Seungheon Doh

Ph.D Candidate @ Music and Audio Computing Lab, KAIST
seungheon.doh@gmail.com
blog • google scholar • github

Research Interest

Music Informational Retrieval, Natural Language Processing, Multi-Modal Deep Learning

My research focuses on the machine's ability to listen to music, express music experience in natural language, and imagine visuals. My research primarily revolves around the field of representation learning between music and multimodal media. The detailed task is as follows:

- Annotation: Music Auto-Tagging, Music Captioning
- Retrieval: Cross-modal Music Retrieval, Conversational Music Retrieval
- Generation: Music Conditioned Media Generation
- Representation Learning: Self-supervision, Disentanglement

Selected Publications

LP-MusicCaps: LLM-based Pseudo Music Captioning
Seungheon Doh, Keunwoo Choi, Jongpil Lee, Juhan Nam
Proceedings of the 24nd International Society for Music Information Retrieval Conference (ISMIR), 2023

Toward Universal Text-to-Music Retrieval
Seungheon Doh, Minz Won, Keunwoo Choi, Juhan Nam
Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023

Textless Speech-to-Music Retrieval Using Emotion Similarity
Seungheon Doh, Minz Won, Keunwoo Choi, Juhan Nam
Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023

Musical Word Embedding: Bridging the Gap between Listening Contexts and Music
Seungheon Doh, Jongpil Lee, Tae Hong Park, and Juhan Nam
Machine Learning for Media Discovery Workshop, International Conference on Machine Learning (ICML), 2020

Academic Papers (peer-reviewed conferences and workshops)

Hi, KIA: A Speech Emotion Recognition Dataset for Wake-Up Words
Taesu Kim*, **SeungHeon Doh***, Gyunpyo Lee, Hyung seok Jun, Juhan Nam, Hyeon-Jeong Suk (*equally contribution)
Proceedings of the 14th Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2022

EMOPIA: A Multi-Modal Pop Piano Dataset For Emotion Recognition and Emotion-based Music Generation
Hung, Hsiao-Tzu and Ching, Joann and **Doh, Seungheon** and Kim, Nabin and Nam, Juhan and Yang, Yi-Hsuan
Proceedings of the 22nd International Society for Music Information Retrieval Conference (ISMIR), 2021

Music Playlist Title Generation Using Artist Information
Haven Kim, **Seungheon Doh**, Junwon Lee, Juhan Nam
1st Workshop on Creative AI Across Modalities (AAAI), 2023

Music Playlist Title Generation: A Machine-Translation Approach
Seungheon Doh, Junwon Lee, Juhan Nam
2nd Workshop on Natural Language Processing for Music and Spoken Audio (NLP4MusA, ISMIR), 2021

Million Song Search: Web Interface for Semantic Music Search Using Musical Word Embedding
Seungheon Doh, Jongpil Lee, Juhan Nam
Late Breaking Demo in the 22st International Society for Music Information Retrieval Conference (ISMIR), 2021

TräumerAI: Dreaming Music with StyleGAN
Dasaem Jeong, **Seungheon Doh**, Taegyun Kwon
Workshop on Machine Learning for Creativity and Design, Neural Information Processing Systems (NeurIPS), 2020

Education

Korea Advanced Institute of Science and Technology (KAIST) 2021 - present

Ph.D Student in Graduate School of Culture Technology
Music and Audio Computing Lab (Advisor: Juhan Nam)

Korea Advanced Institute of Science and Technology (KAIST) 2019 - 2021
MSc. in Graduate School of Culture Technology
Music and Audio Computing Lab (Advisor: Juhan Nam)

Ulsan National Institute of Science and Technology (UNIST) 2014 - 2019
B.S. in School of Business administration & Industrial Design
Academic Performance Scholarship Recipient for every semester

Experience

Naver Corp. Dec 2022 - Mar 2023
Seongnam, South Korea
Research Intern in Now AI Team (Advisor: Jeong Choi)

Computational Intelligence, Learning, Vision, and Robotics (CILVR) Jun 2022 - Aug 2022
New York University, United States
Visiting Student (Advisor: Kyunghyun Cho)

ByteDance Jul 2021 – Jan 2022
Remote (due to COVID-19)
Research Intern in Speech, Audio & Music Intelligence Team (Advisor: Keunwoo Choi, Minz Won)

Music and Audio Research Laboratory (MARL) Dec 2019 - Feb 2020
New York University, United States
Visiting Student (Advisor: TaeHong Park)

Advising

Have advised research internship students. Academic papers have been presented in top-tier conferences and workshops (ISMIR, NLP4MusA).

- Mentor, Gwondae Yong - Computer Science, KAIST (Jun 2023 - Dec 2023)
- Mentor, Junwon Lee - Electronic Engineering, KAIST (Jun 2021 - Sep 2022)
- Mentor, Hounsu Kim - Mechanical Engineering, KAIST (Dec 2020 - Mar 2021)
- Mentor, Nabin Kim - CS/Music, Georgia Institute of Technology (Dec 2020 - Mar 2021)

Teaching Activity

- TA, GCT731 Topics in Music Technology: Generative AI for Music (Mar 2023)
- TA, GCT634 Musical Applications of Machine Learning (Sep 2022)
- TA, GCT634 Musical Applications of Machine Learning (Sep 2021)
- TA, GCT731 Topics in Music Technology: Cognitive Science of Music (Sep-2020)
- TA, GCT576 Social Computing, KAIST GSCT (Sep 2019)

Talk

- Invited Talk, Multimodal Music Retrieval for Listener and Contents Creator, Seoul Univ. (Feb 2023)
- Invited Talk, Music Informational Retrieval with Natural Language Processing, YONSEI Univ.(Dec 2022)
- Invited Talk, Digital Signal Processing and Speech Recognition, SK planet (May 2020)
- Tutorial, Music and Deep Learning, PyconKR (Aug 2019)

Service

- Korean translator, NYU Deep Learning DS-GA 1008, Yann LeCun & Alfredo Canziani (Sep 2020)

Languages & Skill & Interests

- English(fluent), Korean(native)
- Python (Pytorch, Tensorflow), HTML/CSS, React.js, C++
- Media Art, Graphic Design, Prototyping & Making, Piano, Movie, Crossfit, Lifting