

# Hao-Wen (Herman) Dong

PhD Candidate  
Department of Computer Science and Engineering  
University of California San Diego  
9500 Gilman Dr, La Jolla, CA 92093

Email: [salu.hwdong@gmail.com](mailto:salu.hwdong@gmail.com)  
Website: [salu133445.github.io](https://salu133445.github.io)  
Phone: [REDACTED]

## Research Interests

---

Music generation, audio synthesis, generative AI, multimodal learning, music information retrieval

## Education

---

<b>University of California San Diego</b> <i>Ph.D. in Computer Science</i> (expected June 2024) Advisors: <a href="#">Julian McAuley</a> and <a href="#">Taylor Berg-Kirkpatrick</a> Thesis: Empowering Music Creation with Machine Learning	Sep 2019 – present
<b>University of California San Diego</b> <i>M.S. in Computer Science</i>	Sep 2019 – Jun 2021
<b>National Taiwan Normal University</b> <i>Digital Video and Audio Arts Program</i>	Sep 2016 – Jun 2017
<b>National Taiwan University</b> <i>B.S. in Electrical Engineering</i>	Sep 2013 – Jun 2017

## Honors & Awards

---

<b>Rising Stars in AI</b> , <i>AI Initiative, King Abdullah University of Science and Technology (KAUST)</i>	2024
<b>Rising Stars in Data Science</b> , <i>Data Science Institute, University of Chicago and Halicioğlu Data Science Institute, UC San Diego</i>	2023
<b>IEEE SPS Scholarship</b> (\$2,000), <i>IEEE Signal Processing Society</i>	2023
<b>Friends of International Center Fellowship</b> (\$2,000), <i>Division of Graduate Education and Postdoctoral Affairs, UC San Diego</i>	2023
<b>Rising Stars in Signal Processing</b> , <i>International Conference on Acoustics, Speech, and Signal Processing (ICASSP)</i>	2023
<b>Interdisciplinary Research Award</b> (\$500), <i>Graduate and Professional Student Association, UC San Diego</i>	2023
<b>Postdoctoral Networking Tour in Artificial Intelligence</b> , <i>German Academic Exchange Service (DAAD)</i>	2023
<b>AWS Cloud Credit for Research</b> (\$2,500), <i>Amazon Web Services</i>	2023
<b>Government Scholarship to Study Abroad</b> (\$32,000), <i>Taiwan Ministry of Education</i>	2022 – 2024
<b>J. Yang Scholarship</b> (\$25,000), <i>J. Yang and Family Foundation</i>	2021 – 2022

**ECE Department Fellowship** (\$54,198), *Department of Electrical and Computer Engineering, UC San Diego* 2019 – 2020

## Professional Experience

---




<p><b>NVIDIA</b>  <i>Research Intern</i>            Host: Deep Imagination Research Group, NVIDIA Research            Advisors: <a href="#">Siddharth Gururani</a> and <a href="#">Ming-Yu Liu</a>            Topic: Controllable audio generation</p>	Sep 2023 – Dec 2023
<p><b>Adobe</b>  <i>Research Scientist/Engineer Intern</i>            Host: Audio Research Group, Adobe Research            Advisors: <a href="#">Justin Salamon</a> and <a href="#">Oriol Nieto</a>            Topic: Text-to-audio retrieval</p>	May 2023 – Sep 2023
<p><b>Dolby</b>  <i>Speech/Audio Deep Learning Intern</i>            Host: Applied AI Team, Advanced Technology Group            Advisor: <a href="#">Xiaoyu Liu</a>            Topic: Text-to-audio synthesis</p>	Jan 2023 – Apr 2023
<p><b>Amazon</b>  <i>Applied Scientist Intern</i>            Host: Natural Understanding Team, Alexa AI            Advisors: <a href="#">Wenbo Zhao</a> and <a href="#">Gunnar Sigurdsson</a>            Topic: Text-to-audio synthesis</p>	Sep 2022 – Jan 2023
<p><b>Sony</b>  <i>Student Intern</i>            Host: Tokyo Laboratory 30, R&amp;D Center            Advisor: <a href="#">Naoya Takahashi</a>            Topic: Text-queried sound separation</p>	Jun 2022 – Sep 2022
<p><b>Dolby</b>  <i>Deep Learning Audio Intern</i>            Host: Applied AI Team, Advanced Technology Group            Advisor: <a href="#">Cong Zhou</a>            Topic: Music performance synthesis</p>	Jun 2021 – Sep 2021
<p><b>Yamaha</b>  <i>Research Intern</i>            Host: AI Group, Research and Development Division            Advisor: <a href="#">Keiji Saino</a>            Topic: Deep learning based synthesizer</p>	May 2019 – Aug 2019
<p><b>Academia Sinica</b>  <i>Research Assistant</i>            Host: Music and AI Lab, Research Center for IT Innovation            Advisor: <a href="#">Yi-Hsuan Yang</a>            Topics: Music generation and deep generative models</p>	Jul 2017 – Apr 2019


