Mengyu Yang

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

Professional Experience

Dolby Laboratories, San Francisco CA
PhD Research Intern — *May 2025 - Present*
- Multimodal representation learning across spatial audio and real-world videos
- **Mentor**: Gautam Bhattacharya | **Collaborators**: Olha Townsend, Vishnu Raj, Andrea Fannelli

Google Research, Cambridge MA
Student Researcher — *Sep 2023 - May 2024*
- Audiovisual sound source separation and localization on open-domain videos
- **Mentor**: Scott Wisdom | **Collaborators**: Eduardo Fonseca, Arsha Nagrani, Efthymios Tzinis, John Hershey

Publications

Clink! Chop! Thud! - Learning Object Sounds from Real-World Interactions
- ***Yang, M.***; *Chen, Y.; Pei, H.; Agarwal, S.; Balajee Vasudevan, A.; Hays, J.* | ICCV 2025

The Un-Kidnappable Robot: Acoustic Localization of Sneaking People
- ***Yang, M.***; *Grady, P.; Brahmbhatt, S.; Balajee Vasudevan, A.; Kemp, K.; Hays, J.* | ICRA 2024

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

Research

Georgia Institute of Technology

Graduate Research Assistant, Advised by Prof. James Hays *Aug 2022 - Present*

- Visual perception from a multi-modal perspective, with a focus on audio and vision
- Topics include video understanding, audiovisual sound source separation and localization, and acoustic human detection for robotics

Vector Institute for Artificial Intelligence

Research Assistant, Advised by Prof. Alireza Makhzani *Oct 2021 - Feb 2022*

- Few-shot image generation using a StyleGAN2 backbone by learning an implicit representation for unseen support sets

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

Experience

Teaching

Georgia Institute of Technology

Computer Vision, Head TA *Sep 2024 - April 2025*

- Managing team of 12 TAs across various responsibilities including office hours, grading, and project development

Industry

StratumAI

Machine Learning Developer *Feb 2022 - Aug 2022*

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