

“로봇이 처음 보는 물체도 한눈에 정확히 인식!”

GIST, 실시간 AI 비전 기술 개발

- AI융합학과 이규빈 교수팀, 오류 추정 통해 미학습 물체 인식 결과를 스스로 정정하는 AI 기술 개발... 오탐지된 물체를 실시간 정밀 보정하는 기술로 AI 비전의 한계 극복
- 로봇 비전·자율주행·산업 자동화 등 다양한 분야에서 활용 기대... 로봇 분야 세계 최고 권위 IEEE 국제 로봇 자동화 학술대회(ICRA)에서 오는 5월 발표 예정



▲ (뒷줄 왼쪽부터 시계방향으로) AI융합학과 이규빈 교수, 맹제모 박사과정생, 신성호 현대자동차 로보틱스랩 책임연구원, 백승혁 한국기계연구원 선임연구원, 이주순·이상범 석박사통합과정(박사과정생), 김강민 박사과정생

로봇이 새로운 환경에서 원활하게 작업하려면 처음 보는 물체도 빠르고 정확하게 인식하는 능력이 필수이다. 하지만 기존 AI 비전 기술은 사전에 학습한 데이터에 포함된 물체만 인식할 수 있어, 미학습 물체에 대한 인식률이 떨어지는 한계가 있었다.

국내 연구진이 이러한 문제를 해결할 혁신적인 인공지능(AI) 기술을 개발해 로봇 비전, 자율주행 등 다양한 분야에서의 활용이 기대되고 있다.

광주과학기술원(GIST, 총장 임기철)은 AI융합학과 이규빈 교수 연구팀이 **오류 추정을 통해 미학습 물체*의 인식 결과를 정제하는 AI 기술을 개발했다고 밝혔다.**

이 기술은 실시간으로 오탐지된 물체를 삭제하거나 추가하는 기능을 갖추고 있어, 로봇의 시각 인지 능력을 한층 향상시킬 것으로 보인다.

* **미학습 물체**: 학습 과정에 활용되지 않은 물체(학습 데이터에 포함되지 않은 물체)

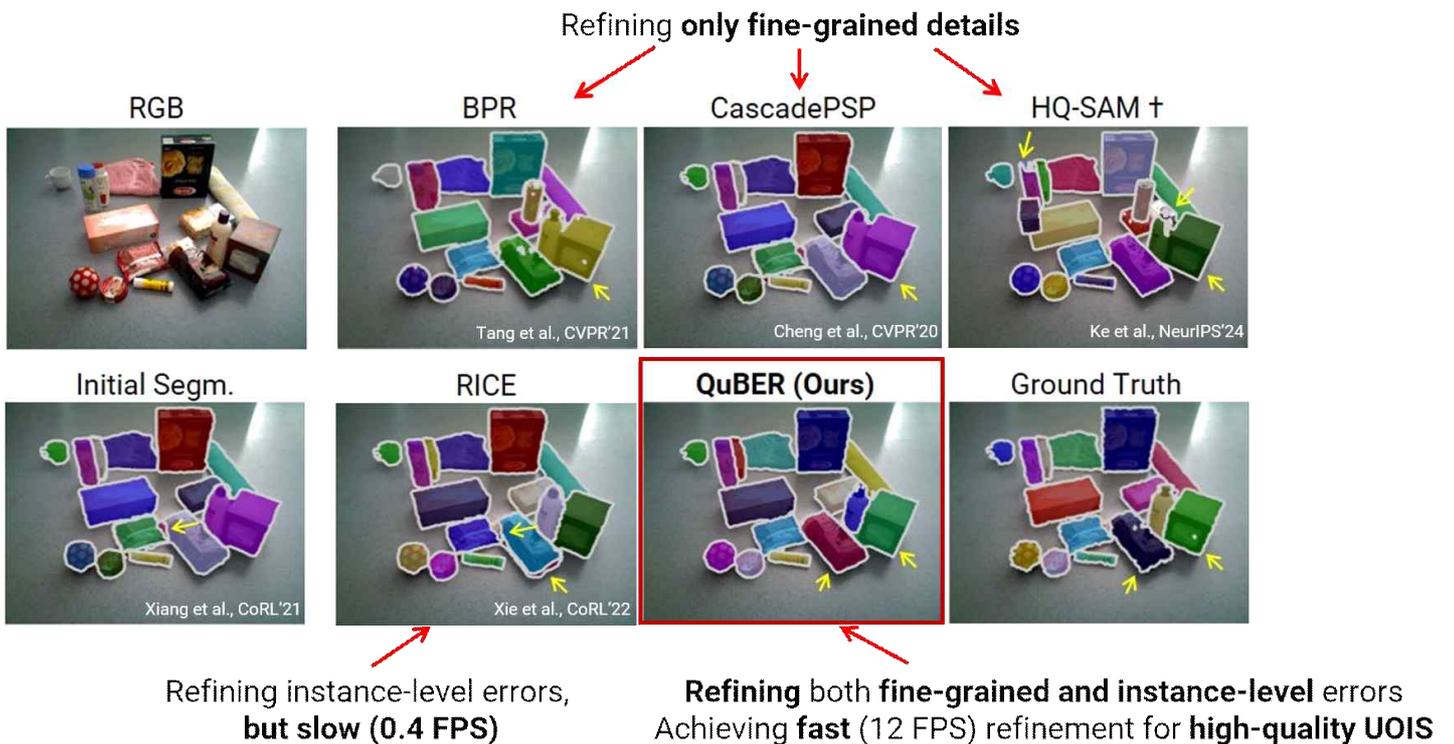
현재 널리 사용되는 AI 비전 기술에는 몇 가지 한계가 있다. Mask R-CNN(2018)* 모델은 사전에 학습된 물체만 인식할 수 있어 새로운 물체를 구별하는 능력이 부족하다. Segment Anything(2023)* 모델은 사용자의 입력(클릭, 박스, 텍스트 등)을 기반으로 임의의 물체를 탐지할 수 있지만, **사람의 도움이 없을 경우 복잡한 환경에서는 정확도가 떨어지는 문제가 있다.**

