*
  - **Service Facade**: Lead the end-to-end development of a proxy layer with Spring WebFlux from scratch used by 4 teams while providing advanced customizable features (Rate Limiter, Circuit Breaker) for each of the upstream service.
  - **VoIP QoE**: Research, design, and integrate Quality of Experience(QoE) metrics for call sessions with Machine Learning into Apple's Retail Platform
  - **Deployment and Monitoring**: Integrate the application into Kubernetes cluster, set up ACL and security groups. Set up Splunk Stats Services with application.
- **Unity Technologies (PyTorch, PyTorch-Lightning, Python, Docker, C#)** — Montréal, Québec
  Machine Learning Engineer (Intern) — *May 2022 - Aug 2022*
  - **Knowledge Distillation**: Modified model architecture supported through experiments which, combined with knowledge distillation, achieving comparable performance to the original model proposed by Oreshkin et al with only 1/20 of size
  - **FPS Improvements**: Supported 5 times as many frames a second as the original model by improving model inference time
- **Unity Technologies (Node.js, JavaScript, C#)** — Montréal, Québec
  Software Developer (Intern) — *May 2021 - Aug 2021*
  - **Download Progressions**: Provided user-facing, event-oriented download progression report for Unity packages
  - **Installation**: Unified installing and caching pipeline for packages from different sources
  - **Testing**: Established decentralized testing framework for the whole team while maintaining 100% code coverage
- **SSENSE (Inversify.js, Node.js, React, Docker, Kubernetes, TypeScript, JavaScript)** — Montréal, Québec
  Full Stack Developer (Intern) — *May 2020 - Aug 2020*
  - **Automation**: Streamlined 10% of the merchandise upload, cutting inter-season product transition by 3 hours daily
  - **Access Control List**: Improved permission level management by integrating ACL(Access Control List) with Google OAuth and dependency injection framework (Inversify.js).
  - **Continuous Monitoring**: Implemented continuous monitoring into Jenkins and generated uptime service level objectives
- **Nuance Automotive (Tornado Web Services, Docker, Python)** — Montréal, Québec
  Backend Developer (Intern) — *Sept 2019 - Apr 2020*
  - **Internal Tools**: Automated internal namespace manager and significantly simplified the onboarding process for employees
  - **Error Tracer**: Designed, implemented and presented an automated error tracer scaled for 9 deployment environments

Skills Summary

- **Programming Languages**: Java, Python, JavaScript, TypeScript, C, C++, Ocaml, C#, MATLAB
- **Tools**: Spring WebFlux, Spring, AWS, GCP, Git, Docker, Kubernetes, Node, Express, SQL, PyTorch, Tensorflow, Django

Projects

- **Search Engine**: An OOP-centric abstracted search-engine library that implements a comprehensive list of Information Retrieval (IR) algorithms such as BM25, Indri, Learning-to-Rank, diversification with full support for IR queries
- **End-to-End Unsupervised ASR**: A full Implementation of Unsupervised Speech Recognition using GAN by Baevski et al
- **Hiss Reduction**: A collection of full MATLAB implementations of audio restoration algorithms
- **Error Tracer (Python, Tornado)**: A software which collects and analyzes internal server error and sends daily reports to the developers. It is connected to the ELK (ElasticSearch, Logstash, Kibanna) stack to collect, filter and analyze the logs
- **effusion.ca (Python, Django)**: Website for my A Cappella group which is built with Django and React and deployed using nginx and Digital Ocean. It is still actively maintained by me
- **JUCE Synth(C++, JUCE)**: A synthesizer built with JUCE and C++, which supports the basic MIDI operations and Open Sound Control with sliders on the UI