## Publications





Google Scholar ID: [tEOa3O4AAAAJ](https://scholar.google.com/citations?user=tEOa3O4AAAAJ)



\* indicates equal contribution




### Conference Papers





Hao-Wen Dong, Xiaoyu Liu, Jordi Pons, Gautam Bhattacharya, Santiago Pascual, Joan Serrà, Taylor Berg-Kirkpatrick, and Julian McAuley, “CLIPSONIC: Text-to-Audio Synthesis with Unlabeled Videos and Pretrained Language-Vision Models,” *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2023. (★ Oral presentation)   



Hao-Wen Dong, Ke Chen, Shlomo Dubnov, Julian McAuley, and Taylor Berg-Kirkpatrick, “Multitrack Music Transformer,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023. (★ Oral presentation)    





Hao-Wen Dong, Naoya Takahashi, Yuki Mitsufuji, Julian McAuley, and Taylor Berg-Kirkpatrick, “CLIPSep: Learning Text-queried Sound Separation with Noisy Unlabeled Videos,” *International Conference on Learning Representations (ICLR)*, 2023.    





Ke Chen, Hao-Wen Dong, Yi Luo, Julian McAuley, Taylor Berg-Kirkpatrick, Miller Puckette, and Shlomo Dubnov, “Improving Choral Music Separation through Expressive Synthesized Data from Sampled Instruments,” *International Society for Music Information Retrieval Conference (ISMIR)*, 2022.  

Hao-Wen Dong, Cong Zhou, Taylor Berg-Kirkpatrick, and Julian McAuley, “Deep Performer: Score-to-Audio Music Performance Synthesis,” *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022.   

Hao-Wen Dong, Chris Donahue, Taylor Berg-Kirkpatrick, and Julian McAuley, “Towards Automatic Instrumentation by Learning to Separate Parts in Symbolic Multitrack Music,” *International Society for Music Information Retrieval Conference (ISMIR)*, 2021.    


Sachinda Edirisooriya, Hao-Wen Dong, Julian McAuley, and Taylor Berg-Kirkpatrick, “An Empirical Evaluation of End-to-End Polyphonic Optical Music Recognition,” *International Society for Music Information Retrieval Conference (ISMIR)*, 2021.  

Hao-Wen Dong, Ke Chen, Julian McAuley, and Taylor Berg-Kirkpatrick, “MusPy: A Toolkit for Symbolic Music Generation,” *International Society for Music Information Retrieval Conference (ISMIR)*, 2020.    

Hao-Wen Dong and Yi-Hsuan Yang, “Convolutional Generative Adversarial Networks with Binary Neurons for Polyphonic Music Generation,” *International Society for Music Information Retrieval Conference (ISMIR)*, 2018.    

Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang, “MuseGAN: Multi-Track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” *AAAI Conference on Artificial Intelligence (AAAI)*, 2018. (★ Oral presentation)   

### Journal Paper

Yin-Cheng Yeh, Wen-Yi Hsiao, Satoru Fukayama, Tetsuro Kitahara, Benjamin Genchel, Hao-Min Liu, Hao-Wen Dong, Yian Chen, Terence Leong, and Yi-Hsuan Yang, “Automatic Melody Harmonization with Triad Chords: A Comparative Study,” *Journal of New Music Research (JNMR)*, 50(1):37–51, 2021. 

