

# “소리만으로 집안일 알아채는 AI” 지스트-한화비전, 국제 AI음향인식대회 1위

- 지스트 오디오지능연구실-한화비전, 실내에서 발생하는 소리 탐지 AI모델 설계
- AI음향인식대회「DCASE 챌린지 2023」서 실내 음향 이벤트 탐지 부문 1위 쾌거



▲ (왼쪽부터) 지스트 송윤아 학생, 김홍국 교수, 김지원, 손상원 학생  
(화면 위부터) 한화비전 송일훈 연구원, 임정은 수석연구원

지스트(광주과학기술원, 총장직무대행 박래길) 공동연구팀이 국제 AI음향인식대회인 「DCASE 챌린지 2023」\*에서 세계 유수 대학과 연구소를 제치고 지난 1일 실내 음향 이벤트 탐지 부문 1위의 성과를 거뒀다.

\* DCASE 챌린지: 국제 음향 장면 및 이벤트 탐지 분류 경진대회(Challenge on Detection and Classification of Acoustic Scenes and Events)

이 대회는 국제전기전자공학회(IEEE) 산하 신호처리소사이어티(AASP)가 2013년부터 개최하는 세계적인 대회로, 인공지능(AI)을 이용해 소리를 듣고 상황을 판단하는 음향 인식 기술을 겨룬다.

지스트 전기전자컴퓨터공학부 김홍국 교수가 지도하는 오디오지능연구실 학생들과 한화비전 AI연구소(소장 노승인) 연구원으로 구성된 '지스트-한화비전팀(지스트 김지원, 손상원, 송윤아 학생, 한화비전 송일훈 연구원, 임정은 수석연구원)'은 실내 음향 이벤트 탐지 부문에서 단일 모델 분야 1위, 앙상블 모델 분야 2위를 달성하며 연구성과의 우수성을 입증했다.

실내 음향 이벤트 탐지 부문에서는 실내 환경에서 발생하는 진공청소기, 설거지 소리, 개 짖는 소리, 물 흐르는 소리 등 10가지 소리를 AI 기술로 탐지하고 구별하는 능력으로 순위를 매긴다.

지스트-한화비전팀은 다양한 오디오 지능을 구현할 수 있는 AI 기술을 접목함으로써 성능을 대폭 높였다. 구체적으로는 ▲인공지능이 추론한 의사 정답을 학습에 활용하는 준지도 학습 기술, ▲사전 학습 훈련 모델\*의 추론 결과를 기존 모델의 추론 결과와 융합하는 기술, ▲성능 향상을 위해 데이터를 정제하는 기술, 그리고 ▲인공지능의 성능을 향상하는 여러 기법들을 병합하는 앙상블 기술을 적용했다.

이 기술을 활용하면 카메라로 볼 수 없는 상황에서도 소리만 듣고도 무슨 일인지 탐지할 수 있어 실내 감시나 시스피커 등 다양한 어플리케이션에 활용될 것으로 기대된다.

\* 사전 학습 훈련 모델 : 대량의 데이터셋에 대해 학습된 대용량의 모델

김홍국 교수는 "이번 대회를 통해 얻은 경험과 기술은 한화비전이 개발하는 CCTV에서 음향 이벤트를 탐지하는데 적용될 예정"이라며, "나아가 소셜미디어 콘텐츠에서 음성구간 및 음향 이벤트를 검출하는 기술 등 더 효율적이고 사용자 친화적인 서비스를 개발하기 위해 노력하겠다"고 말했다.

지스트 학생들은 "한화비전 연구원들과의 토론과 김홍국 교수님의 지도 덕분에 좋은 결과를 얻을 수 있었다"며, "이번 성과에 만족하지 않고 오디오 지능에 대한 AI 모델을 꾸준히 발전시킬 수 있도록 연구에 박차를 가하겠다"고 포부를 밝혔다.

한편, 지스트 오디오지능연구실(AiTeR)에서는 음성.음향 인식을 위한 AI 모델을 연구하고 있다. 또한 음향 이벤트 탐지, 음성 합성, 음성 잡음 제거, 음성 인식, 이상 상황 감지, 생체정보 처리 등 다양한 연구를 MIT와 같은 국내외 산학연기관과 공동으로 진행하고 있다.

이번 연구는 한화비전, 과학기술정보통신부 및 정보통신기획평가원의 미디어 콘텐츠 음성 언어 현지화 기술개발 사업의 지원으로 수행됐다.