

NAIRL Frontier LETTER

2025 | Vol.01

특집기사

All about NAIRL

뉴럴 스케일링 법칙 초월

AI 효율성과 성능의 한계를 넘어,
새로운 스케일링 패러다임으로

로봇 파운데이션 모델

범용성과 자율성을 갖춘 차세대 AI
로봇의 핵심 기반 기술 개발

초고차원 멀티모달 파운데이션 모델

현실 세계를 정밀하게 반영하는 고차원
생성형 AI 기술의 새로운 지평

NAIRL Insight Talk

최예진

스탠퍼드대학교 교수 | NVIDIA 수석 과학자

“한국도 독자적 언어모델 개발해야”



National AI
Research Lab



ALL ABOUT NAIRL

- 04 · 발간사
- 06 · 격려사
- 08 · 지난해 10월 28일 개소식 개최
대한민국 AI 연구 구심점 '국가AI연구거점' 출범하다
- 10 · ① **세부과제**
뉴럴 스케일링 법칙 초월 연구
AI 효율성과 성능의 한계를 넘어, 새로운 스케일링 패러다임으로
- 12 · ② **세부과제**
로봇 파운데이션 모델 연구
범용성과 자율성을 갖춘 차세대 AI 로봇의 핵심 기반 기술 개발
- 14 · ③ **세부과제**
초고차원 멀티모달 파운데이션 모델 연구
현실 세계를 정밀하게 반영하는 고차원 생성형 AI 기술의 새로운 지평

RESEARCH HIGHLIGHT

- 16 · 성영철 KAIST 전기및전자공학부 교수
가사 도우미 로봇 한층 더 가까이
- 20 · 여진영 연세대 첨단컴퓨팅학부 교수
'행동 전에 생각하는 AI',
월드 모델로 진화하는 웹 에이전트
- 24 · 정든솔 국가AI연구거점 박사후연구원
제로샷 인간-객체 상호작용 탐지를 위한 지역성
및 관계성 주입 네트워크
Locality-Aware Interaction Adapter for
Zero-shot HOI Detection



NAIRL INSIGHT TALK

- 28 · 최예진 교수가 제시하는
AI 스케일링 혁신과
한국의 독자적 AI 개발 전략
"한국도 독자적 언어모델 개발해야"

NAIRL IMPACT

- 32 · 디지털 헬스케어 혁신 기업,
에버엑스EverEx
AI로 재활의 미래를 바꾸다

NAIRL BULLETIN

- 36 · 'AI 석학세미나 및 기술협력
네트워킹 스페셜 토크 시리즈' 개최
- 38 · 국가AI연구거점,
'연구성과 공유 및 참여기업
네트워킹 데이' 개최



NAIRL Frontier letter | 2025 | Vol.01

발간인 국가AI연구거점 김기웅 센터장
발행일 2025년 8월
발행처 서울특별시 서초구 태봉로 108 서울 AI허브 별관 2층
기획 국가AI연구거점 연구전략팀 나현대 연구조교수
제작 (주)동아에스앤씨

발간사



국가AI연구거점(NAIRL)의 대표 소식지, ‘NAIRL Frontier Letter’ 창간호 발행을 진심으로 기쁘게 생각합니다. ‘프론티어(Frontier)’라는 이름에는 기술과 연구의 최전선에서 새로운 지평을 열어가겠다는 저희의 확고한 의지가 담겨 있습니다. 이는 단순한 정보 전달을 넘어, 도전과 혁신, 그리고 미래 기술의 이미지를 담아 국가 R&D의 중심에서 AI의 새로운 역사를 써 내려가겠다는 상징성을 내포합니다. 본 뉴스레터는 국가AI연구거점의 공식 소통 창구로서, 우리 국가AI연구거점의 연구 성과와 세계적 기술 트렌드를 정기적으로 제 공함으로써 산업계, 학계, 연구계 간의 협력 체계를 강화하고 지식 확산을 촉진하는 데 이바지할 것입니다.

국가AI연구거점은 대한민국이 AI 강국으로 도약하기 위한 핵심적인 전략적 허브로 구축되었습니다. 지난 해 7월 개소해 10월 28일 개소식을 거행하였으며, 국내 카이스트, 연세대, 고려대, POSTECH 컨소시엄 소속 45분의 연구진과 해외 우수 대학의 19분의 연구진들이 뉴럴 스케일링 법칙 초월 연구, 로봇 파운데이션 모델 연구, 초고차원 멀티모달 파운데이션 모델 연구를 수행하고 있습니다. 지난 6월 19일에는 주요 연구진들과 참여기업 및 서울시 서초구청 관계자들을 모시고 네트워킹하는 오픈랩 행사를 개최하였습니다.

본 뉴스레터의 ‘NAIRL Research Highlight’ 섹션은 학술적 성과를 선정하여 조명하고, 그 연구적 의의와 산업적 함의를 심층적으로 해설하는 내용을 담습니다. 대한민국이 AI 분야에서 진정한 ‘프론티어’가 되기 위해서는 단순한 점진적 개선을 넘어, ‘큰 과제’에 집중하고 ‘원천 기술’을 확보하는 것이 필수적입니다. 이러한 정신으로 매호 새로운 내용의 원천 연구 성과를 소개하도록 하겠습니다. 또한, ‘NAIRL Insight Talk’을 통해 국내외 전문가들의 귀중한 통찰을 공유하고, ‘NAIRL Impact’ 섹션에서는 실제 현장에서의 AI 활용 사례를 소개함으로써, 뉴스레터가 AI 3대 강국 비전을 달성하기 위한 우리의 ‘산·학·연 원팀’ 체제를 공고히 하는 데 소통의 창구가 되도록 하겠습니다.

독자 여러분의 적극적인 참여와 관심은 AI 혁신의 여정에서 가장 큰 동력이 될 것입니다. 미래를 향한 공동의 혁신과 협력, 그리고 국가적 발전을 기대하며, 모든 분께 깊은 감사의 말씀을 전합니다.

국가AI연구거점 센터장 김기웅

격려사



국가AI연구거점의 'NAIRL Frontier Letter' 창간호 발행을 진심으로 축하드립니다.

우리는 인공지능(AI)이 우리 경제·사회 전반에 걸쳐 혁신을 선도하고, 패러다임의 변화를 이끌어가고 있는 격변의 시기에 살고 있습니다. AI 기술이 이례적인 속도로 빠르게 발전하며, AI 연구자는 물론 이해 관계자들 간의 정보 공유와 협력의 중요성이 그 어느 때보다 커지고 있는 상황에서, 국가AI연구거점이 새로운 소통의 창구를 마련하게 된 것은 매우 뜻깊고 시의적절한 일이라고 생각합니다.

'NAIRL Frontier Letter'는 단순한 뉴스레터를 넘어, AI 분야의 최전선에서 이뤄지는 혁신적인 AI 연구 성과와 미래 비전 등을 폭넓게 공유하는 지식의 플랫폼이 될 것으로 기대합니다. 국내외 최고 수준의 연구진들이 축적해 온 귀중한 통찰과 경험이 이 지면을 통해 더 넓은 연구 생태계로 확산되어, 우리나라 AI 연구의 질적 도약과 글로벌 경쟁력 강화에 크게 기여하리라 확신합니다. 또한, 국내 AI 연구진들은 물론 산업계, 정책 입안자들에게도 최신 동향과 핵심 이슈들을 효과적으로 전달하는 가교 역할을 할 것으로 확신합니다.

아울러, 인공지능이 빠른 발전과 함께 사회적 가치와 책임, 그리고 신뢰라는 새로운 과제를 함께 안고 있는 지금, 연구자와 정책 담당자, 산업계가 긴밀히 소통하며 신뢰 기반의 발전 방향을 모색하는 것이 중요합니다. 'NAIRL Frontier Letter'는 단순히 기술의 진보를 알리는 것에 그치지 않고, 지향해야 할 바람직한 AI 활용과 윤리적 기준 등을 함께 논의하는 장이 되기를 기대합니다.

앞으로 이 뉴스레터가 AI 연구의 새로운 이정표가 되어, 대한민국이 글로벌 AI 강국으로 도약하는 데 중요한 역할을 해나갈 것으로 믿습니다. 과학기술정보통신부는 국가AI연구거점과 함께 AI 혁신 생태계 조성과 확장을 위해 적극 노력하겠습니다. 국가AI연구거점의 지속적인 발전을 기원하며, 'NAIRL Frontier Letter'의 성공적 출발을 다시 한번 축하드립니다.

과학기술정보통신부 정보통신정책실장 **송상훈**

지난해 10월 28일 개소식 개최

대한민국 AI 연구 구심점 ‘국가AI연구거점’ 출범하다



대한민국을 대표하는 인공지능(AI) 연구 구심점인 국가 AI연구거점이 출범했다. 지난해 10월 28일 과학기술정보통신부와 정보통신기획평가원이 서울 서초구 서울AI허브에서 국가AI연구거점 개소식을 개최했다. 오세훈 서울시장, 이광형 KAIST 총장 등이 참석한 가운데 국가AI연구거점의 성공적 출범을 축하하며 대한민국 인공지능 세계 3대 강국 도약을 위한 의지를 다졌다.

국가AI연구거점의 주관기관인 KAIST의 이광형 총장은 “이곳에서 국내외 인공지능 연구자들이 교류하며 창의적 인공지능 연구를 펼치길 바란다”고 환영사로 밝혔다. 이어 국가AI연구거점이 자리한 서울시의 오세훈 시장은 축하사에서 “국가AI연구거점에 기반하여 서울시가 세계적인 인공지능 연구자들이 모여드는 국제 인공지능 중심지로 성장할 수 있도록 전폭 지원하겠다”고 강조했다.

