Stanley Z. Hua

Data Enthusiast ◇ Toronto, Canada
437-986-3444 | stanley.hua@mail.utoronto.ca | linkedin | github | website

Education

University of Toronto Toronto, CA
Honours BS Computer Science Specialist, Statistics Minor (GPA: 3.86/4) *Sept. 2019 – May 2024 (Expected)*

Technical Skills

Languages: Python, SQL, Shell Script, Git, Javascript, HTML/CSS, C/C++, Java, R, Assembly
Software Tools: GitHub, Perforce, AWS, Azure ML
Data Skills: Data Preprocessing, Data Visualization, Machine Learning, Computer Vision
Python Libraries: Pandas, NumPy, Scikit-Learn, PyTorch, Tensorflow, Matplotlib, Dask, psycopg2, OpenCV

Experience

Software Engineer Intern May 2022 – Present
Intel Corporation *Toronto, CA*

- Refactored a core tool that allows users to easily access data for benchmarking experiments from the database and file system, resulting in improved code architecture, efficiency, test coverage (99%) and backwards-compatibility.
- Saved over 1.5K hours of cloud compute, by developing a tool (in Python/SQL) to estimate runtime of queued benchmarking jobs to prevent users from abusing the high priority job queue.
- Constructed a ChartJS dashboard to monitor the number of CRON jobs starting every hour, and a tool to assist in rescheduling CRONs. This led to a decrease in volume of jobs launched per minute, increasing stability of jobs.

Junior ML Specialist May. 2022 – Present
The Hospital for Sick Children *Toronto, CA*

- Showed that (MoCo) self-supervised pretraining improves in-domain **and** out-of-domain performance on view labeling for renal ultrasound videos.
- Implemented CNN to forecast ER patient volumes, and explored Bayesian methods for confidence interval estimation (GP, Bayesian NNs).

AI Research Student Sept. 2021 – May. 2022
The Hospital for Sick Children (Goldenberg Lab) *Toronto, CA*

- Adapted video-based deep learning methods (Conv.Pooling, CNN-LSTM, TSM) to predict a kidney disease from medical (ultrasound) images taken over multiple hospital visits using PyTorch.
- Demonstrated a positive finding that single-visit models are enough to predict the kidney disease.

AI Research Student May 2021 – Sept. 2021
University of Toronto (Moses Lab) *Toronto, CA*

- Created the **CytoImageNet** dataset (890K images, 894 classes) from 20 TB of open-source microscopy images.
- Showed that CytoImageNet-pretrained features are competitive with ImageNet features on downstream datasets.
- The CytoImageNet dataset has attracted attention on Kaggle (9244 views, 509 downloads).

First-Author Publications

CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning
NeurIPS 2021 *Learning Meaningful Representations of Life Workshop* (Online)

From Single-Visit to Multi-Visit Image-Based Models: Single-Visit Models are Enough to Predict Obstructive Hydronephrosis
SIPAIM 2022 (in Valparaiso, Chile)

Conference Posters

Supervised Contrastive Learning for Improved View Labeling in Pediatric Renal Ultrasound Videos
ISBI 2023 (in Cartagena, Colombia)