GIST wins Judge's Award at international AI sound recognition competition: School of Electrical Engineering and Computer Science Professor Hong Kook Kim's Lab, 'AunionAI' team members Yun-Ah Song and Do-Hyeon Lee

- Won first place in the Language Query-based Audio Source Separation (LASS) category of the 'DCASE Challenge 2024' hosted by the IEEE Signal Processing Society (AASP) and also received the Judges' Award

- Presented a new technique for data augmentation based on the Large Language Model (LLM) and achieved results… "As a case where the lab's achievements led to the possibility of commercialization, it will contribute to creating a more convenient and safer life."

▲ Yun-Ah Song, a student in the combined master's and doctoral program at GIST Audio Intelligent Technology & Research Lab (AiTeR, Professor Hong Kook Kim), is receiving the Judge's Award at the DCASE Challenge 2024 workshop.

The Gwangju Institute of Science and Technology (GIST, President Kichul Lim) announced that the 'GIST-AunionAI' team (Yun-Ah Song and Do-Hyeon Lee, combined master's and doctoral programs), comprised of students from the Audio Intelligent Technology & Research Lab (AiTeR, Professor Hong Kook Kim) in the School of Electrical Engineering and Computer Science, won the Judges' award at the 'International AI

Sound Recognition Competition (DCASE Challenge 2024*)' workshop hosted by the Audio and Acoustic Signal Processing (AASP) Society of the Institute of Electrical and Electronics Engineers (IEEE).

The 'GIST-AunionAI' team, with the support of AunionAI, founded by Professor Hong Kook Kim, won first place in the 'Language-Queried Audio Source Separation (LASS: Language-Queried Audio Source Separation, Task 9)' category of the 'DCASE Challenge 2024' in July. In addition, at this workshop held in October, they won the 'Judges' Award', beating out leading research institutes at home and abroad, thereby being recognized for their excellent achievements in various criteria such as research innovation, technological completeness, and paper quality.

* DCASE Challenge: Challenge on Detection and Classification of Acoustic Scenes and Events

The Judges' Award, which the 'GIST-AunionAI' team received this time, is evaluated based on the technical report submitted to the DCASE 2024 Challenge and is awarded to only one team per task, based on the originality and excellence of the research, separate from the challenge ranking.

'Language Query-based Audio Source Separation (LASS) Technology' is a technology that separates audio signals that match the text entered by the user.

For example, if you input a text query such as 'Someone is cutting something and there is a crunching sound', it will extract only the crunching sound from the audio mixed with various kitchen noises. This technology provides the basis for developing a generative AI model that connects language and audio, and can be utilized in various application fields such as automatic audio editing, multimedia content search, and augmented listening.

In this workshop, the 'GIST-AunionAI' team presented a high-performance language-queried audio source separation technology by grafting AI technologies that can express various audio intelligence. (Paper: Large-Language-Model-Based Caption Augmentation for Language-Queried Audio Source Separation)

The research team improved the AI model through ▴ LLM (Large Language Model)-based prompt technology and data augmentation technology, ▴ pre-learning training model* and existing model inference result fusion technology, and ▴ ensemble technology to improve AI capabilities.

