

# Mengyu Yang

MACHINE LEARNING · COMPUTER VISION

✉ my.yang@gatech.edu | 🏠 mengyu.page | 🌐 mengyu-yang

## Education

---

### Georgia Institute of Technology

Atlanta, U.S.A.

*Ph.D. in Machine Learning*

Aug 2022 - Present

*Advisor: James Hays*

### University of Toronto

Toronto, Canada

*B.A.Sc. in Engineering Science with Honours (Specialization in Machine Intelligence)*

Sep 2017 - Apr 2021

**Thesis:** Building a Dataset for Music Analysis and Conditional Generation

## Publications

---

### TriBERT: Full-body Human-centric Audio-visual Representation Learning for Visual Sound Separation

- Rahman, T.; Yang, M.; Sigal, L. | NeurIPS 2021

### Soloist: Generating Mixed-Initiative Tutorials from Existing Guitar Instructional Videos Through Audio Processing

- Wang, B.; Yang, M.; Grossman, T. | 2021 ACM Conference on Human Factors in Computing Systems (CHI '21)

### Mask-Guided Discovery of Semantic Manifolds in Generative Models

- Yang, M.; Rokeby, D.; Snelgrove, X. | 4th Workshop on Machine Learning for Creativity and Design at NeurIPS 2020

### Musical Speech: A Transformer-based Composition Tool

- d'Eon, J.\*; Dumpala, S.\*; Sastry, C.\*; Oore, D.; Yang, M.; Oore, S. (\*Equal contribution) | NeurIPS 2020 Demonstration Track

## Research

---

### Georgia Institute of Technology

*Graduate Research Assistant, Advised by Prof. James Hays*

Aug 2022 - Present

- Learning robotic human-awareness through ambient environment noise

### Vector Institute for Artificial Intelligence

*Research Assistant, Advised by Prof. Alireza Makhzani*

Oct 2021 - Feb 2022

- Worked on a few-shot generation network that implicitly models unseen support sets to generate samples based on only a few example images
- Experimented with a StyleGAN2 network to take in a support set rather than single images during training and augmented various aspects of the model to stabilize training

### Vector Institute for Artificial Intelligence

*Research Intern, Co-advised by Prof. Leonid Sigal and Prof. Sageev Oore*

May 2021 - Feb 2022

- Designed and implemented cross-modal retrieval experiments for a multi-modal representation learning model, demonstrating its generalizability and the semantic meaningfulness of learned representations compared to baselines
- Experimented with a model that jointly learns cross-modal generation between video and audio, based on a GAN architecture guided by conditioning on learned features of the two modalities

## **Vector Institute for Artificial Intelligence**

### ***Undergraduate Thesis, Advised by Prof. Sageev Oore***

*Sep 2020 - May 2021*

- Built a dataset of solo piano recordings containing multi-track data of fundamental musical structural information, to address shortcomings faced by current deep learning music models which lack structural knowledge and cohesion
- Trained a Transformer model for harmonizing an input melody to use within a larger system that translates human voice into music

## **BMO Lab in Creative Research in the Arts, Performance, Emerging Technologies and AI**

### ***Research Intern, Advised by Prof. David Rokeby***

*May 2020 - Nov 2020*

- Designed an optimization-based method, guided by a custom objective function, to learn manifolds within the latent space of StyleGAN2 that correspond to localized changes in the output images (e.g. latent vectors within the manifold only change the mouth region of the same image of a face)
- Presented work as first author at the Workshop on Machine Learning for Creativity and Design at NeurIPS 2020

## **Dynamic Graphics Project, University of Toronto**

### ***Research Assistant, Advised by Prof. Tovi Grossman***

*Sep 2019 - Sep 2020*

- Created an algorithm that segments audio within guitar tutorial videos into musically meaningful phrases, used within a music learning system that allows guitar learners to easily navigate through the lesson
- Conducted a technical evaluation on the segmentation algorithm by developing tests to measure precision, recall, F1, and boundary similarity against human-labelled ground truths, with results exceeding baseline performance

## **Dynamic Graphics Project, University of Toronto**

### ***Research Assistant, Advised by Prof. Khai N. Truong***

*May 2019 - Sep 2019*

- Developed a webcam tool for face detection and pupil tracking to detect when the user has incorrect gaze response and head posture, implemented within a system for teaching piano sight reading
- Designed and implemented a dynamic-programming algorithm for identifying correctly played notes from noisy audio data, achieving **100%** accuracy on all testing examples

# **Experience**

---

## TEACHING

### **Division of Engineering Science, University of Toronto**

#### ***ESC101 & ESC102 Teaching Assistant***

*Sep 2019 - Apr 2020*

- Taught engineering design principles to classes of 20-30 students by leading individual group sessions
- Graded field note reports and core competency evaluations; provided feedback on design showcase presentations and suggested areas of improvement

## INDUSTRY

### **StratumAI**

#### ***Machine Learning Developer***

*Feb 2022 - Aug 2022*

- Developed and applied machine learning methods to create resource models for the mining industry