## Workshop Papers .....

Weihan Xu, Julian McAuley, Shlomo Dubnov, and Hao-Wen Dong, “Equipping Pretrained Unconditional Music Transformers with Instrument and Genre Controls,” *IEEE Big Data Workshop on AI Music Generation (AIMG)*, 2023. [📄](#) [🔗](#)

Hao-Wen Dong, Gunnar A. Sigurdsson, Chenyang Tao, Jiun-Yu Kao, Yu-Hsiang Lin, Anjali Narayan-Chen, Arpit Gupta, Tagyoung Chung, Jing Huang, Nanyun Peng, and Wenbo Zhao, “CLIPSynth: Learning Text-to-audio Synthesis from Videos using CLIP and Diffusion Models,” *CVPR Workshop on Sight and Sound (WSS)*, 2023. [📄](#) [🔗](#) [📺](#)

## Demo Papers .....

Hao-Wen Dong, Wen-Yi Hsiao, and Yi-Hsuan Yang, “Pypianoroll: Open Source Python Package for Handling Multitrack Pianoroll,” *ISMIR Late-Breaking Demos*, 2018. [📄](#) [🔗](#) [🔗](#)

Hao-Wen Dong, Wen-Yi Hsiao, Li-Chia Yang, and Yi-Hsuan Yang, “MuseGAN: Demonstration of a Convolutional GAN Based Model for Generating Multi-track Piano-rolls,” *ISMIR Late-Breaking Demos*, 2017. [📄](#) [🔗](#) [🔗](#)

## Nonrefereed Papers .....

Hao-Wen Dong and Yi-Hsuan Yang, “On Output Activation Functions for Adversarial Losses: A Theoretical Analysis via Variational Divergence Minimization and An Empirical Study on MNIST Classification,” *arXiv preprint arXiv:1901.08753*, 2019. [📄](#) [🔗](#) [🔗](#)

Hao-Wen Dong and Yi-Hsuan Yang, “Training Generative Adversarial Networks with Binary Neurons by End-to-end Backpropagation,” *arXiv preprint arXiv:1810.04714*, 2018. [📄](#) [🔗](#) [🔗](#)

## Talks

---

### Tutorial Talk .....

Hao-Wen Dong and Yi-Hsuan Yang, “Generating Music with GANs: An Overview and Case Studies,” *Tutorials of the 20th International Society for Music Information Retrieval Conference (ISMIR)*, Delft, The Netherlands, Nov. 4, 2019. (3-hour tutorial) [📄](#) [🔗](#) [🔗](#)

### Invited Talks .....

“Generative AI for Music and Audio,” *CS Seminar, Department of Computer Science, Northwestern University*, Evanston, IL, USA, Nov. 15, 2023. [📄](#)

“Generative AI for Music and Audio,” *Rising Stars in Data Science Workshop, Data Science Institute, University of Chicago*, Chicago, IL, USA, Nov. 14, 2023. [📄](#)

“Generative AI for Music and Audio,” *DSP Seminar, Center for Computer Research in Music and Acoustics (CCRMA), Stanford University*, Palo Alto, CA, USA, Nov. 10, 2023. [📄](#)

“Learning Sound Separation and Synthesis from Videos using Pretrained Language-vision Models,” *Seminar, Mitsubishi Electric Research Laboratories (MERL)*, Cambridge, MA, USA, Oct. 30, 2023. [📄](#)

“Generative AI for Music and Audio,” *Seminar, Institute of Information Science, Academia Sinica*, Taipei, Taiwan, Sept. 18, 2023. [📄](#) [📺](#)

“Generative AI for Music and Audio,” *Seminar, Department of Computer Science and Information Engineering, National Taiwan University*, Taipei, Taiwan, Sept. 15, 2023. [📄](#) [📺](#)