또한 국가AI연구거점의 책임자인 김기응 KAIST 교수가 국가AI연구거점의 운영계획을 제시했다. 국가AI연구거점

에는 KAIST, 고려대, 연세대, POSTECH 컨소시엄을 중심으로 국내외 AI 연구진이 참여해 파괴적 혁신을 지향하는 뉴럴 스케일링 법칙 초월 연구, 로봇 파운데이션 모델 연구 등을 진행한다. 2028년까지 정부가 440억 원, 지자체가 506억 원을 투입해 총 946억 원 규모로 운영될 예정이다.

서울AI허브에 자리한 국가AI연구거점은 국내외 우수 연구진이 역동적으로 교류하며 세계적 인공지능 국제공동 연구를 수행하고, 국제 인공지능 지도자 양성 기능과 인공지능 산학연 생태계를 집약하는 온라인 체제 기반(플랫폼) 역할을 한다. 아울러 국가AI연구거점을 통해 미국, 캐나다, 프랑스, 아랍에미리트 등에 있는 해외 우수기관의 연구자들도 일정 기간 국내에 상주하여, 도전적인 인공지능 국제공동연구를 수행하며, 글로벌 포럼 개최 등을 통해 국제 협력 관계망도 구축하고 지속적으로 확장해나갈 계획이다. [NAIRL](#)





① 세부과제

뉴럴 스케일링 법칙 초월 연구

AI 효율성과 성능의 한계를 넘어, 새로운 스케일링 패러다임으로

인공지능 모델의 성능 향상은 일반적으로 더 큰 모델, 더 많은 데이터, 더 높은 연산 자원을 요구한다. 이는 곧 ‘뉴럴 스케일링 법칙[Neural Scaling Law]’으로 불리는 법칙으로 설명되며, 모델의 성능은 규모에 따라 일정한 곡선으로 향상되지만, 그에 따른 비용과 자원 소모 또한 기하급수적으로 증가하는 한계를 갖고 있다. 국가 AI 연구거점이 수행하는 이 연구는 바로 이러한 한계를 극복하고, 자원 효율성과 성능을 동시에 확보하는 지속가능한 인공지능 기술 개발을 목표로 하고 있다.

KAIST, 고려대, 연세대, POSTECH 등이 참여하는 본 과제는 크게 네 개의 세부 분야에서 진행되고 있다. 먼저 기존 대형 모델의 연산량을 줄이고도 고성능을 유지할 수 있도록, 한국어 특화 언어모델의 경량화와 지식 증류(distillation) 기술을 개발하고 있다. 특히 프롬프트 증류, 무학습 기반의 다기준 토큰 병합, 이중 비트 양자화 기법을 활용하여 모델 추론 과정에서의 연산 자원을 최소화하면서도 정보 손실을 줄이는 방법을 연구하고 있다. 또한 확산 모델과 같은 생성형 모델이 학습한 내부 지식을 분석해 이를 보다 작고 효율적인 모델로 전이시키는 기술도 함께 개발 중이다.



1세부의 책임자인 양은호 KAIST 교수.

두 번째로, 새로운 학습 알고리즘을 통해 스케일링 법칙의 제약을 완화하고자 한다. L2O(Learning to Optimize) 기반의 모델 업데이트 방식과 강화학습을 활용한 연산 자원 분배 기법, 그리고 학습 데이터의 정보량을 기반으로 한 데이터 정제 및 압축 기술을 통해, 학습 속도를 향상시키고 불필요한 데이터나 연산을 줄이려는 시도를 이어가고 있다. 특히 다중 에이전트 기반의 협력 알고리즘, 고해상도 생성 모델에 적용 가능한 차등 프라이버시 기술 등은 향후 실제 응용 단계에서도 큰 파급력을 가질 것으로 기대된다.

세 번째로, 뉴럴 스케일링의 구조적 한계를 극복하기 위한 새로운 모델 아키텍처 개발에 집중하고 있다. 그래프 신호처리를 기반으로 셀프 어텐션(Self-Attention) 구조를 개선하거나, 연산 복잡도가 $O(N^2)$ 인 기존 트랜스포머 구조를 $O(N)$ 수준으로 줄이는 효율적인 멀티모달 아키텍처를 설계함으로써, 연산량은 줄이면서도 긴 시퀀스 입력에 대한 성능을 유지할 수 있는 모델을 구현하고자 한다. 또한 객체 중심의 표현 학습을 통한 비지도 의미 분할 기술과, 추론 효율을 높이는 추측적 디코딩(speculative decoding) 전략도 개발 중이다.

끝으로 효율적인 기초 데이터 세트를 구축해 뉴럴 스케일링 법칙을 넘어서는 기반을 제공하고자 한다. 정보량 분석에 기반한 데이터 증류, 가상 영상 생성 및 웹 합성 데이터 세트 구축, 고효율 영상 데이터 압축 등의 기술을 이용해 고품질의 초소형 데이터 세트를 제작하고, 이를 통해 경량 모델이 성능 저하 없이 훈련될 수 있도록 돕는다. 특히 한국어 데이터 세트의 다변화와 평가 벤치마크 구축은 국내 AI 생태계의 자립성을 높이는 데도 기여하고 있다.

이러한 통합적 접근을 통해 본 연구는 산업 전반에서 활용 가능한 실시간 경량 AI 시스템 개발은 물론, 지속가능하고 친환경적인 AI 기술 발전이라는 큰 방향성을 실현하고자 한다. 기존의 한계를 뛰어넘는 새로운 패러다임의 인공지능, 그것이 바로 뉴럴 스케일링 법칙 초월 연구가 향하는 지점이다.

1세부를 총괄하는 양은호 KAIST 교수는 다음과 같은 포부를 밝힌다. “AI 연구에 필요한 투자금이 점점 커지고, 자본의 격차가 기술 발전의 비대칭을 심화시키는 현실에서, 비용 효율적인 AI 연구는 지속 가능한 AI 발전을 위해 반드시 필요하다고 생각합니다. 특히, AI 자본 경쟁에서 뒤처질 수밖에 없는 우리나라 환경에서, 우리 연구팀이 효율적인 연구 전략을 통해 국내 AI 경쟁력 제고에 실질적으로 기여하길 기대합니다. 한정된 자원 속에서도 창의적이고 혁신적인 접근으로 대한민국 AI 연구의 새로운 가능성을 열어가겠습니다.” **NAIRL**



② 세부과제

로봇 파운데이션 모델 연구

범용성과 자율성을 갖춘 차세대 AI 로봇의 핵심 기반 기술 개발

로봇 기술은 이제 개별 작업 수행을 넘어, 변화하는 현실 환경에 능동적으로 대응하고 스스로 판단하며 학습하는 방향으로 진화하고 있다. 본 연구는 이러한 흐름에 맞춰, 다양한 환경에서 인지·의사결정·제어를 통합적으로 수행할 수 있는 범용 파운데이션 모델을 구축하고자 한다. 목표는 기존의 단일 과제 특화형 모델을 넘어, 스스로 사고하고 적응하며 협업할 수 있는 차세대 인공지능 로봇의 기반 기술을 마련하는 것이다.

연구 1단계에서는 멀티모달 기반 인식, 범용 로봇 제어, 체화된 인공지능 등 세 가지 축을 중심으로 파운데이션 모델의 기초 역량을 확보하는 데 집중하고 있다. 개방어휘 물체 탐지 및 분할, 음성과 소리 정보를 통합하는 양방향 이해 모델, 그리고 다양한 센서 정보를 융합하는 인터페이스 기술을 개발하여 복잡한 현실 환경 속에서도 안정적인 인지 기능을 구현한다. 동시에 로봇의 동작을 생성하고 재구성하는 기술과 강화학습 기반 의사결정 모델을 통해, 범용적이고 적응 가능한 제어 능력을 확보하고 있다. 또한, 로봇이 상식과 윤리, 도메인 지식을 바탕으로 상황을 해석하고 장기적으로 학습할 수 있도록, 세계상(world model) 기반의 체화형 AI 기술도 함께 개발하고 있다.

2단계에서는 이러한 핵심 기술들을 실세계 환경에 적용하고, 성능과 효율성을 동시에 고도화하는 것을 목표로 한다. 특히 실제 환경에서의 센서 정보 처리, 실시간 의사결정, 에너지 효율성과 연산 자원 최소화를 고려한 모델 양자화, 그리고 로봇 행동의 정밀 제어 등 다양한 요소를 융합해 실행 가능한 고도화된 로봇 파운데이션 모델을 구현할 예정이다. 체화된 인공지능 측면에서는 상식·윤리 추론의 현실 적용, 장기기억 처리, 도메인 지식 통합 및 세계상 학습 기술의 진화를 통해, 인간과 자연스럽게 상호작용하며 학습할 수 있는 로봇의 토대를 다진다.

본 과제는 POSTECH, 연세대, 고려대, KAIST 등이 참여하고 있으며, 각각 로봇의 다양한 모션 제어, 체화형 AI, 멀티모달 센서 기반 인지 등 각기 다른 전문영역에서 상호보완적 역할을 수행하고 있다. 특히, 프롬프트 기반의 멀티모달 이해와 로봇 정책 학습 간의 연계를 위한 대규모 시뮬레이션 환경과 벤치마크 세트도 구축하는 중이다.

