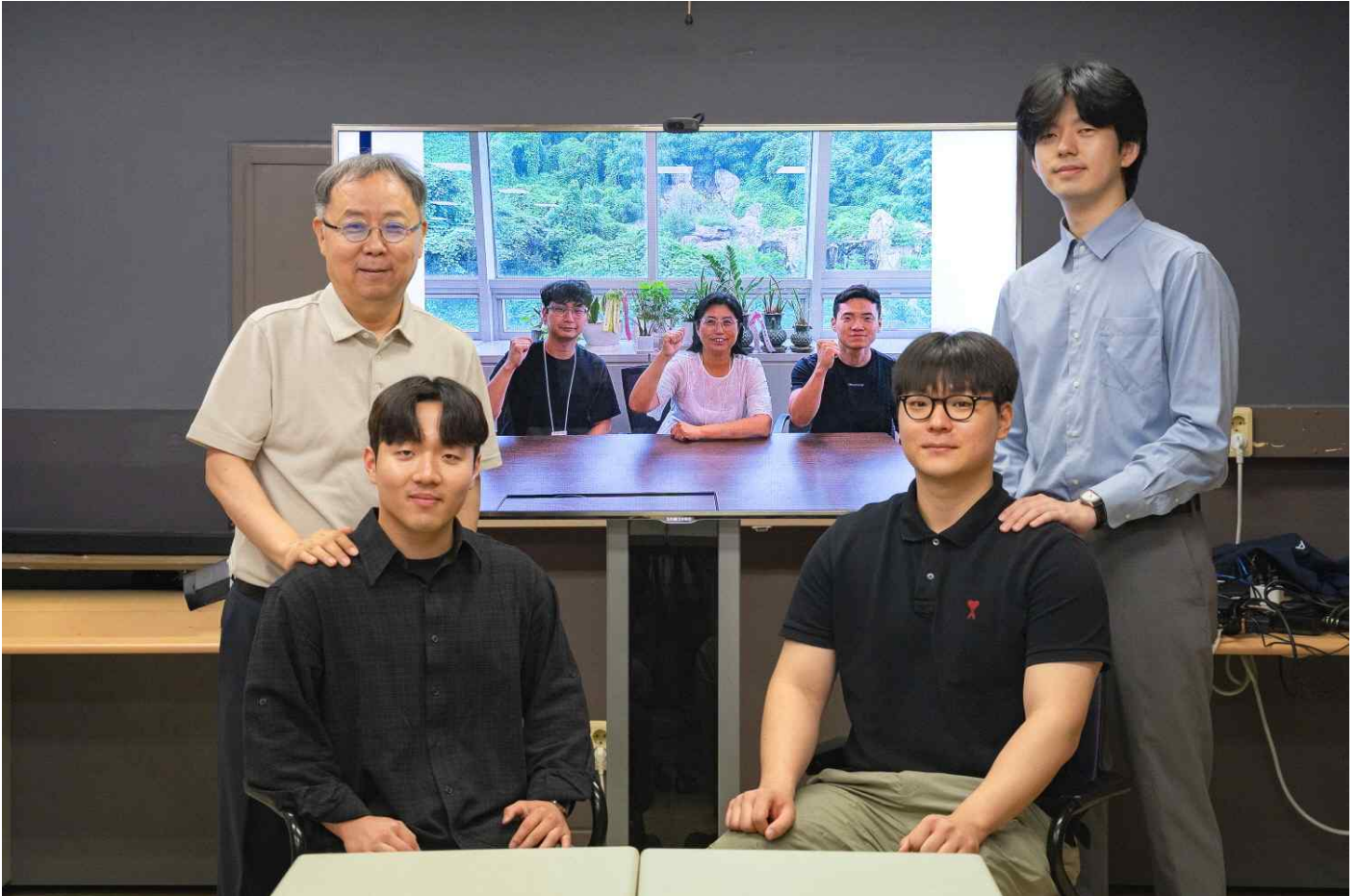


# “새로운 소리 배워도 기존 소리 인식 능력은 그대로”

## GIST-한화비전, 국제 음향 AI 경진대회

### ‘DCASE 2026’ 세계 1위

- 김홍국 교수팀한화비전 R&D센터 공동연구팀 제출 시스템 4개 모두 시스템 랭킹 1~4위 석권
- 새로운 환경에서도 기존 지식 유지하는 ‘음향 연속학습 기술’ 개발, AI의 고질적 한계 ‘치명적 망각’ 극복... 지능형 CCTV·AI 영상보안 등 차세대 보안 솔루션 적용 기대



▲ GIST-한화비전 공동연구팀 단체사진. (뒷줄 왼쪽) 전기전자컴퓨터공학과 김홍국 교수, (뒷줄 오른쪽) 박종연 석박통합과정생, (앞줄 왼쪽) 임도현 석사과정생, (앞줄 오른쪽) 박상원 석사과정생, (화면 속 왼쪽부터) 한화비전 R&D센터 AI연구소 고경득 책임연구원, 임정은 AI연구소장, 금형철 선임연구원

광주과학기술원(GIST·지스트, 총장 임기철)은 전기전자컴퓨터공학과 김홍국 교수 연구팀과 한화비전 R&D센터 AI연구소 공동연구팀이 세계적 권위의 음향 인공지능(AI) 경진대회 ‘디케이스 2026 챌린지(DCASE 2026 Challenge)’에서 공식 팀 랭킹(Teams Ranking) 1위를 차지했다고 밝혔다.