- “Generative AI for Music and Audio,” *Seminar, Graduate Institute of Communication Engineering, National Taiwan University*, Taipei, Taiwan, Sept. 11, 2023. [📄](#) [▶](#)
- “Generating Multitrack Music using Deep Learning,” *Seminar, AI Research Center, National Institute of Advanced Industrial Science and Technology (AIST)*, Tokyo, Japan, Sept. 12, 2022. [📄](#)
- “MusPy: Symbolic Music Processing in Python,” *Music + AI Reading Group, Mila and Vector Institute*, Online, Apr. 8, 2022. [📄](#) [▶](#)
- “Towards Automatic Instrumentation by Learning to Separate Parts in Symbolic Multitrack Music,” *J. Yang Scholars Symposium, UC San Diego*, Online, Sept. 25, 2021. [📄](#) [▶](#)
- “MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” *Music AI Day, Academia Sinica*, Taipei, Taiwan, Jan. 3, 2019. [📄](#)
- “MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” *The 28th Generative Music Informatics Workshop*, Hiroshima, Japan, Aug. 20, 2018. [📄](#)
- “MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” *The 5th Taiwanese Music and Audio Computing Workshop*, Taipei, Taiwan, May 24, 2018. [📄](#)
- “MuseGAN: Multi-track Sequential Generative Adversarial Networks for Symbolic Music Generation and Accompaniment,” *AAAI Seminar, Program Office of AI Research, Taiwan Ministry of Science and Technology*, Hsinchu, Taiwan, Mar. 24, 2018. [📄](#)

## Posters .....

- Hao-Wen Dong, “Learning Text-queried Sound Separation and Synthesis using Unlabeled Videos and Pretrained Language-Vision Models,” *Rising Stars in Data Science Workshop*, Chicago, IL, USA, Nov. 14, 2023. [📄](#)
- Hao-Wen Dong, “Learning Text-queried Sound Separation and Synthesis using Unlabeled Videos and Pretrained Language-Vision Models,” *Speech and Audio in the Northeast (SANE)*, Brooklyn, NY, USA, Oct. 26, 2023. [📄](#)
- Hao-Wen Dong, “Empowering Music Creation with Machine Learning,” *ICASSP Rising Stars in Signal Processing Workshop*, Rhodes, Greece, June 9, 2023. [📄](#)
- Hao-Min Liu, Hao-Wen Dong, Wen-Yi Hsiao, and Yi-Hsuan Yang, “Lead sheet and Multi-track Piano-roll generation using MuseGAN,” *GPU Technology Conference Taiwan (GTC Taiwan)*, Taipei, Taiwan, May 30, 2018. [📄](#)

## Teaching Experience

<b>Teaching Assistant</b> <i>Principles of Machine Learning: Learning Algorithms (CSE 251A), UC San Diego</i> Student evaluations: 100% recommendations (71.4% strong recommendations)	Winter 2022
<b>Teaching Assistant</b> <i>Recommender Systems &amp; Web Mining (CSE 258), UC San Diego</i> Student evaluations: 100% recommendations (78.6% strong recommendations)	Fall 2021
<b>Teaching Assistant</b> <i>Topics in CSE: Advanced Statistical NLP (CSE 291), UC San Diego</i> Student evaluations: 100% recommendations (75% strong recommendations)	Spring 2021

<b>Teaching Assistant</b> <i>Introduction to Machine Learning (CSE 151A), UC San Diego</i>	Winter 2021
<b>TA Lead on Remote Teaching</b> <i>Recommender Systems &amp; Web Mining (CSE 258), UC San Diego</i>	Fall 2020
<b>Teacher</b> <i>AMC 8 Training Camp (2-hour weekly math classes), Kang Chiao International School</i>	Fall 2015