이 연구는 향후 자동화 공정, 스마트 물류, 재난 대응,

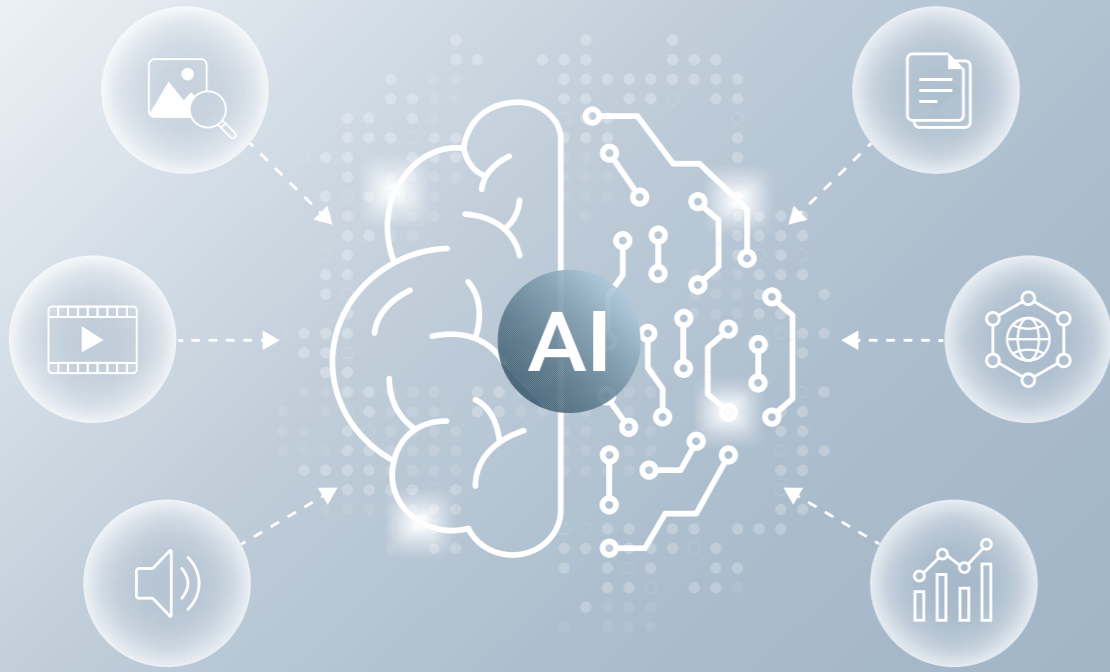
공공 안전, 스마트홈 등 다양한 응용 분야에서 지능형 로봇의 실질적인 활용을 가능케 할 것으로 기대된다. 특히 인간 중심의 윤리적 사고와 적응적 학습을 겸비한 로봇의 구현은 로봇이 단순 기계적 도구를 넘어 협력자(co-agent)로서 자리 잡는 데 핵심적인 역할을 하게 될 것이다.

2세부 책임자인 조민수 POSTECH 교수 다음과 같은 포부를 밝힌다. “본 세부과목으로 하는 로봇 파운데이션 모델은 가상세계에 머물렀던 기존의 인공지능 기술을 현실로 가져오는 차세대 인공지능 기술의 핵심 프로젝트입니다. 인공지능 기술이 대한민국 미래 세대를 좌우할 것으로 기대되는 이 시기에 국가시거점의 핵심연구를 맡게 된 것에 큰 책임을 느끼면서도, 인공지능 분야에서 이미 세계적인 성과를 보이고 있는 한국의 연구진들과 연구거점에 참여하고 있는 해외석학들이 있기에 기대도 큼니다. 2세부 연구를 통해 대한민국이 로봇 인공지능 기술의 핵심거점으로 발돋움할 수 있도록 최선을 다하겠습니다.”

NAIRL



2세부의 책임자인 조민수 POSTECH 교수.



③ 세부과제

초고차원 멀티모달 파운데이션 모델 연구

현실 세계를 정밀하게 반영하는 고차원 생성형 AI 기술의 새로운 지평

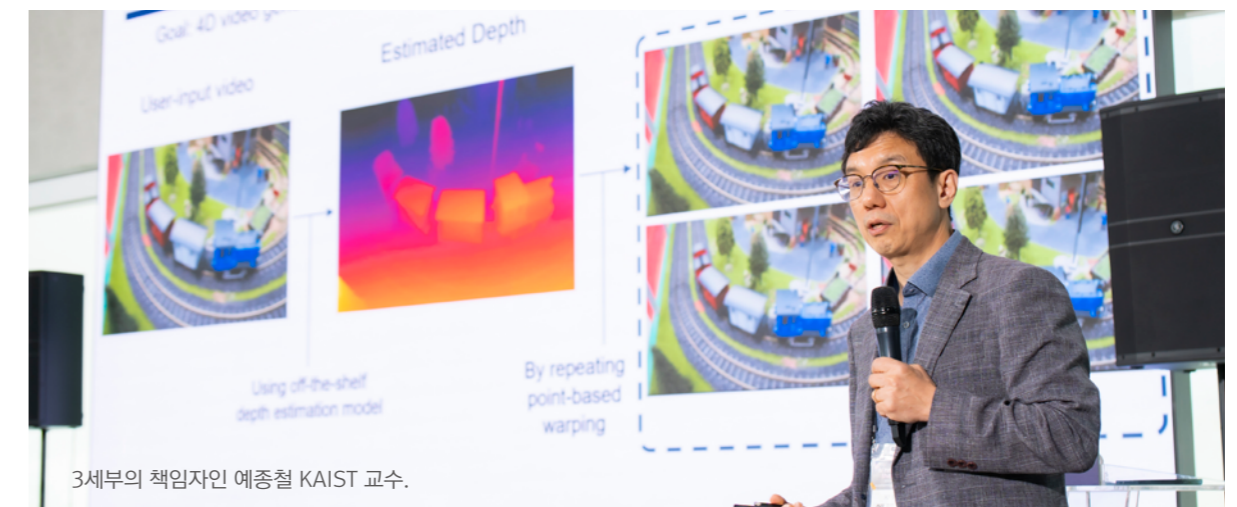
생성형 인공지능이 고도화될수록 현실 세계의 물리 법칙, 다양한 센서 정보, 개인 프라이버시, 고차원 시계열 및 3D·4D 데이터까지 포괄적으로 이해하고 반영할 수 있는 모델이 요구되고 있다. 3세부과제는 이러한 흐름 속에서, 현실 세계와 정밀하게 상호작용 가능한 고차원 멀티모달 파운데이션 모델을 구축하는 것을 목표로 한다.

이 연구는 영상, 의료, 분자 구조, 시계열 데이터, 프라이버시 보호 등 다양한 도메인을 아우르는 초고차원 생성형 AI 기술의 핵심 기반을 마련하고자 한다. 특히 다양한 유형의 현실 센서 데이터와 물리적 환경, 인간의 선호와 윤리까지 반영하는 통합적 접근이 강조된다.

KAIST는 멀티모달 이미지-텍스트 데이터 세트를 활용해 디퓨전 트랜스포머^{DIT, Diffusion Transformer} 기반의 텍스트-영상 생성 모델과 4D 생성 모델을 개발하고 있다. 텍스트 입력의 품질을 높이기 위해 프롬프트 리캡셔닝^{prompt recaptioning} 기법을 적용하고 있으며, 비전-언어 모델과 연동해 텍스트, 비디오, 의료 데이터를 생성하는 연구도 병행하고 있다. 또한 실사형 3D 모션 예측 기술과 고자기장 MRI 데이터를 활용한 의료 영상 생성 모델도 동시에 개발하고 있다.

POSTECH은 미시 세계의 물리 법칙을 반영한 3차원 분자 구조 생성 모델을 개발하고 있다. 물리적으로 안정적인 분자 구조를 생성하기 위해 외부 시뮬레이터와 상호작용하는 대화형 학습 구조를 도입했으며, 강화학습 기반의 커리큘럼 학습 기법도 적용해 모델의 정확성과 안전성을 높이고 있다.

연세대는 사용자의 선호와 현실 기반 요구를 반영한 강화학습 기반 최적화 기법(RLHF, DPO 등)을 통해 실제 활용도 높은 생성형 AI 모델을 개발하고 있다. 다수 시점에서의 이미지 정보를 통합해 정제되지 않은 영상으로부터 정밀한 3D·4D 재구성을 수행하며, 멀티오브젝트 상황에서도 선호 기반 판단이 가능한 아키텍처 설계에 주력하고 있다.



3세부의 책임자인 예종철 KAIST 교수.

고려대는 시계열 생성에 특화된 LLM 기반 구조를 개발하고 있으며, 시간적 역학 특성을 반영한 데이터 토큰화, 시계열 예측, 다변량 생성 등에서 도메인 법칙을 효과적으로 적용하고 있다. 아울러, 회전·이동 불변성을 반영한 SE(3) 등가 신경망, 해밀토니안 기반 물리 특성 예측 등도 병행하여 고정밀 과학 데이터를 생성할 수 있는 모델을 구현하는 중이다.

한편, 실세계에 적용 가능한 프라이버시 보호 기술도 이 연구의 핵심 영역이다. LoRA^{Low-Rank Adaptation} 기반의 실시간 개념 제거, 메타러닝 기반의 원샷^{one-shot} 프라이버시 표

현 제거 기술, 사용자 요청 기반 표현 제거 등을 통해, 프라이버시 친화적 생성형 AI를 구현할 계획이다.

