# 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: 3.90/4.0
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
- media understanding and intelligence

## WORK EXPERIENCE

**Signal Analysis and Interpretation Laboratory (SAIL)**                August 2022 - Present
*PhD Researcher, Advisor: Professor Shrikanth Narayanan*        *University of Southern California*

- Understanding the relationships in data 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.

**Collaborative Drug Discovery Clinic Team**                         August 2021 - May 2022
*Project Manager*                                                        *Harvey Mudd College*

Led a team of six undergraduate students on a collaboration with CDD, a biotechnology company for proposing diffusion models for optimized, novel molecular synthesis.

## RECENT PROJECTS

- *Is the Glass Half-Full or Half-Empty? A Mixture-Of-Tasks Perspective on Missing Modality* (Under Review at ICLR 2024):
  We reformulate the problem of vision-language missing modality robustness, which allows us to draw parallels to a rich line of work: modality competition. With this formulation, we show how to beat the current state-of-the-art in missing modality.
- *Better Initialization for Early-Fusion VLMs*:
  Early-fusion models provide a nice balance of performance and parameter efficiency. Our recent work has identified a core problem with these architectures—namely their initialization from vision models (ViT, ViT-MAE) have caused a weakness in their text features. Since text features tend to dominate image features in most multimodal classification datasets using SOTA unimodal models, we should achieve much higher performance by fixing this initialization problem.

- *Extending V-Usable Information for Multimodal Models to Understand Modal Relationships*: Why are certain modal relationships easy for transformers to learn? Why are others hard? We extend V-usable information to multimodal datasets to study this. We explore parameter-efficient methods to address overfitting so computations of V-usable information can be more precise.

## PUBLICATIONS:

- 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 MuddSub in 2018, 2019, 2021