이번 성과는 새로운 환경의 소리를 추가로 학습하면서도 기존에 익힌 소리 정보를 잃지 않는 ‘연속 학습(Continual Learning)’ 기반 음향 AI 기술력을 세계적으로 인정 받은 결과다.

세계전기전자공학회(Institute of Electrical and Electronics Engineers, IEEE)가 주관

하는 ‘DCASE(Detection and Classification of Acoustic Scenes and Events) Challenge’는 음향 환경과 다양한 소리를 분석·인식하는 분야의 대표적인 국제 AI 경진대회로, 전 세계 연구자와 기업 연구팀이 참가해 음향 AI 기술력을 겨루는 대회다.

공동연구팀은 대회 7개 과제(Task) 중 과제 7인 ‘도메인 비의존 오디오 분류를 위한 점진 학습(Domain-Agnostic Incremental Learning for Audio Classification)’부문에 참가했다.

입력 소리가 어떤 환경에서 수집된 것인지 알려주지 않는 ‘도메인 비의존(Domain-Agnostic)’ 조건에서 아기 울음소리, 개 짖는 소리, 화재 경보음 등 10종의 소리를 정확하게 분류해야 하는 과제다.

기존 오디오 인식 AI는 녹음 장비나 장소, 주변 소음 등 환경이 달라지면 성능이 저하되거나, 새로운 데이터를 학습하는 과정에서 이전에 학습한 내용을 잊어버리는 ‘치명적 망각(Catastrophic Forgetting)’ 문제를 겪어와 현장 적용에 한계가 있었다.

공동연구팀은 이러한 문제를 해결하기 위해 새로운 환경을 학습하면서도 기존에 학습한 소리 인식 능력을 유지할 수 있는 ‘연속 학습(Continual Learning)’ 기술을 개발했다.

연구팀은 각 학습 단계에서 학습한 AI 모델을 그대로 보존해 새로운 학습이 기존 지식을 훼손하지 않도록 했다.

특히 원본 데이터를 반복 저장하거나 재학습하지 않고도 과거에 학습한 소리의 특징을 재현해 활용하는 ‘딥인버전(DeepInversion) 기반 생성형 리플레이’ 기법을 적용해, 새로운 환경을 학습하는 과정에서도 기존 지식이 사라지는 문제를 최소화했다.

또한 여러 AI 모델의 예측 결과를 종합해 최종 판단을 내리는 ‘앙상블 기법’을 적용해 하나의 모델만 사용할 때보다 높은 정확도와 안정적인 분류 성능을 구현했다.

그 결과 최종 제출 시스템이 평균 정확도 79.62%를 기록하며 공식 팀 랭킹(Teams Ranking) 1위에 올랐다.

아울러 3개의 단일 시스템과 이를 결합한 앙상블 시스템 등 제출한 4개 시스템이 공식 시스템 랭킹(Systems Ranking) 1위부터 4위까지를 모두 차지하며 기술 완성도를 입증했다.

김홍국 교수는 “이번 성과는 새로운 환경이 추가될 때마다 전체 모델을 다시 학습하지 않고도 기존 지식을 유지할 수 있는 연속 학습 기술의 가능성을 보여준 사례”라며 “AI가 변화하는 환경에서도 효율적으로 학습할 수 있음을 입증했다는 점에서 의미가 있다”고 말했다.

특히 이번 기술은 한화비전의 지능형 영상보안 기술과 결합해 영상과 음향 정보를 동시에 분석하는 차세대 AI 보안 솔루션으로 발전할 수 있을 것으로 기대된다.

한편, 이번 연구는 한화비전 ‘CCTV 및 온디바이스(On-device) 환경을 위한 경량 딥러닝 기반 음원 분류와 다국어 임베딩 및 음성 인식 모델 개발’ 과제와 산업통상자원부 ‘교통약자를 위한 AI 기반 능동형 운전자 어시스턴스 시스템/서비스 개발’ 과제의 지원을 받았다.

연구 성과 발표와 시상이 진행되는 ‘DCASE 2026 워크숍’은 오는 10월 28일부터 29일까지 미국 보스턴에서 개최될 예정이다.