

# SEUNGHEON DOH

Ph.D Candidate @ Music and Audio Computing Lab, KAIST

seungheon.doh@gmail.com

[blog](#) • [google scholar](#) • [github](#)

## RESEARCH INTEREST

---

Music Informational Retrieval, Multimodal Deep Learning, Natural Language Processing

My research focuses on the machine's ability to listen to music, express music experience in natural language, and imagine visuals. A key aspect of my research lies in representation learning, particularly in bridging the gap between music and multi-modal media. Presently, my primary focus is on **multi-turn conversation** and **multi-modality**. I aim for machines to comprehend diverse modalities during conversations, particularly in the context of music generation and retrieval, facilitating the discovery of music through dialogue. The detailed task is as follows:

- Annotation: Music Understanding, Captioning, and Reasoning
- Retrieval: Cross-modal Music Retrieval, Multi-turn Conversational Music Retrieval
- Generation: Multi-turn Conversational Generation, Reference & Text Conditioned Generation

## SELECTED PUBLICATIONS

---

### Enriching Music Descriptions with a Finetuned-LLM and Metadata for Text-to-Music Retrieval

Seungheon Doh, Minhee Lee, Dasaem Jeong, Juhan Nam

Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2024  
(Submitted)

### LP-MusicCaps: LLM-based Pseudo Music Captioning

Seungheon Doh, Keunwoo Choi, Jongpil Lee, Juhan Nam

Proceedings of the 24th International Society for Music Information Retrieval Conference (ISMIR), 2023  
([Best Paper Award - Long Talk](#))

### 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

### 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

## ACADEMIC PAPERS (PEER-REVIEWED CONFERENCES AND WORKSHOPS)

---

### The Song Describer Dataset: a Corpus of Audio Captions for Music-and-Language Evaluation

Ilaria Manco, Benno Weck, Seungheon Doh, Yixiao Zhang, Dmitry Bogdanov, Yusong Wu, Ke Chen, Philip Tovstogan, Emmanouil Benetos, Elio Quinton, George Fazekas, Juhan Nam, Minz Won

Neural Information Processing Systems, Workshop on Machine Learning for Audio (NeurIPS), 2023  
(Submitted)

### Annotator Subjectivity in the MusicCaps Dataset

Minhee Lee, Seungheon Doh, Dasaem Jeong

2nd Workshop on Human-Centric Music Information Research (HCMIR), 2023

### Music Playlist Title Generation Using Artist Information

Haven Kim, Seungheon Doh, Junwon Lee, Juhan Nam

1st Workshop on Creative AI Across Modalities (AAAI), 2023

### 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

### Music Playlist Title Generation: A Machine-Translation Approach

**Seunghoon 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**

**Seunghoon Doh**, Jongpil Lee, Juhan Nam

Late Breaking Demo in the 22st International Society for Music Information Retrieval Conference (ISMIR), 2021

**Musical Word Embedding: Bridging the Gap between Listening Contexts and Music**

**Seunghoon Doh**, Jongpil Lee, Tae Hong Park, and Juhan Nam

Machine Learning for Media Discovery Workshop, International Conference on Machine Learning (ICML), 2020

**TräumerAI: Dreaming Music with StyleGAN**

Dasaem Jeong, **Seunghoon 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)

Topic: Multimodal Video Recommendation System (User-to-Item)

**Computational Intelligence, Learning, Vision, and Robotics (CILVR)**

Jun 2022 - Aug 2022

New York University, United States

Visiting Student (Advisor: Kyunghyun Cho)

Topic: Self-supervised Learning Model for Music

**ByteDance / Tiktok**

Jul 2021 - Jan 2022

Remote (due to COVID-19)

Research Intern in Speech, Audio & Music Intelligence Team (Advisor: Keunwoo Choi, Minz Won)

Topic: Multimodal Music Retrieval System (Text-to-Music, Speech-to-Music)

**Music and Audio Research Laboratory (MARL)**

Dec 2019 - Feb 2020

New York University, United States

Visiting Student (Advisor: TaeHong Park)

Topic: Word Embedding for Music Annotation and Retrieval

---

OPEN-SOURCE SOFTWARE

- **LP-MusicCaps** Music Captioning System (Released 2023, 146+ stars):
- **TTMR** Text-to-Music Retrieval System (Released 2022, 92+ stars):

---

ADVISING

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

- Gwondae Yong - Computer Science, KAIST (Jun 2023 - Dec 2023)
- Minhee Lee - Computer Science, Sogang Univ. (May 2023 - Oct 2023)

- Junwon Lee - Electronic Engineering, KAIST (Jun 2021 - Sep 2022)
- Hounsung Kim - Mechanical Engineering, KAIST (Dec 2020 - Mar 2021)
- 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, Music and Audio Research Group (MARG), Seoul National 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

---

- Reviewer, ICASSP 2024
- 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

#### REFERENCES

---

- |  |                        |
|--|------------------------|
| • Juhan Nam, Professor, KAIST  | juhan.nam@gmail.com    |
| • Keunwoo Choi, Senior Research Director/Engineer, Gaudio Lab/Prescient Design | gnuchoi@gmail.com      |
| • Minz Won, Research Scientist, ByteDance/Tiktok                               | minz.s.won@gmail.com   |
| • Jongpil Lee, CEO, Neutune  | jongpillee@neutune.com |