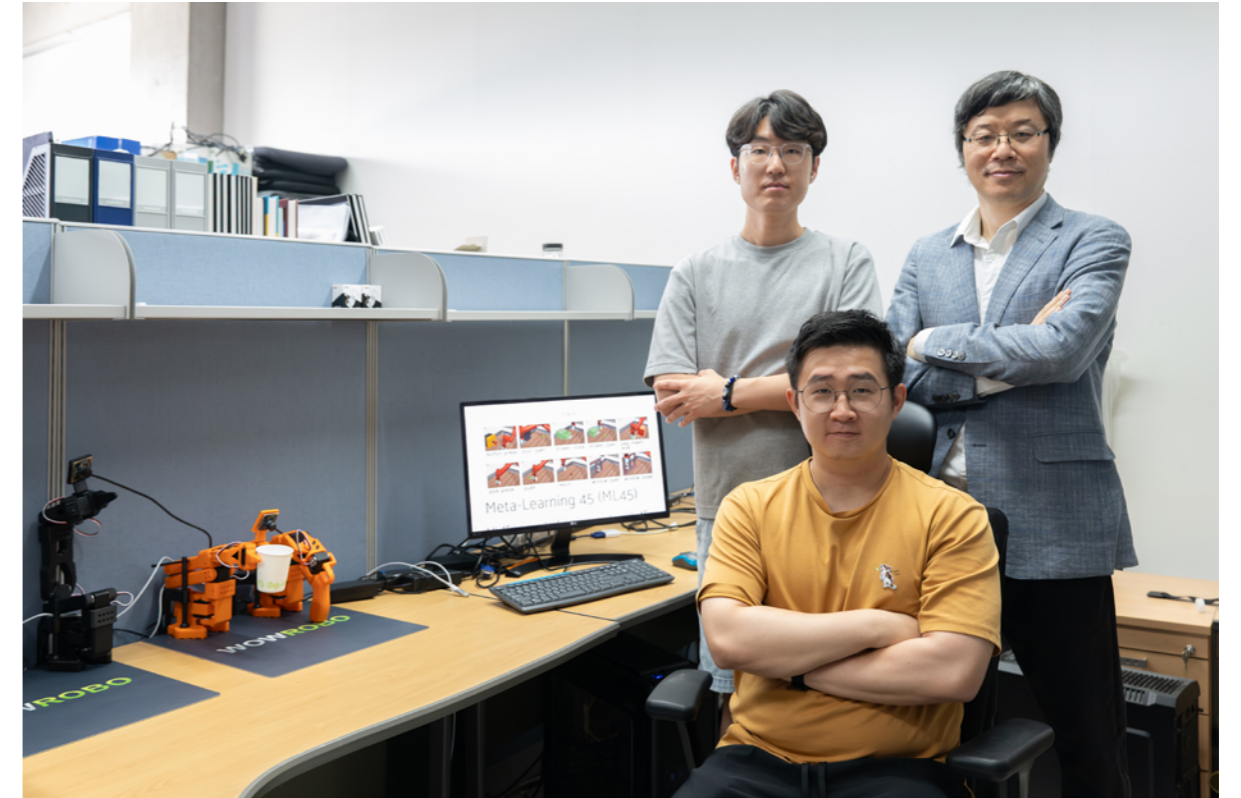
또한, 시간과 공간을 모두 반영하는 4D 생성 파운데이션 모델 개발도 병행하고 있다. 이 모델은 기존의 2D/3D 생성 기술을 넘어 시간 변화와 카메라 시점 이동까지 고려해 고차원적 데이터 생성을 가능케 한다. 더 나아가 단순한 형상 복원을 넘어 의미론적 정보까지 포함할 수 있는 예측 가능한 표현 설계로 확장되고 있다.

이 연구는 향후 정밀 의료, 신약 설계, 로봇틱스, 스마트 제조, 보안 친화적 AI 서비스 등 다양한 분야에 응용될 수 있다. 현실 세계의 복잡한 구조와 윤리적 요구까지 반영할 수 있는 고성능 파운데이션 모델 개발은 생성형 AI의 신뢰성과 적용 범위를 획기적으로 넓히는 기반이 될 것이다.

3세부 연구책임자인 예종철 KAIST 교수는 다음과 같은 포부를 밝힌다. “본 연구팀은 영상, 바이오, 도시 분야를 중심으로 물리학적 원리를 바탕으로 한 고차원 생성·추론 기술을 개발하고 있습니다. 모노컬러 영상의 2D 트랙을 활용한 상대 깊이 추론, BFN+DPO 기반의 다중 물성 분자 생성, 3D 의미 지도를 통한 보행자 행동 생성, 그리고 최대 256배까지 확대 가능한 초해상도 Chain-of-Zoom 기법 등을 통해 다양한 현실 문제에 접근하고 있으며, 디지털 트윈, 신약 개발, 도시계획, 고해상도 콘텐츠 생성 등 광범위한 응용 가능성을 제시합니다.” **NAIRL**

성영철

KAIST 전기및전자공학부 교수

가사 도우미 로봇
한층 더 가까이

최근 성영철 교수 연구팀은 다중 태스크 강화학습에서 나타나는 부정적 전이와 난이도 편향 문제를 해결하는 데 성공했다. 특히 다중 태스크 강화학습의 대표적 테스트 벤치마크인 Meta-World MT10/MT50 벤치마크에 적용해 뛰어난 성과를 거두었다.

로봇이 생각 능력과 운동 능력을 갖추려면?

일상 생활에서 집안일을 도와주는 가사 도우미 로봇은 어떻게 만들 수 있을까? 우선은 사람과 소통해야 하니, 보고 들을 수 있어야 할 것이다. 시각으로 본 이미지에서 각 물체를 찾아내고 무엇인지 알아맞춰야 하는데, 이것은 컴퓨터 비전^{computer vision}의 영역으로 이미 상당히 발전해 있다. 그리고 우리가 언어로 말했을 때 그 음성 신호를 의미의 텍스트로 바꿔야 한다. 이것은 음성인식 및 음성·문자 변환^{text-to-speech/speech-to-text}의 문제인데, 관련 기술 또한 이미 상당히 발전해 있다.

그렇다면 남은 것들은 무엇일까? 로봇이 우리가 시킨 일이 무엇인지 알아듣고 스스로 집안 환경 속에서 행동을 취하며 그 일을 달성할 수 있는 ‘생각 능력’이 필요하고, 또한 이 생각으로부터 나오는 행동들을 자신의 팔과 다리를 움직이면서 실행해야 하는 ‘운동 능력’이 필요하다. 전자에 관해서는 현재 거대언어모델을 이용하여 스스로 생각하며

계획하는 AI 에이전트가 그 방법론이 될 수 있으며, 로봇 팔다리를 움직이는 운동 능력에 관한 후자의 연구는 소위 강화학습의 연구 영역이다.

지금까지 중국의 유니트리^{unitree} 로봇이나 테슬라의 아틀라스를 보면, 주로 걷고 뛰고 중심 잡는 수준에 머물러 있음을 알 수 있다. 하지만, 현실에 적용되는 휴머노이드 로봇이 어떤 일을 수행하려면 단순히 걷고 뛰는 수준을 넘어서 팔과 손으로 다양한 동작을 할 수 있어야 한다. 인간처럼 손으로 물체를 잡고 놓아야 하며, 문도 열고 닫아야 하고, 서랍도 열고 닫을 수 있으며, 물건을 밀고 당길 수도 있어야 한다. 이러한 강화학습의 영역이 바로 다중 태스크 강화학습인데, 이는 어려운 연구 문제로 현재 활발한 연구가 이루어지고 있다.

다중 태스크 강화학습의 문제를 해결하다

다중 태스크 강화학습은 하나의 알고리즘으로 다양한



태스크를 모두 수행할 수 있도록 하는 것이다. 즉 로봇이 다수의 동작 과제를 하나의 정책(policy)으로 통합 수행할 수 있어야 한다. 다중 태스크 강화학습의 경우 각각의 태스크에서는 잘할 수 있지만, 다양한 태스크를 모두 잘하도록 학습시키면, 부정적 전이(negative transfer)와 난이도 편향이란 2가지 문제가 발생한다. 부정적 전이는 한 태스크의 학습이 다른 태스크의 학습을 방해하는 문제이고, 난이도 편향은 학습이 쉬운 태스크에 편중되어 쉬운 태스크들만 잘하는 문제이다.



기존에 다중 태스크 강화학습 시 나타나는 문제를 해결하기 위해 구배(gradient)를 이용하거나 모듈을 이용하는 등 다양한 시도가 있었지만, 그리 효과적이지 못하였다. 각 태스크에서 얻은 손실의 구배를 조절하거나 조합해서 부정적 전이를 줄이고자 했지만, 각 태스크 간 구배가 상충하거나 방향이 엇갈리면 성능 향상에 한계가 있었다. 또 태스크마다 다른 서브 모듈을 두고 상황에 따라 선택적으로 사용하는 방식에서는 모듈 수가 많아지면 학습과 활용이 복잡해지고, 일반화에도 제약이 따랐다.

이에 KAIST 성영철 교수 연구팀은 2024년 태스크의 난이도를 이용한 스케줄링과 네트워크 리셋 기법을 이용하여, 이런 문제를 해결하고자 노력했다. 특히 쉬운 태스크보다 어려운 태스크를 먼저 스케줄링하여 난이도 편향을 줄이고, 학습 중에 인공지능망의 가중치를 주기적으로 초기화하는 방식을 통해 편향이나 부정적 전이를 없앨 수 있었다. 관련 연구 결과는 인공지능 최대 학회 중 하나인 ICML (International Conference on Machine Learning) 2024에서 발표되어 큰 주목을 받았다.



성영철 교수가 스마트 정보 시스템 연구실의 연구원들과 함께한 모습.

‘적응형 보상 스케일링(ARS)’을 넘어 ‘메타 강화학습’까지

2025년에는 연구팀이 한층 더 발전된 기법인 ‘적응형 보상 스케일링(Adaptive Reward Scaling, ARS)’을 개발해 태스크별 보상값을 과거 히스토리 기반으로 동적으로 조절하며, 네트워크 리셋과 함께 작동해 학습 초기 편향을 억제하고 전체 성능을 고르게 끌어올리는 데 성공하였다. 각 태스크의 보상값을 적응적으로 스케일링하면서 인공지능망 리셋을 병행하여 다중 태스크 강화학습에서 부정적 전이와 난이도 편향 문제를 근원적으로 해결한 셈이다.