이를 해결하기 위해 이미지와 초기 예측 데이터를 기반으로 오류를 정정하는 기술이 개발돼 왔다. 하지만 기존 기술은 세부적인 부분만 수정하거나, 잘못된 인식을 정정할 수 있어도 작동 속도가 느리다는 단점이 있었다.

* **Mask R-CNN(He et al. TPAMI'18):** 객체 분할(instance segmentation) 작업에 가장 널리 사용되는 모델로, 물체 영역 후보군(Region of Interest)을 예측하고 해당 영역을 바탕으로 분류(Classification), 바운딩 박스 예측(Bounding Box Regression), 분할 영역 예측(Segmentation Mask Prediction)을 수행하는 구조를 가지고 있다. 많은 객체 분할 모델들이 본 논문을 기반으로 하여 개발되고 있다.

* **Segment Anything(Kirillov et al. ICCV'23):** 컴퓨터 비전 분야의 파운데이션 모델(e.g. 자연어 분야의 GPT)의 필요로 인해 개발된 모델로 수백만개의 대규모 데이터셋으로 학습되었으며 이를 통해 다양한 사용자의 입력(클릭, 박스, 텍스트, 등)에 따라 실시간으로 임의의 객체를 분할할 수 있는 모델이다.

연구팀은 이러한 문제를 해결하기 위해 **빠르고 정확한 오류 정정 기술을 적용한 'QuBER' 모델을 개발했다.**



▲ **개발 모델인 QuBER와 기존 모델간 비교.** 기존의 미학습 물체 분할 정제 모델들의 경우 물체들의 세세한 부분만 정제할 수 있거나, 객체 수준의 정제가 가능하지만 실행 시간이 너무 오래 걸린다는 한계가 있다. 본 연구팀에서 개발한 모델인 QuBER는 빠른 실행 속도를 가짐과 동시에 물체의 세세한 부분부터 물체를 추가하거나 삭제하는 등의 객체 수준의 정제도 가능하다는 장점을 가지고 있다.

