

# Stanley Z. Hua

Toronto, Canada

(+1) 4379863444 | [stanley.hua@mail.utoronto.ca](mailto: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:** Airflow, AWS, Docker, GitHub, Figma

**Data Skills:** Data Engineering, Data Visualization, Machine Learning, Computer Vision, NLP

**Python Libraries:** Pandas, NumPy, Scikit-Learn, PyTorch, Tensorflow, Matplotlib, Dask, psycopg2, OpenCV

## EXPERIENCE

---

### Data Science Intern

June 2023 – Present

*Wealthsimple*

Toronto, CA

- Spearheaded design, implementation and deployment of a model to an API endpoint that detects if a user entered invalid account details when initiating an account transfer.
- Simulation showed potential saving of \$12M in flagged transfers annually.
- Coleading project to create a scalable and all-purpose labeling engine using LLMs.

### Junior ML Specialist

May 2022 – Aug 2023

*The Hospital for Sick Children*

Toronto, CA

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

### Software Engineer Intern

May 2022 – May 2023

*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.

### 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 the need for kidney surgery from medical (ultrasound) images taken over multiple hospital visits using PyTorch.
- Demonstrated positive finding that single-visit conv. models are enough to predict the need for surgery.

### 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 *LMRL* Workshop (Online)

### From Single-Visit to Multi-Visit Image-Based Models: Single-Visit Models are Enough to Predict Obstructive Hydronephrosis

SIPAIM 2022 (Valparaiso, Chile)

## CONFERENCE POSTERS & PRESENTATIONS

---

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

**\*Longitudinal Image-Based Prediction of Surgical Intervention in Hydronephrosis Patients: Perhaps Earlier Decision-Making Is Possible!**

ESPU 2023 (Lisbon, Portugal)

\*Not presenter

## INVITED TALKS

---

**CytoImageNet: A large-scale pretraining dataset for bioimage transfer learning**

Models, Inference & Algorithms Seminar, Broad Institute

Boston, USA, October 25th, 2023