

DANIEL YANG

dyang165@usc.edu ◊ dyang165.github.io ◊ (925)-918-3353

EDUCATION

University of Southern California

Expected May 2028

Ph.D. in Computer Science, GPA: 3.85/4.0

Signal Analysis and Interpretation Laboratory (SAIL), advised by Shrikanth Narayanan

National Science Foundation Graduate Research Fellow

Harvey Mudd College

May 2022

B.S., Joint Major in Computer Science and Mathematics, GPA: 3.93/4.0

Graduated with High Distinction, Engineering Departmental Honors

RESEARCH INTERESTS

- multimodal transformers
- vision-language modeling
- generative AI
- foundation models

WORK EXPERIENCE

Apple Inc

May 2024 - Present

AIML Research Intern, Siri and Information Intelligence

Working on creating generative, multimodal foundation models.

Signal Analysis and Interpretation Laboratory (SAIL)

August 2022 - Present

PhD Researcher, Advisor: Professor Shrikanth Narayanan

University of Southern California

- Understanding the relationships in data at the sample level that cause multimodal transformers to create constructive or destructive representations
- Distilling key insights on modality relationships into inductive biases for multimodal transformers

Music Information Retrieval Laboratory

January 2019 - May 2022

Undergraduate Researcher, Advisor: Professor TJ Tsai

Harvey Mudd College

Researched scalable cross-modal systems for music retrieval classification, and generation.

RECENT PROJECTS

- *Extending V-Usable Information for Multimodal Models to Understand Modal Relationships:* Why are certain multimodal relationships easy for transformers to learn? Why are others hard? We use V-Usable Information to analyze when modalities are necessary from a granular, sample-level perspective. We find that in many cases, multimodal representations from transformers fail to provide more information than unimodal ones. We find that paired multimodal data is often unnecessary and can even hurt downstream task performance.
- *GTI-MM: Robust Visual Recognition through Generative Transformer Imputation in Multi-modal Learning with Visual Modality Missing* (Under Review at CVPR 2024): We show that data imputation using generative models can be effective in improving model robustness to missing modality, especially in low-resource settings.

PUBLICATIONS:

- Yoonsoo Nam, Adam Lehavi, Daniel Yang, Digbalay Bose, Swabha Swayamdipta, Shrikanth Narayanan. Does Video Summarization Require Videos? Quantifying the Effectiveness of Language in Video Summarization. Proceedings of International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2024
- Daniel Yang, Aditya Kommineni, Mohammad Alshehri, Nilamadabh Mohanty, Vedant Modi, Jonathan Gratch, Shrikanth Narayanan. Context Unlocks Emotions: Text-based Emotion Classification Dataset Auditing with Large Language Models. Affective Computing and Intelligent Interaction (ACII), 2023.
- Daniel Yang, K. Ji, T. Tsai. A Study of Parallelizable Alternatives to Dynamic Time Warping for Aligning Long Sequences, IEEE Transactions on Audio, Speech, and Language Processing (TASLP), 2022.
- Daniel Yang, Arya Goutam, Kevin Ji, and TJ Tsai. Large-Scale Multimodal Piano Music Identification Using Marketplace Fingerprinting. Algorithms, 15(5): 146, 2022.
- Daniel Yang, T. Tsai. Composer Classification with Cross-Modal Transfer Learning and Musically Informed Augmentation. Proceedings of the International Society for Music Information Retrieval (ISMIR), 2021. Candidate for best student paper.
- Daniel Yang, Kevin Ji, and TJ Tsai. Aligning Unsynchronized Part Recordings to a Full Mix Using Iterative Subtractive Alignment. Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2021, pp. 810-817.
- Kevin Ji, Daniel Yang, and TJ Tsai. Piano Sheet Music Identification Using Marketplace Fingerprinting. Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2021, pp. 326-333.
- Kevin Ji, Daniel Yang, and TJ Tsai. Instrument Classification of Solo Sheet Music Images. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 2021, pp. 546-550.
- Daniel Yang and TJ Tsai. Piano Sheet Music Identification Using Dynamic N-gram Fingerprinting. Transactions of the International Society for Music Information Retrieval, 4(1): 42-51, 2021.
- Daniel Yang, Kevin Ji, and TJ Tsai. A Deeper Look at Sheet Music Composer Classification Using Self-Supervised Pretraining. Applied Sciences, 11(4): 1387, 2021.
- Daniel Yang and TJ Tsai. Camera-Based Piano Sheet Music Identification. Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2020, pp. 481-488.
- TJ Tsai, Daniel Yang, Mengyi Shan, Thitaree Tanprasert, and Teerapat Jenrungrot. Using Cell Phone Pictures of Sheet Music To Retrieve MIDI Passages. IEEE Transactions on Multimedia, 22(5): 3115-3127, 2020.
- Daniel Yang, Thitaree Tanprasert, Teerapat Jenrungrot, Mengyi Shan, and TJ Tsai. MIDI Passage Retrieval Using Cell Phone Pictures of Sheet Music. Proceedings of the International Society for Music Information Retrieval Conference (ISMIR), 2019, pp. 916-923.

AWARDS

- **NSF Graduate Research Fellowship:** Awarded for artificial intelligence
- **Best Student Paper Candidate:** “Composer Classification with Cross-Modal Transfer Learning and Musically Informed Augmentation” is one of three papers nominated as candidate for best student paper at ISMIR 2021
- **Harvey Mudd Shanahan Project Fund:** Ten-thousand dollar award for student-directed projects. Received three times for co-founded MuddSub robotics team in 2018, 2019, 2021