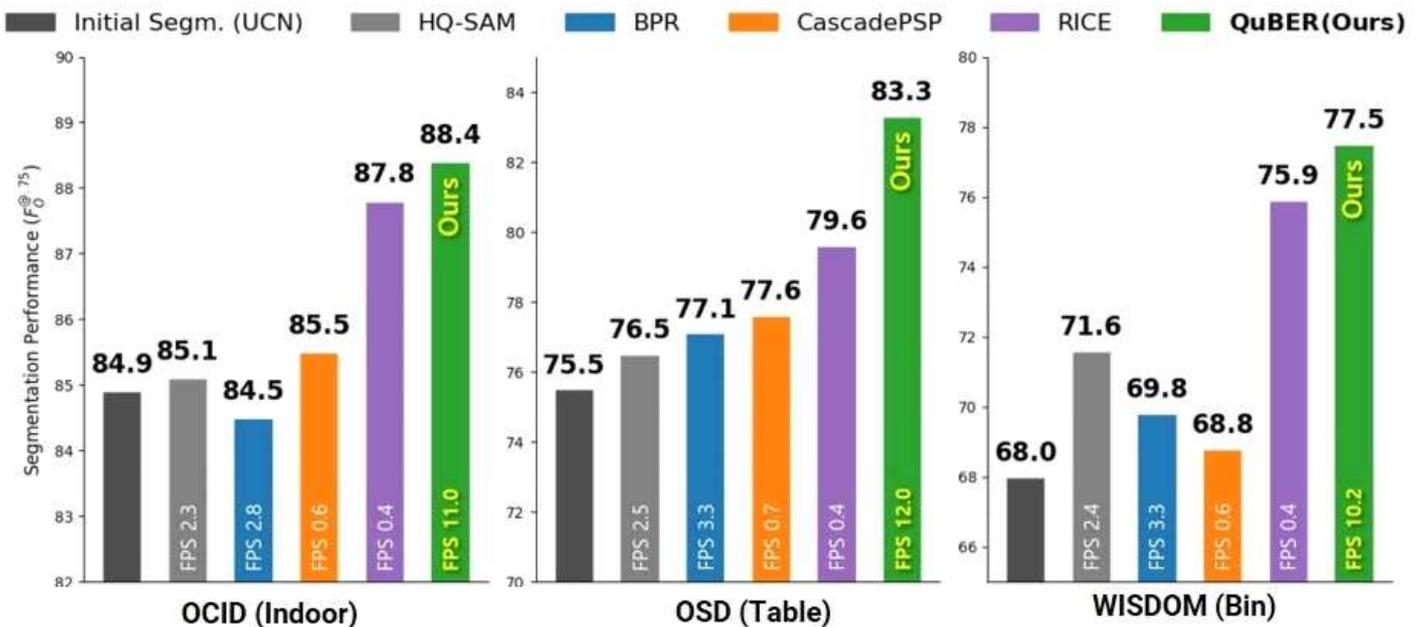
QuBER 모델은 RGB-D(Red·Green·Blue·Depth, 컬러+깊이) 이미지와 초기 예측 데이터를 활용해 '4자 경계 오류(Quadruple Boundary Error)'를 분석함으로써 물체 인식의 정확도를 높이는 것이 특징이다. 이 기술을 통해 처음 보는 미학습 물체도 실시간으로 빠르고 정확하게 인식할 수 있다.

이처럼 4자 경계 오류(Quadruple Boundary Error)는 AI 비전 기술의 정확도를 높이는 데 핵심적인 기법으로 주목받고 있다.

이 기술은 AI의 초기 예측값과 실제 데이터(Ground Truth) 간의 차이를 ▲ True Positive 경계(올바르게 검출된 경계) ▲ False Negative 경계(검출해야 했지만 놓친 경계) ▲ False Positive 경계(잘못 검출한 경계) 그리고 ▲ True Negative 경계(검출하지 않아야 할 부분을 정확히 비검출한 경계) 등 네 가지 경계 기준으로 분석해 물체 인식의 오류를 효율적으로 정정한다.

연구팀이 개발한 QuBER 모델은 가려진 물체가 많아 인식이 어려운 상황에서도 빠르고 정확한 분할 능력을 보이며 세계 최고 수준의 정확도를 기록했다.

▲ 실내 환경(OCID)에서 88.4%의 인식률 ▲ 테이블 위 물체들(OSD)에서 83.3%의 정확도 ▲ 상자 안의 복잡한 물체들(WISDOM)에서 77.5%의 정확도를 보였다.



▲ 미학습 물체 분할 정제 모델 간 성능 비교 그래프. 좌측부터 OCID, OSD, WISDOM 데이터셋에 대해 미학습 물체 분할 정제 모델들 간의 성능을 비교한 그래프이다. 기존 모델들에 비해 QuBER 모델 (Ours)가 더 높은 성능을 보임을 알 수 있다.

이규빈 교수는 "이번 연구를 통해 로봇이 처음 보는 물체도 정확하고 효율적으로 인식할 수 있다는 가능성을 확인했다"며, "이 기술이 다양한 로봇 작업에 적용돼 새로운 환경에서도 안정적으로 동작하는 로봇 개발에 큰 역할을 할 것"이라고 기대감을 나타냈다.

GIST AI 융합학과 이규빈 교수가 지도하고 백승혁 박사가 박사과정 재학 중 수행한 이번 연구는 산업통상자원부와 과학기술정보통신부의 지원을 받았으며, 로봇 분야 세계 최고 권위의 학회인 IEEE 국제 로봇 자동화 학술대회(ICRA, International Conference on Robotics and Automation)에서 2025년 5월 발표될 예정이다.

논문의 주요 정보

1. 논문명, 저자정보

- 저널명 : IEEE International Conference of Robotics and Automation (h5-index: 122 , 2025년 기준, 로봇 분야 최고 학회)
- 논문명 : High-quality Unknown Object Instance Segmentation via Quadruple Boundary Error Refinement
- 저자 정보 : 백승혁(제1저자, 한국기계연구원, GIST 박사 졸업), 이규빈(교신저자, GIST AI융합학과), 이상범 (제2저자, GIST AI융합학과), 김강민(제3저자, GIST AI융합학과), 이주순 (제4저자, GIST AI융합학과), 신성호 (제5저자, 현대자동차 로보틱스랩, GIST 박사 졸업), 맹제모 (제6저자, GIST AI융합학과)