연구팀은 이 알고리즘을 로봇 손의 다양한 작업을 시뮬레이션하는, 다중 태스크 강화학습의 대표적 테스트 벤치마크인 Meta-World MT10/MT50 벤치마크에 적용해 뛰어난 성과를 거두었다. 특히 열 개의 손동작으로 구성된 MT10에서 99.5%의 태스크 성공률을 보여, MT10 문제를

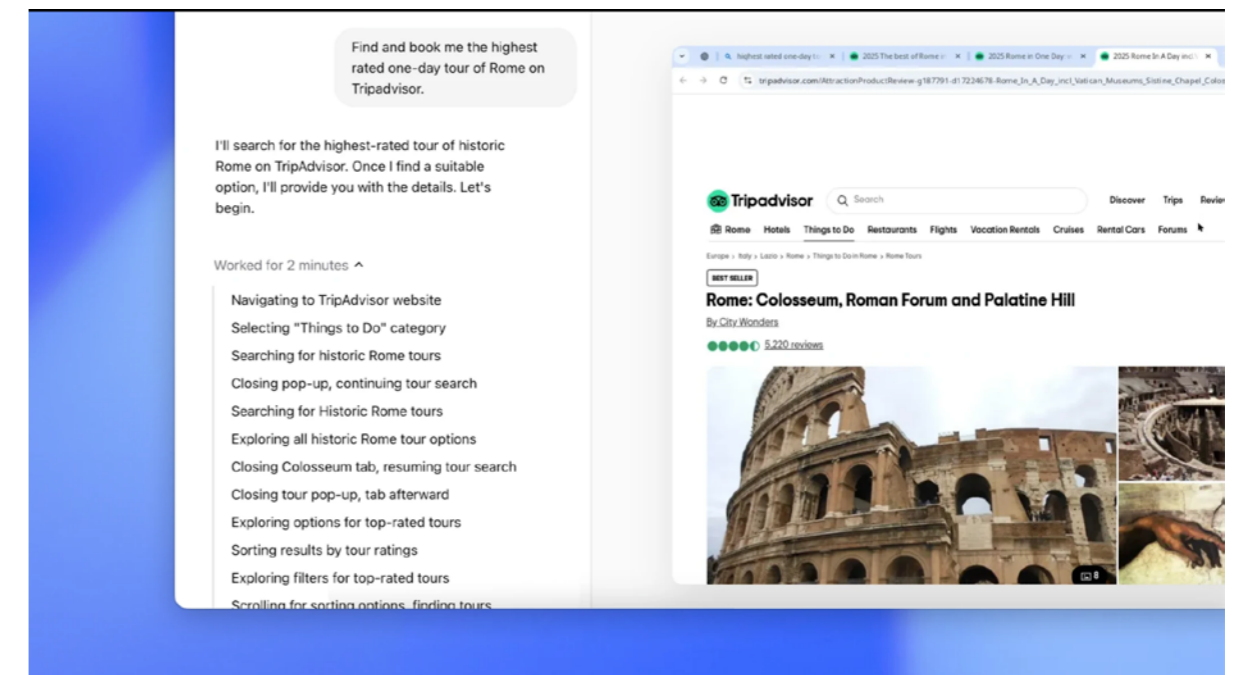
세계 최초로 풀었으며, 이보다 훨씬 많은 50개 손작업으로 구성된 MT50 벤치마크에서 기존 최고 64.2%에서 88.9%로 성능을 24.7%p 향상시켰다. 이 연구 결과는 7월 중순에 열린 ICML 2025에 발표되었다.

앞으로 연구팀은 가사 도우미 로봇을 개발하는 데 필요한 ‘메타 강화학습(Meta Reinforcement Learning)’과 심투리얼(Sim-to-Real)에도 힘을 계획이다. 메타 강화학습은 로봇이 새로운 집안 환경이나 작업을 빠르게 학습할 수 있도록 돕고, 심투리얼은 시뮬레이터에서 훈련한 AI를 실제 현실 환경에서도 안정적으로 작동하게 해준다. 이 두 기술을 결합함으로써, 로봇은 낯선 상황에서도 빠르게 적응하고 사람과 자연스럽게 협업할 수 있는 지능을 갖추게 된다. 가사 도우미 로봇이 가정에 필요한 가전처럼 도입될 날도 머지않았다.

NAIRL

여진영

연세대 첨단컴퓨팅학부 교수

‘행동 전에 생각하는 AI’,
월드 모델로 진화하는 웹 에이전트

OpenAI가 최근 공개한 ‘Operator’ 에이전트 인터페이스. 사용자의 지시에 따라 웹 브라우저를 직접 탐색하며 작업을 수행하는 자율형 에이전트 구조로, 복잡한 다단계 업무 처리 가능성을 보여준다.

Operator의 등장, 그리고 AI는 이제 ‘웹’을 직접 움직인다

OpenAI는 최근 ‘Operator’라는 새로운 형태의 웹 에이전트를 공개했다. 사용자의 명령을 이해한 후, 웹 브라우저 내에서 직접 클릭하고 입력하며 다양한 온라인 작업을 수행하는 자율형 시스템이다. 단순한 명령어 실행이 아니라, 복잡한 단계의 조합을 처리하는 능력이 요구되는 구조다. 예컨대 사용자가 “뉴욕행 항공권을 예약하고, 일정에 맞춰 호텔도 잡아줘”라고 입력하면, Operator는 실제 웹사이트에서 이 모든 과정을 ‘사람처럼’ 수행한다.

이는 웹 에이전트 기술이 새로운 국면에 접어들었음을 상징한다. 이제 AI는 대화형 응답만이 아니라, 실제 디지털 환경을 직접 조작하며 복합적인 업무를 수행하는 존재로 진화하고 있다. 그러나 과연 이러한 시스템은 정말로 ‘실수 없이’ 작동할까?

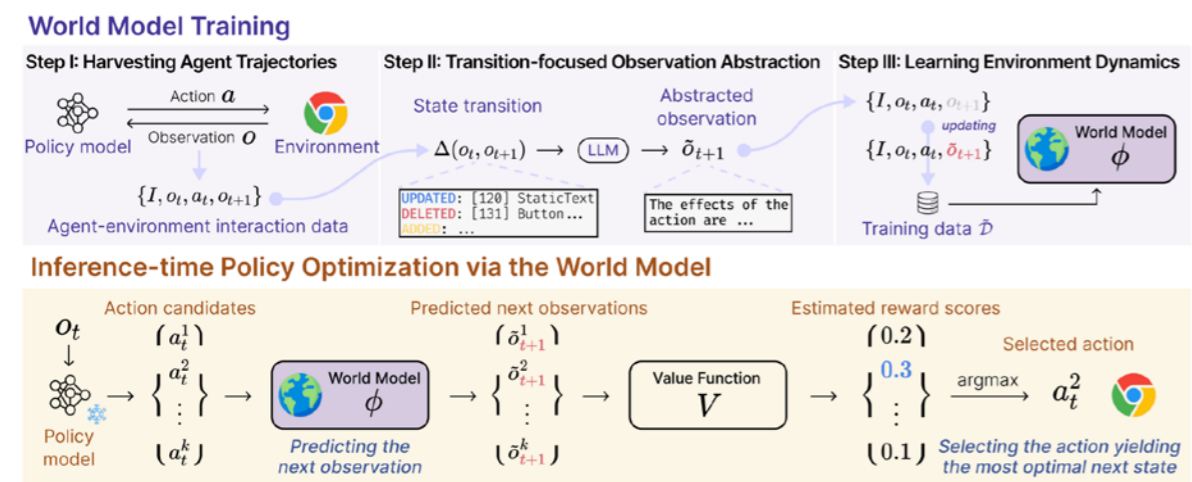
LLM의 놀라운 능력, 그런데 왜 웹에서는 빠듯할까?

거대 언어 모델(Large Language Models, LLMs)은 수학, 글쓰기,

프로그래밍 등 다양한 영역에서 인간 수준의 성과를 보여주고 있다. 하지만 ‘웹 탐색’이라는 일상적인 업무 앞에서는 여전히 허점을 드러낸다. 쇼핑몰에서 물건을 고르고, 항공권을 예약하는 단순한 작업에서조차도 자주 실수가 발생하며, 복잡한 과업에서는 더욱 불안정해진다.

연세대 언어 및 범용인공지능 연구실을 이끌고 있는 여진영 교수.





연세대 연구팀이 개발한 WMA의 프레임워크 개요. 상단은 웹상의 상태 변화를 예측하기 위한 월드 모델 학습 과정을, 하단은 예측된 결과를 바탕으로 최적의 행동을 선택하는 정책 최적화 과정을 나타낸다.

문제의 핵심은 ‘예측’이다. 인간은 클릭하기 전에 머릿속으로 다음 화면을 대략 예측하고, 예상과 다르면 행동을 수정한다. 반면 대부분의 LLM 기반 웹 에이전트는 현재 화면만을 기반으로 행동을 결정하고, 그 결과를 미리 상상하거나 시뮬레이션하지는 못한다. 이로 인해 동일 항공편을 두 번 예약하거나, 환불 불가 상품을 잘못 구매하는 일이 발생한다. GPT-4 같은 강력한 모델조차 이런 실수에서 완전히 자유롭지 않다.

이는 단순한 기능 부족이 아니라, 인간처럼 ‘이 행동을 하면 무슨 일이 벌어질까?’를 미리 생각하는 능력, 즉 월드 모델(World Model)의 부재에서 기인한다.

World Model for Web Agents: 행동 전에 생각한다, AI도

세계적인 AI 석학 얀 르쿤은 “AGI(범용 인공지능)의 핵심은 세상의 변화를 예측할 수 있는 능력, 즉 월드 모델(World Model)”이라고 강조해왔다. 단순한 문장 생성이 아닌, 환경과 상호작용하는 시스템에서는 이 예측 능력이 필수적이다. 연세대학교 연구팀은 이에 착안해 WMA(World-Model-Augmented 웹 에이전트)라는 새로운 구조를 제안했다.

WMA의 핵심은 ‘행동 전에 결과를 예측해보는 것’이다. 사용자의 지시를 받은 에이전트는 다양한 행동 후보들에 대해 그 결과(다음 웹페이지 상태)를 시뮬레이션하고, 가장 바람직한 행동을 선택한다. 이는 마치 인간이 ‘이 버튼을 누르면 어떤 화면이 나올까?’를 머릿속으로 그려보는 것과 유사하다.

예측의 효율성도 중요하다. 단순히 HTML 전체를 예측



이번 연구논문의 제1 저자인 채형주 연구원.