* pre-learning training model: a large-capacity model learned on a large data set
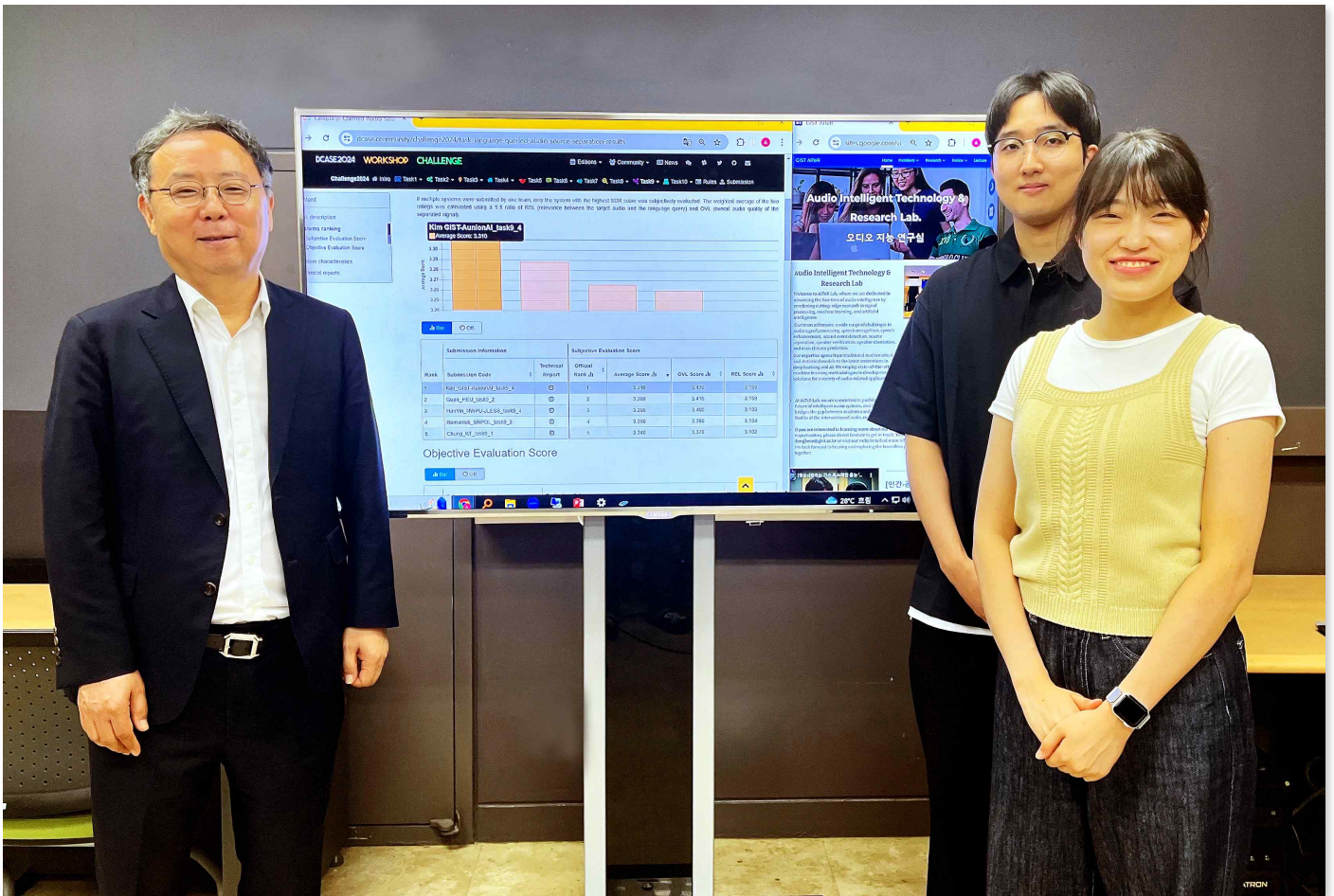
▲ (From left) Professor Hong Kook Kim, student Do-Hyeon Lee, and student Yun-Ah Song of the GIST Audio Intelligent Technology & Research (AiTeR)

Professor Hong Kook Kim said, "The significance of the AI model developed through collaboration between GIST Lab and Aunion AI Co., Ltd. is very great in that it does not remain in the lab but suggests the possibility of commercialization. In particular, we will continue to improve the LLM-based audio generation and recognition AI model and apply it to various fields to contribute to the development of technology for a more convenient and safe life."

The 'GIST-AunionAI' team members said, "Thanks to the guidance and generous support of Professor Hong Kook Kim, we were able to achieve good results in international competitions. Based on the achievements so far, we plan to further deepen our research in the field of audio artificial intelligence and contribute to the practical application of audio intelligence technology and the creation of social value."

Meanwhile, the GIST Audio Intelligent Technology & Research Lab (AiTeR, Professor Hong Kook Kim), which is researching various AI models related to voice and audio, is conducting various researches such as language query-based voice source separation, as well as acoustic event detection, voice synthesis, voice noise removal, voice recognition, abnormal situation detection, multilingual recognition, and translation, in collaboration with domestic industries, universities, and research institutes, as well as overseas research institutes such as the Massachusetts Institute of Technology (MIT) in the United States.

'Language query-based audio source separation' was conducted with the support of the MIT International Joint Research Project, the GIST Science and Technology Innovation Foundation's 'Practical Research and Development Project', and the Research and Development Special Zone Promotion Foundation's 'Science and Technology Project Opening the Future of the Region' project.

The DCASE workshop was held in Tokyo, Japan from October 23 to 25, and the award ceremony was held on the 25th at the 'Shinagawa Season Terrace' building.

GIST
Since 1993