## Mentoring Experience

---

### Research Mentoring .....

<b>Amit Namburi</b> , undergraduate student, <i>UC San Diego</i> , Topic: Music recommender systems	2023 – present
<b>Wanning Lu</b> , undergraduate student, <i>Early Research Scholars Program (ERSP), UC San Diego</i> , Topic: Music performance rendering	2023 – present
<b>Minh Hoang Phan</b> , undergraduate student, <i>Early Research Scholars Program (ERSP), UC San Diego</i> , Topic: Music performance rendering	2023 – present
<b>Diego Reyes</b> , undergraduate student, <i>Early Research Scholars Program (ERSP), UC San Diego</i> , Topic: Music performance rendering	2023 – present
<b>Wei Ji Chen</b> , undergraduate student, <i>UC San Diego</i> , Topic: Music generation	2023 – present
<b>Weihan Xu</b> , graduate student, <i>Duke University</i> , Topic: Music generation	2023 – present
<b>Chris Francis</b> , graduate student, <i>UC San Diego</i> , Topic: Music generation	2023 – present
<b>Sanjayan Sreekala</b> , graduate student, <i>UC San Diego</i> , Topic: Audio synthesis	2023
<b>Chih-Pin Tan</b> , research assistant, <i>Academia Sinica</i> , Topic: Music generation	2022 – 2023

### Career Mentoring .....

<b>WiMIR Mentoring Program</b> , <i>Women in Music Information Retrieval</i> 2023: Julien Guinot 2022: Kavya Saxena	2022 – present
<b>project Tyra Mentoring Program</b> , <i>Taiwanese Young Researcher Association</i> 2023: Tien Yu Chang, Hao-Cheng Chang, Yun Ho 2022: Jhan-Shuo Liu, Hou-Chun Liu, Chia-Yuan Hsu 2021: Yen-Chen Shih, Li-Ching Chen, Yu-Chia Chen, Yu-Yang Kao, Yen-Fu Lu	2021 – present

## Professional Service

---

<b>Conference Reviewer</b> , <i>ICASSP (2024), ICLR (2024), NeurIPS (2022–2023), ISMIR (2020–2023), AIMC (2022–2023), ICML (2023), CHI Late-Breaking Work (2022), ISMIR Late-Breaking Demos (2021)</i>
<b>Journal Reviewer</b> , <i>TASLP (2023), TOMM (2023), EURASIP JASMP (2022–2023), SoftwareX (2022–2023), EURASIP JASP (2022), PJS (2022), JOCCH (2022), TISMIR (2022), IEEE Transactions on Games (2020), Scientific Reports (2020), Neurocomputing (2019)</i>
<b>Book Reviewer</b> , “ <i>Generative AI with Python and TensorFlow 2</i> ” by Joseph Babcock and Raghav Bali, Packt Publishing, 2021

## Music, Films & Dramas

---

Bass Singer, <i>NTUChorus</i>	2013 – 2017
Producer, “ <i>FragMemory</i> ”, short film, color, 39min ▶	2017
Conductor, <i>NTUEE Concert Choir</i> ▶	2017
Student Conductor, <i>NTUChorus</i>	2015 – 2016
Student Conductor, <i>NTUChorus Summer Concert</i> , Taiwan National Concert Hall ▶ ▶	2016
Student Conductor, <i>NTUChorus Winter Concert</i> , Taiwan National Concert Hall ▶	2016
Bass Leader, <i>NTUChorus</i>	2014 – 2015
Director and Writer, “ <i>Eye of Horus</i> ”, Sophomore Drama, NTUEE Night ▶	2015
Leading Actor, “ <i>The Magus</i> ”, Freshman Drama, NTUEE Night ▶	2014
Music Director, “ <i>Life is ___.</i> ”, Graduation album for NKNUSH Class of 2013 ▶	2013
Writer, Vocal and Drums, “ <i>No longer ever</i> ”, Graduation song for NKNUSH Class of 2013 ▶	2013

## References

---

### Julian McAuley

*Professor*

Department of Computer Science and Engineering  
University of California San Diego

[jmcauley@ucsd.edu](mailto:jmcauley@ucsd.edu)

██████████

### Taylor Berg-Kirkpatrick

*Associate Professor*

Department of Computer Science and Engineering  
University of California San Diego

[tberg@ucsd.edu](mailto:tberg@ucsd.edu)

██████████

### Yi-Hsuan Yang

*Professor*

Department of Electrical Engineering  
National Taiwan University

[yhyangtw@ntu.edu.tw](mailto:yhyangtw@ntu.edu.tw)

██████████

### Shlomo Dubnov

*Professor*

Department of Music  
University of California San Diego

[sdubnov@ucsd.edu](mailto:sdubnov@ucsd.edu)

██████████