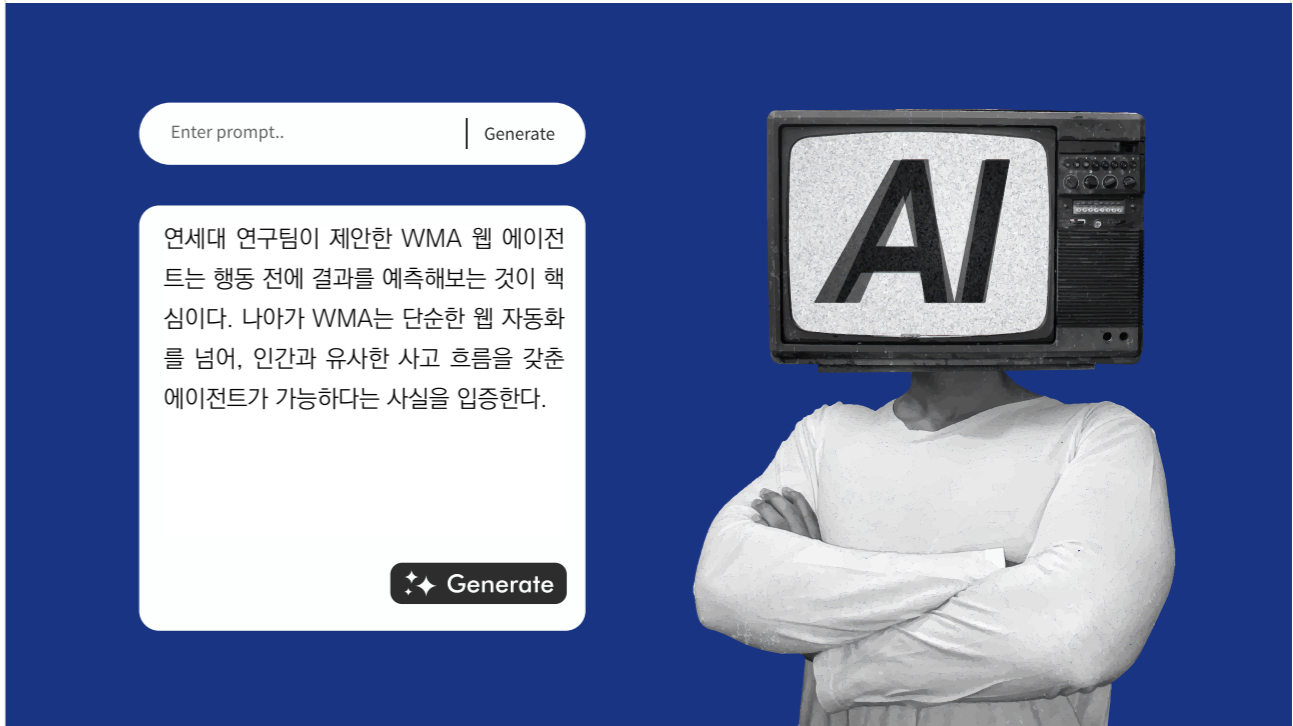
하면 비효율이 심각해진다. 연구팀은 이에 대한 해법으로 전이 중심 관찰 요약(transition-focused observation abstraction)을 제안했다. 이전과 이후 웹페이지의 상태를 비교하여 변화된 부분만 추출하고, 이를 자연어로 요약하는 방식이다.

“이전과 다음 웹페이지의 차이만 뽑아 요약하는 방식은, 마치 사람이 ‘뭐가 달라졌지?’를 빠르게 파악하는 방식과 닮아 있다.”

이 접근은 모델이 학습할 내용을 정밀하게 좁혀주며, 연

아니라, 스스로 전략을 세우고 수정하며 협력할 수 있는 존재로 진화할 것이다.

무엇보다 중요한 것은, 이러한 구조가 웹 탐색에 국한되지 않는다는 점이다. 월드 모델 기반 구조는 물리 세계의 로봇 시스템에도 적용 가능하다. 예컨대, 창고 물류 로봇이 다음 행동의 결과를 미리 예측하고 최적의 경로를 선택하거나, 안내 로봇이 엘리베이터 호출과 사용자 응대를 동시에 조율하는 것도 가능하다.



산 자원을 크게 절감한다. 동시에 예측된 변화는 ‘이 행동이 사용자 목표 달성에 기여하는가?’를 판단하는 가치 평가 단계로 연결된다.

Operator와 WMA, 그 의미는 단순한 자동화를 넘는다

WMA는 단순한 웹 자동화를 넘어, 인간과 유사한 사고 흐름(예측 → 판단 → 실행)을 갖춘 에이전트가 가능하다는 사실을 입증한다. Operator가 보여준 잠재력은 WMA와 같은 기술이 실현 가능한 기반 위에 있음을 보여준다. 앞으로의 웹 에이전트는 단순히 명령을 수행하는 도구가

즉, WMA는 디지털 세계의 에이전트에서 물리 세계의 안드로이드로 이어지는 기술적 징검다리라 할 수 있다.

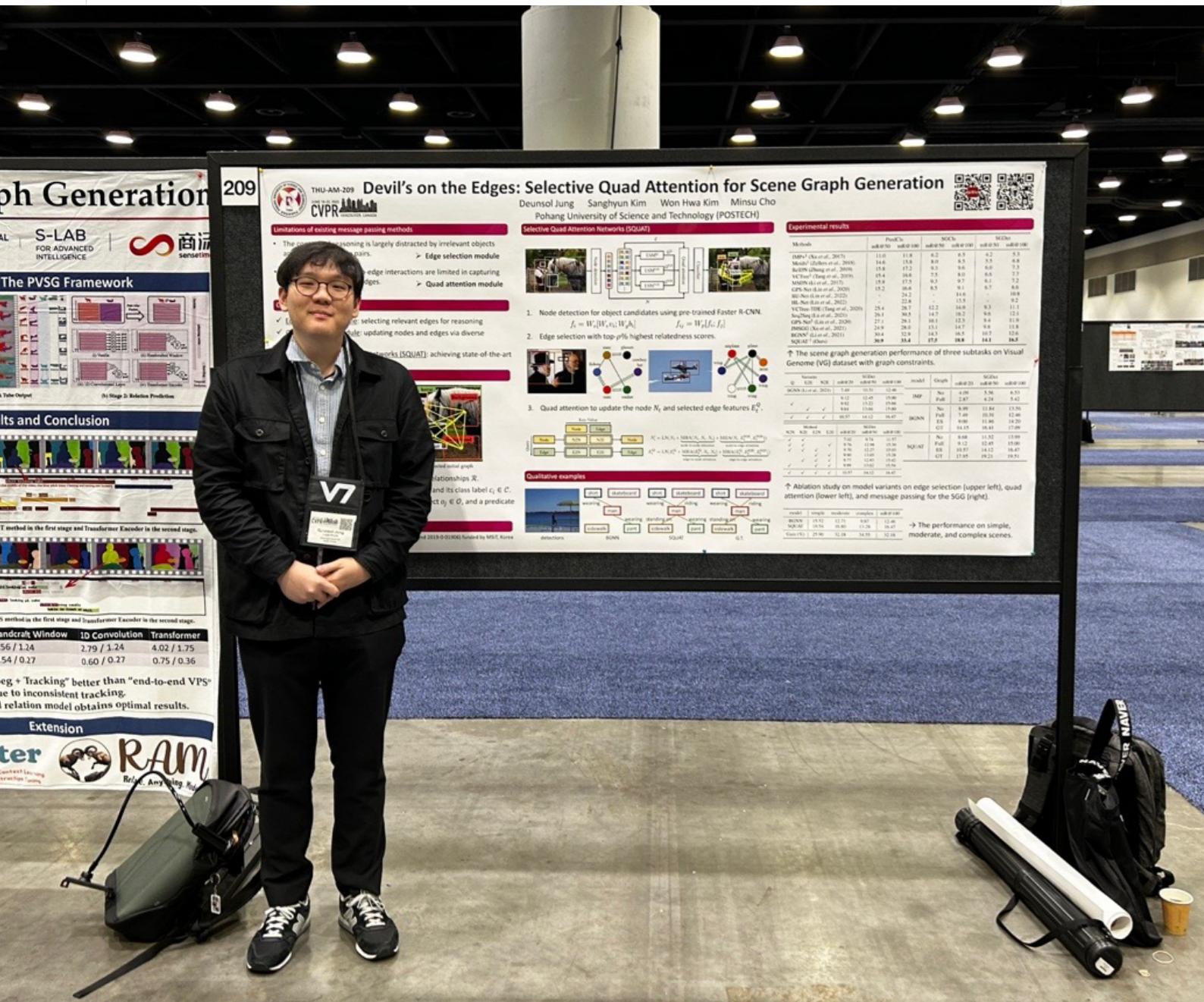
“우리가 원하는 에이전트는 단순히 클릭을 잘하는 도구가 아니라, 실수하지 않고 스스로 고민하는 ‘디지털 동료’이다.”

