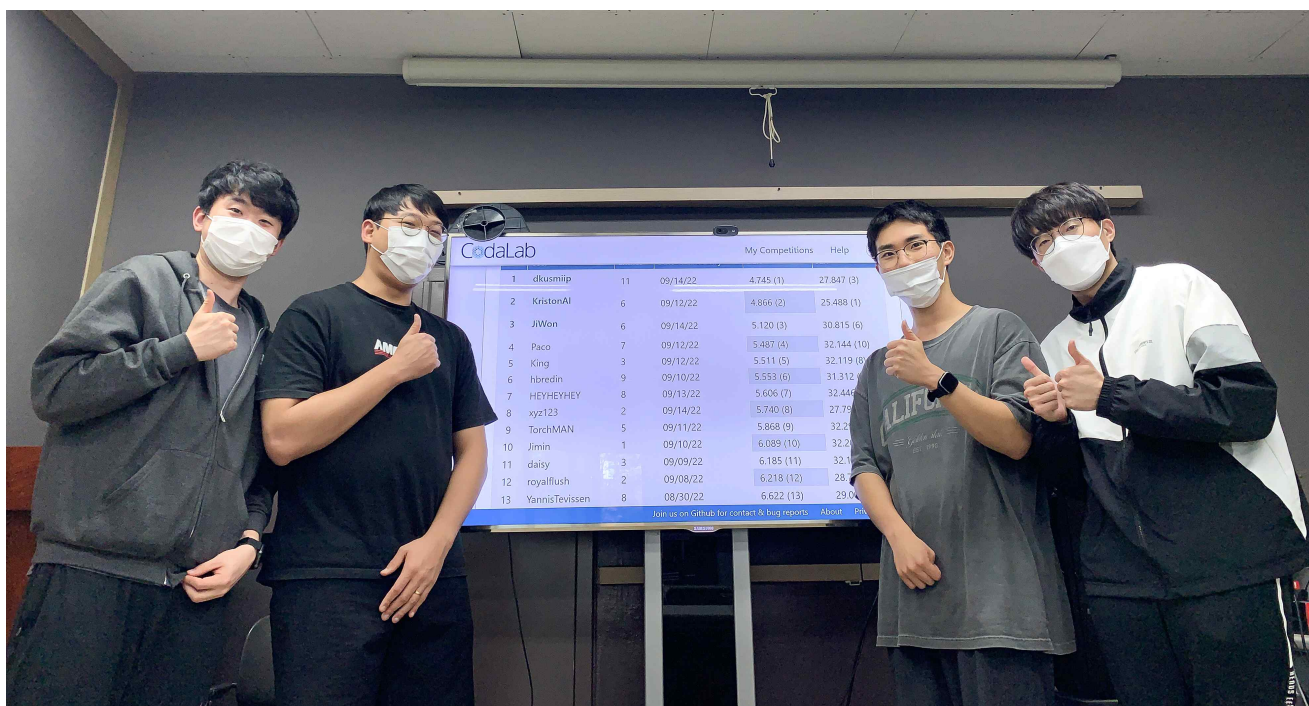# Hosted by Oxford University, **G**IST Audio Intelligent Lab won 3rd place in 「VoxSRC Challenge 2022」

- School of Electrical Engineering and Computer Science Professor Hong Kook Kim's research team, students and researchers, won 3rd place in the speaker division

- Designing an artificial intelligence model that classifies each person's voice section from the voices of multiple speakers

▲ The GIST-AiTeR team, composed of students and researchers from the Audio Intelligent Lab (AiTeR) of Professor Hong-Guk Kim at the School of Electrical Engineering and Computer Science Professor, won 3rd place in the 'VoxSRC Challenge 2022' track 4 'Speaker Segmentation', hosted by the VGG Group at Oxford University in the UK. did. (From right, master's student Kyung-wan Park, integrated student Dong-gun Park, researcher Ye-chan Yoo, and master's student Ji-won Kim

GIST (Gwangju Institute of Science and Technology, President Kiseon Kim) students who study the field of audio intelligence won a prize at the International Speaker Recognition competition.

The GIST-AiTeR team (master's student Kyung-wan Park, integrated student Dong-gun Park, researcher Ye-chan Yoo, and master's student Ji-won Kim), composed of students and researchers belonging to the Audio Intelligent Lab (AiTeR) of Professor Hong-Guk Kim of the School of Electrical Engineering and Computer Science, is hosted by the VGG Group of Oxford University in the UK. They won 3rd place in 「VoxSRC Challenge 2022」 * Track 4 'Speaker Division' held online from July 19th to August 14th.

* VoxCeleb Speaker Recognition Challenge (VoxSRC): To encourage the development of technology in the field of speaker recognition using big data recorded in various environments by the Visual Geomery Group (VGG) group of Oxford University, UK, which researches and develops image processing AI models. International Speaker Recognition Contest was held every year since 2019.

'Speaker segmentation technology' is a technology to effectively distinguish 'who spoke when' by dividing the audio containing the voices of two or more speakers

for each individual speaker. This technology is essential for voice-related applications such as business meetings and interviews for recognizing audio data conducted interactively by multiple speakers.

In this competition, the GIST-AiTeR team developed a technology to classify speech sections by speakers from the conversational voices of multiple speakers. To do this required ▲ technology to remove ambient noise from voice signals ▲ technology to recognize the speaker using the transformer technique among the AI models ▲ echnology for discriminating speech sections for each speaker and dividing sections where the speakers are uttering at the same time ▲ echnology to correct errors for separated vocalization sections using the clustering technique ▲ an artificial intelligence model including ensemble technology that merges various techniques to improve the ability of artificial intelligence.

▲ integrated student Dong-gun Park of the GIST-AiTeR team who participated in 「VoxSRC Challenge 2022」, is giving a presentation.

The participating students of the GIST-AiTeR team said, "We were able to get good results at the international competition thanks to the generous support from my advisor and the school. Multi-speaker conversational speech recognition based on artificial intelligence based on the experience gained and developed technology while preparing for this competition. We will endeavor to publish excellent thesis and commercialize it by improving performance so that it can be used in real life by grafting speech synthesis adaptive to speaker changes."

This competition was held for the purpose of business meeting or speaker classification using interview data to develop an artificial intelligence learning model using speaker recognition data. and at the <Interspeech 2022 Conference>, an international academic conference related to AI.

Meanwhile, the GIST Audio Intelligent Lab (AiTeR, Supervisor Hong-Guk Kim), to which the GIST-AiTeR team belongs, is conducting various studies such as removing not only the speaker but also voice noise, voice recognition, acoustic event detection, abnormal situation detection, and vocal cord disease detection.

This study is a 2022 research project of the Ministry of Culture, Sports and Tourism and the Korea Creative Content Agency to develop robots and service contents for children's reading activities based on artificial intelligence and was carried out with the support of the Ministry of Science and ICT and the Information and Communication Planning and Evaluation Institute's media content voice language localization technology development project.

GIST
Since 1993