MENGYI SHAN

mshan@hmc.edu 909-288-9488
340 E. Foothill Blvd. Claremont, CA 91711

EDUCATION

Harvey Mudd College, Claremont, CA Aug. 2017 – May 2021

- B.S. Double Major in Computer Science and Mathematics.
- Humanity, Social Science and Arts Concentration in Linguistics.
- GPA: 3.93/4.00. Major GPA: 4.00/4.00.

RESEARCH EXPERIENCE

Music Information Retrieval Lab, Student Researcher, Harvey Mudd College Jan. 2019 – Now
Advisor: Professor TJ Tsai

- **Score Following:** Developed a system that aligns a piano recording with scanned sheet music images to automatically generate a score following video. Proposed a novel alignment algorithm to handle repeats and jumps in time series of musical performance. (Paper, poster and talk at https://program.ismir2020.net/poster_1-07.html)
- **Audio Cross-Verification**: Designed an audio tampering detection system to protect world leaders from fake speech recordings. Improved the alignment algorithm to handle various types of tampering operations in audio. Constructed an LSTM attribution system based on alignment path for frame level classification.

Vobile Group, Remote Research Consultant, Santa Clara, CA Aug. 2020 – Now
Advisor: Professor George Montanez

- **Cover Song Identification**: Developed a Temporal Pooling CNN and assessed different audio features to enable the detection of subtle or complex variations of recordings. Analyzed existing literature and incorporated metadata and lyrical analysis
- **Multimedia Copyright Protection**: Extended Vobile's Mediawise fingerprinting algorithm to multimedia domains including raw video. Designed an automatic copyright protection system.

Senior Thesis Research, Student Researcher, Harvey Mudd College Sep. 2019 – May 2020
Advisor: Professor Weiqing Gu, Professor Nicholas J. Pippenger

- **Algorithmic Mathematical Derivation**: Extensively inspected existing literature on the current status of 3D object classification as a computer vision task. Summarized and documented detailed mathematical derivation for differential geometric algorithms.
- **Geometric Unified Representation:** Designed a representation to characterize surfaces with differential geometric measures. Defined and computed feature matrix and shape derivatives to describe a 3D point cloud. Applied these representations to deep learning architecture.

TEACHING EXPERIENCE

Teaching Assistant, Department of CS and Mathematics, Harvey Mudd College Sep.2018 – Now

- Tutor and grade courses for math and computer science Departments. Tutored courses include Artificial Intelligence, Machine Learning Seminar, Big Data, Discrete Mathematics,

Data Structure, Algorithms, Computability and Logic, Mathematical Analysis, and Abstract Algebra.

- Maintain course websites and code base. Design and develop programming assignments. Provide individual assistance on final projects.

Academic Excellence Team Member, Harvey Mudd College Sep. 2019 – Now

- Recommended and selected to join the special tutor team to tutor core mathematics courses (Calculus, Linear Algebra, ODE) for incoming freshmen and sophomores.
- Receive specific weekly training for tutoring lower-level classes.
- Organize tutor schedule and develop course material in close contact with professors.

Industry Experience

Software Engineering Intern, Google Inc., Los Angeles, CA May 2019 – Aug. 2019

- Experimented with importing data from manual tests. Created automatic pipeline workflow to integrate test log data from a database to an interactive dashboard.
- Tracked manual testing progress and ingested it inside of the scenario/feature coverage solution. Summarized and transformed scenario coverage data of test suites.

Student Researcher, Wolfram Research Inc., Boston MA Jun. 2018 – Jul. 2018

- Built a punctuation restoration system for English text with deep learning architectures.
- Designed, constructed and compared several recurrent neural net models.
- Presented research poster and oral talk at Wolfram Research.

Publications and Presentations

A Cross-Verification Approach for Protecting World Leaders from Fake and Tampered Audio, **Mengyi Shan**, TJ Tsai. Submitted to the International Conference on Acoustics, Speech, and Signal Processing (ICASSP) 2021

Automatic Generation of Piano Score Following Videos, **Mengyi Shan**, TJ Tsai. Submitted to Transactions of the International Society for Music Information Retrieval.

Improved Handling of Repeats and Jumps in Audio-Sheet Image Synchronization, **Mengyi Shan**, TJ Tsai. In Proceedings of the Conference of the International Society for Music Information Retrieval (ISMIR), 2020.

Geometric Unified Methods in 3D Object Classification, **Mengyi Shan**, Weiqing Gu. Poster and presentation accepted to the Southern California Applied Mathematics Symposium (SoCAMS), 2020.

Camera-Based Sheet-MIDI Passage Retrieval Using Bootleg Score Features, TJ Tsai, Daniel Yang, **Mengyi Shan**, Thitaree Tanprasert, Teerapat Jenrungrot. IEEE Transactions on Multimedia, 2020.

MIDI Passage Retrieval Using Cell Phone Pictures of Sheet Music, Daniel Yang, Thitaree Tanprasert, Teerapat Jenrungrot**, Mengyi Shan**, TJ Tsai. In Proceedings of the Conference of the International Society for Music Information Retrieval (ISMIR), 2019.

Course Projects

Environmental Audio Tagging, Team Project 2020

Designed deep convolutional transform learning structure to work on 2D log Mel spectrograms and 1D raw wave forms. Combined the two results to generate classification prediction. Implemented various data augmentation techniques to improve the performance.

Manifold Surface Representation of Human Speech, Team Project 2020
Visualized and characterized short segments of human speech recordings by representing them as 3D surfaces along time and intensity. Applied smoothing and spline techniques to approximate signal surfaces as geometric manifold.

Light Weight Singing Voice Detection in Popular Music, Individual Project 2020
Extracted various features from popular music dataset for frame level human voice detection. Experimented with and compared traditional machine learning algorithms including random forest as well as time series approaches like Hidden Markov Model.

HONORS

Chavin Prize, Harvey Mudd College Department of Mathematics Sep. 2020
Awarded to one student each year for the best senior thesis in mathematical science. First student in ten years to receive this prize as a junior undergraduate student.

Meritorious Award, The Mathematical Contest in Modeling (MCM), COMAP Apr. 2019
Awarded to 5% of the participating teams for excellent reports in modelling and problem solving.

Robert James Prize, Harvey Mudd College Department of Mathematics Sep. 2018
Awarded to two sophomores each year for outstanding performance in mathematics.

R.I.F. Scholarship, Harvey Mudd College Department of Mathematics Sep. 2018
Awarded to one student each year who demonstrates exceptional mathematical ability by distinguished performance in mathematical competition.

Honorable Mention, Putnam Math Competition, American Math Association Feb. 2018
Individually rank 93rd nationally out of over 5000 participants. Individually rank 3rd in college. College team rank 9th out of 575 institutes.

2nd Prize Team Round, Harvard-MIT Mathematics Tournament (HMMT) Feb. 2016

1st Prize, Chinese National Mathematics Olympiad Oct. 2015

5th Place, Individual Final Round, Princeton University Mathematics Competition (PuMAC) Nov. 2012

SKILLS AND LANGUAGES

Coding: Python, Java, C/C++, HTML/CSS, MATLAB, R, Mathematica, LaTeX

Spoken language: English, (native) Mandarin Chinese, Japanese

ACTIVITIES

Music: 15 years' experience playing Guzheng (Chinese zither). Highest professional certificate offered by China Conservatory of Music. Led concert performance in the Chinese Great Hall of People as the principal performer. One year's experience of private voice lesson. Half years' experience of piano.