WMA는 그 첫걸음을 보여줬다. 앞으로의 AI는 단순한 응답 기계가 아니라, 스스로 판단하고 협력하는 진정한 파트너로 발전할 것이다. 그 변화는 지금, 우리가 마주하는 웹 브라우저 위에서 조용히 시작되고 있다. [NAIRL](#)

정든술
국가AI연구거점 박사후연구원

제로샷 인간-객체 상호작용 탐지를 위한
지역성 및 관계성 주입 네트워크

Locality-Aware Interaction Adapter for Zero-shot HOI Detection



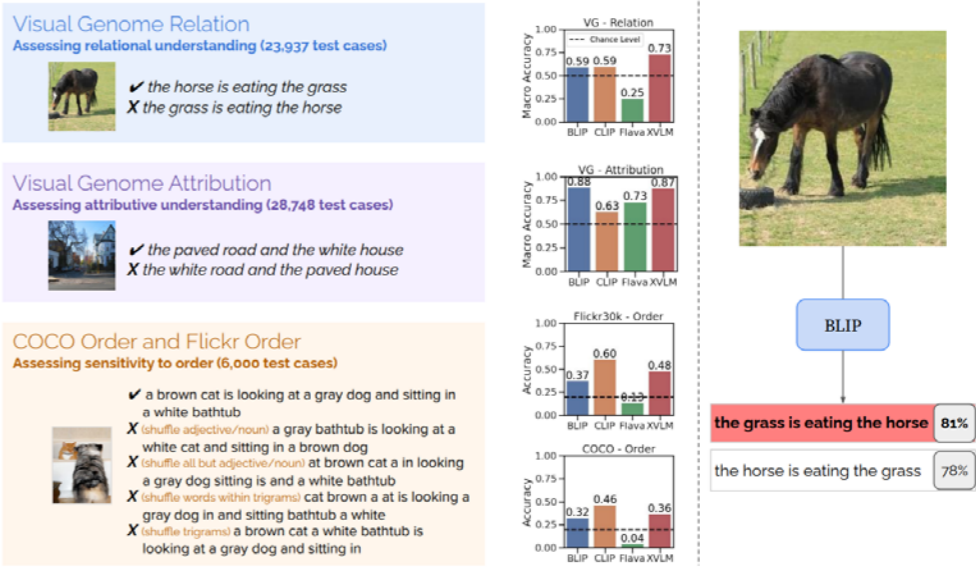
최근 CLIP¹이나 BLIP² 같은 대규모 비전-언어 모델
Large Vision-Language Model, LVLM이 기반 모델foundation model로서

다양한 컴퓨터 비전 문제를 푸는 데 도움을 주고 있다. 그
러나, 이러한 LVLM은 이미지와 이미지 전체를 묘사하는
설명 쌍으로 학습하여 이미지의 세세한 부분이나 객체 간
의 관계를 이해하는 데 어려움이 있다는 것이 밝혀지고 있
다. 특히, 어떤 논문[1]에서는 LVLM이 이미지의 구조적인
정보를 이해하는 것이 아닌 그저 'bag-of-objects', 즉 이
미지 내 객체의 집합만 이해하고 있다고 이야기한다. 예
를 들어, [그림 1]의 오른쪽을 보면, 말이 풀을 먹고 있는the
horse is eating the grass 이미지임에도 BLIP은 풀이 말을 먹고
있다the grass is eating the horse는 현실적이지 않은 설명의 점수
가 더 높게 평가되고 있다. 즉, LVLM이 이미지를 제대로
이해하는 것이 아니라 풀과 말 등 객체들이 존재한다는 사
실만 이해하고 점수를 낸다는 것이다.

이러한 문제를 해결하기 위해 최근에는 객체 간의 관계
가 레이블 된 장면 그래프scene graph 혹은 인간-객체 상호작
용Human-Object Interaction, HOI 데이터를 이용해 LVLM을 미세
조정fine-tuning하는 연구들이 속속 발표되고 있다[2, 3]. 이

렇게 이미지 내 객체 간의 관계를 이해하는 것은 이후 기
반 모델의 발전에 큰 역할을 할 것이다.

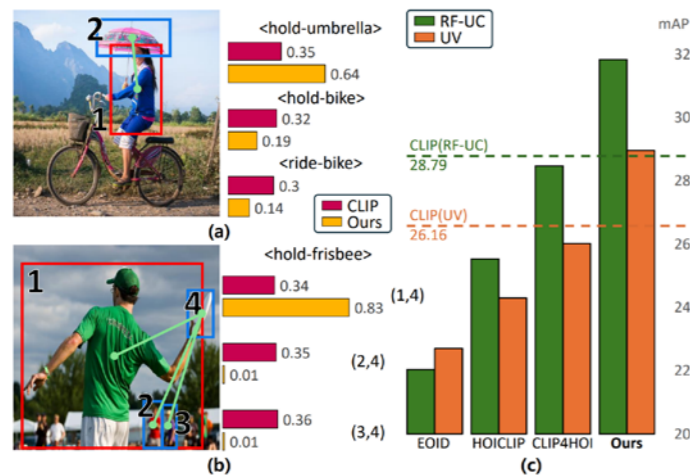
인간-객체 상호작용 탐지^{HOI detection}(HOI 탐지)는 이미
지 내 객체 간 관계 중에서도 인간과 타 객체 사이의 상호
작용을 찾는 문제이다. 최근에는 CLIP 등의 LVLM을 이용
해 학습 중 본 적 없는 상호작용 클래스에 일반화하는 제
로샷 HOI 탐지 문제에 대한 관심이 높아지고 있다. 이런
접근 방식에서는 이미지 내 인간과 객체의 시각적 특징
visual feature을 CLIP의 비주얼 인코더visual encoder로 추출하
고, 이를 통해 상호작용 특징feature을 구성한다. 이후 이 상
호작용 특징feature과 사전에 정의된 객체-상호작용 쌍 클래
스의 텍스트 특징text feature을 비교하여 유사도 점수score를
계산하고 상호작용 클래스를 예측한다. 예를 들어, [그림
2]의 (a)을 보면 여자와 우산의 시각적 특징을 이용해 상호
작용 특징을 만들고, 이를 hold-bike, ride-bike 같은 객체-
상호작용 쌍의 텍스트 특징과 비교하여 클래스를 추정한
다. 그러나, [그림 2]에서 보듯이 이 방식은 객체 간의 관계
가 지역적인 정보를 받아들이지 못해 이미지 내 존재하는
다른 객체(자전거)로 오해한다든지([그림 2] (a)), 다른 객



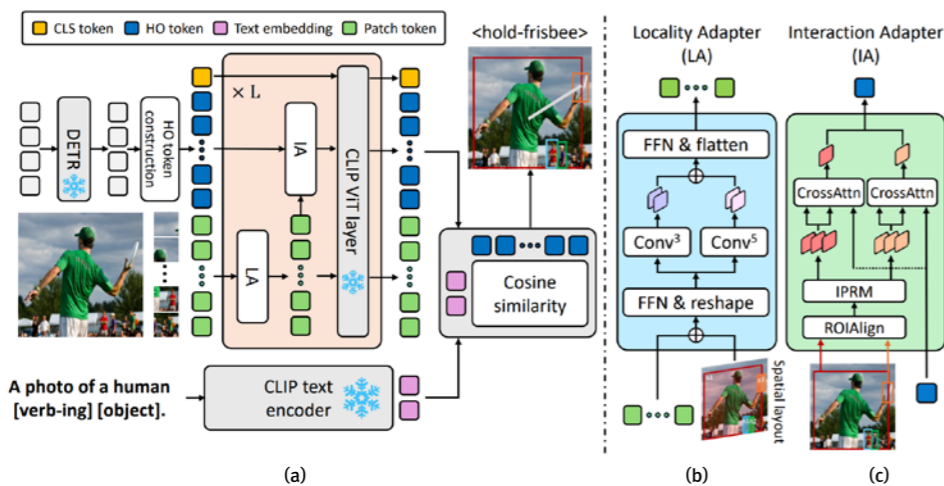
[그림 1] 기존 기-학습된 BLIP과 같은 LVLM은 이미지를 제대로 이해하고 있다기보다는 이미지 내 존재하는 객체들만 이해한다.

출처: [1]

¹ CLIP은 OpenAI에서 개발한 모델로, 이를 그대로 '언어와 이미지 간의 대조 학습(contrastive learning)'을 기반으로 사전학습된 비전-언어 모델
² BLIP은 Salesforce Research에서 발표한 모델로, CLIP 이후 등장한 LVLM 중 하나.



[그림 2] (a, b) CLIP과 같은 LVM을 활용한 모델은 지역적인 정보를 정확히 알지 못해 HOI 탐지에 문제가 발생한다. (c) 그 이후에 나온 모델들도 CLIP보다 성능이 떨어지고 있다.



[그림 3] (a) LAIN의 전체적인 구조도 (b) 지역성 적응기의 구조도 (c) 관계성 적응기의 구조도

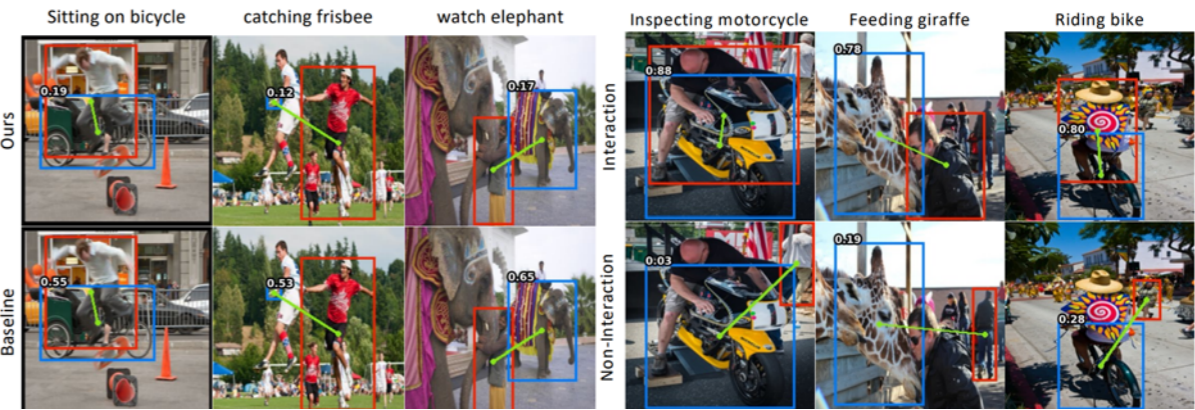
체와의 관계에 더 높은 점수를 주는 식([그림 2] (b))의 문제가 생긴다. 우리는 이러한 문제를 해결하기 위해 적응기 adapter 구조를 활용하여 CLIP에 지역적인 정보를 주입하고 이를 통해 HOI 탐지 성능을 개선하는 Locality-Aware Interaction Network(이하 LAIN)를 제안하였다.

LAIN은 지역성 적응기[locality adapter]([그림 3] (b))와 관계성 적응기[interaction adapter]([그림 3] (c))라는 두 종류의 적응기를 활용한다. 먼저 트랜스포머 구조로 이루어진 CLIP의 비주얼 인코더[visual encoder]에서 추출된 시각 토큰[visual token]을 지역성 적응기에 입력하여 CLIP에는 부족한 지역성을 주입한다. 지역성 적응기는 시각적 특징[visual feature]을 각자

형태로 되돌리고, 이미 학습된 객체 탐지기에서 나온 객체 탐지 결과를 연결[concat]하여 3×3, 5×5의 합성곱 신경망 convolution을 통과시킨다. 이렇게 합성곱 신경망을 통과시킨 시각 토큰[visual token]은 탐지된 객체의 위치 정보를 포함하고, 합성곱을 통해 지역성이 강화된다. 이렇게 지역성이 주입된 시각 토큰과 인간-객체 쌍 토큰(HO 토큰)을 관계성 적응기에 입력한다. 관계성 적응기에서는 HO 토큰에 해당하는 인간과 객체 탐지 결과를 이용해 시각 토큰에서 ROI Align[Region Of Interest Align]으로 국소적 특징[local feature]을 추출한다. 인간과 객체의 국소적 특징으로부터 상호작용 패턴을 추출하고 이를 HO 토큰에 주입한다. 이를 통해

| Method | RF-UC | | | NF-UC | | | UO | | | UV | | | UC | | |
|------------------------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------|-------|-------|
| | Unseen | Seen | Full | Unseen | Seen | Full | Unseen | Seen | Full | Unseen | Seen | Full | Unseen | Seen | Full |
| FCL [19] | 13.16 | 24.23 | 22.01 | 18.66 | 19.55 | 19.37 | 15.54 | 20.74 | 19.87 | - | - | - | - | - | - |
| ATL [18] | 9.18 | 24.67 | 21.57 | 18.25 | 18.78 | 18.67 | 15.11 | 21.54 | 20.47 | - | - | - | - | - | - |
| RLIP [56] | 19.19 | 33.35 | 30.52 | 20.27 | 27.67 | 26.19 | - | - | - | - | - | - | - | - | - |
| GEN-VLKT [30] | 21.36 | 32.91 | 30.56 | 25.05 | 23.38 | 23.71 | 10.51 | 28.92 | 25.63 | 20.96 | 30.23 | 28.74 | - | - | - |
| LOGICHOI [26] | 25.97 | 34.93 | 33.17 | 26.84 | 27.86 | 27.95 | 15.67 | 30.42 | 28.23 | - | - | - | - | - | - |
| ADA-CM [23] | 27.63 | 34.35 | 33.01 | 32.41 | 31.13 | 31.39 | - | - | - | - | - | - | - | - | - |
| EoID [52] | 22.04 | 31.39 | 29.52 | 26.77 | 26.66 | 26.69 | - | - | - | 22.71 | 30.73 | 29.61 | 23.01 | 30.39 | 28.91 |
| HOICLIP [37] | 25.53 | 34.85 | 32.99 | 26.39 | 28.10 | 27.75 | 16.20 | 30.99 | 28.53 | 24.30 | 32.19 | 31.09 | 23.15 | 31.65 | 29.93 |
| CLIP [39] | 28.79 | 22.00 | 23.36 | 28.52 | 22.06 | 23.36 | 28.66 | 22.29 | 23.36 | 26.16 | 22.90 | 23.36 | 24.28 | 23.12 | 23.36 |
| CLIP4HOI [36] | 28.47 | 35.48 | 34.08 | 31.44 | 28.26 | 28.90 | 31.79 | 32.73 | 32.58 | 26.02 | 31.14 | 30.42 | 27.71 | 33.25 | 32.11 |
| BCOM [†] [46] | 28.52 | 35.04 | 33.74 | 33.12 | 31.76 | 32.03 | - | - | - | - | - | - | - | - | - |
| CMMP [24] | 29.45 | 32.87 | 32.18 | 32.09 | 29.71 | 30.18 | 33.76 | 31.15 | 31.59 | 26.23 | 32.75 | 31.84 | 29.60 | 32.39 | 31.84 |
| LAIN | 31.83 | 35.06 | 34.41 | 36.41 | 32.44 | 33.23 | 37.88 | 33.55 | 34.27 | 28.96 | 33.80 | 33.12 | 31.64 | 35.04 | 34.36 |
| LAIN [†] | 36.57 | 38.54 | 38.13 | 37.52 | 35.90 | 36.22 | 40.78 | 36.96 | 37.60 | 32.05 | 38.04 | 37.20 | 32.25 | 37.95 | 36.81 |

[표 1] LAIN의 정량 결과



[그림 4] LAIN의 정성 결과

HO 토큰에 인간과 객체의 지역적 특징을 이용하여 관계성이 주입된다. 결과적으로 LAIN은 이렇게 두 종류의 적응기를 통해 CLIP의 특징에 지역성과 관계성을 주입하였으며, 제로샷 HOI 탐지 분야에서 [표 1]과 같이 최고 성능을 달성하였다. [그림 4]에서 볼 수 있듯이 [그림 2]의 기존 모델과는 달리 지역적으로 관계없는 객체와의 관계성이나 틀린 관계에 해당하는 점수가 현저히 낮아지는 것을 확인할 수 있다. 본 연구는 김상현 학생의 주도로 조민수 교수 지도하에 공동 작업을 통해 완성되었고, 연구 결과는 CVPR 2025에서 발표[4]되었다.

이후에도 계속해서 이미지 내 관계 탐지에 관한 연구를 계속하고 있다. 특히 최근에는 3D 장면에서의 관계 탐지에 관심이 있어 해당 방향으로 연구하는 중이다. 3D 장면에서의 관계 탐지는 최근 로봇 조작 혹은 경로 탐색 등 다

양한 문제에서 활용되고 있어 로봇 기반 모델 제작에 도움이 될 것이다. [NAIRL](#)

참고문헌

- [1] Yuksekgonul et al., “When and Why Vision-Language Models Behave Like Bag-of-Words, and What to Do About it?”, ICLR 2023, Stanford University
- [2] Herzig et al., “Incorporating Structured Representations into Pre-trained Vision & Language Models Using Scene Graphs”, EMNLP 2023, Tel-Aviv University
- [3] Mitra et al., “Compositional Chain-of-Thought Prompting for Large Multimodal Models”, CVPR 2024, University of California
- [4] Kim et al., “Locality-Aware Zero-Shot Human-Object Interaction Detection”, CVPR 2025, POSTECH

최예진 교수가 제시하는
AI 스케일링 혁신과 한국의 독자적 AI 개발 전략

“한국도 독자적 언어모델 개발해야”

최예진 교수

Dieter Schwarz Foundation HAI Professor |
Professor of Computer Science |
Senior Fellow, Stanford HAI |
Distinguished Scientist of Language and
Cognition Research, NVIDIA.

© John D. and Catherine T. MacArthur Foundation

AI 석학으로 유명한 최예진 스탠퍼드대 교수가 지난 6월 5일 국가AI연구거점을 방문했다. 최 교수는 ‘지능형 알고리즘으로 스케일링 법칙을 뛰어넘기(Bending Scaling Laws with Brighter Algorithms)’라는 주제로 강연한 뒤 인터뷰에 응했다.

‘타임’ AI 분야 100대 인물, 엔비디아에 LLM 선임 디렉터로 합류

지난 1월 스위스 다보스에서 열린 세계경제포럼(WEF)에서 대표 인공지능^{AI} 세션인 ‘범용인공지능^{AGI}의 새벽’에 연사로 참여, 지난해 가을부터 AI 반도체의 선두 주자 엔비디아에 합류, 2023년 미국 ‘타임’지로부터 AI 분야 100대 인물로 선정, 2022년 ‘천재들의 상’이라 불리는 ‘맥아더 펠로십’ 수상… 이 화려한 이력의 한국인이 바로 스탠퍼드대 인간중심AI연구소 최예진 교수다.

AI와 기계학습, 자연어처리^{NLP} 분야에서 세계적 권위자인 최 교수는 엔비디아의 대형언어모델^{LLM} 연구 부문 선임 디렉터를 겸하며 스탠퍼드대 교수로 재직 중이다. 서울대 컴퓨터공학과를 졸업하고 미국 코넬대에서 박사 학위를 받은 그는 스토니브룩대 교수와 워싱턴대 교수, 옥스퍼드대 AI윤리연구소 선임연구원 등을 거쳤다. 지금까지 세계 최고 권위의 AI 학회인 ACL^{Association for Computational Linguistics}, NeurIPS^{Neural Information Processing Systems}, CVPR^{Computer Vision and Pattern Recognition} 등 6곳에서 총 8개의 최우수·우수 논문상을 수상할 정도로 연구 실적도 화려하다.

창의적 연구철학과 협업 중요성 강조

먼저 ‘천재들의 상’이라는 애칭으로 불리는 맥아더 펠로십을 수상한 배경에 대해 들어봤다. 최 교수는 워싱턴대 재직 시절인 2022년 10월 미국 맥아더 재단으로부터 자연어처리 분야에서 AI가 상식 기반 추론을 수행하도록 하는 창의적 연구를 인정받아 약 5년에 걸쳐 지원금을 받게 됐다. 이에 최 교수는 “여기서 언급된 ‘천재’란 단순히 수

학 문제 풀이 능력이 뛰어난 사람과 같은 전통적 의미보다 새로운 영역에서 도전적이고 창의적인 연구를 수행하는 능력을 지닌 사람을 뜻한다”며 “기존에 남들이 시도하지 않은 분야에 대한 탐험을 중시하며, 실패를 두려워하지 않고 삶의 의미를 찾는 자세로 연구에 임한다”고 밝혔다. 유행을 따라가기보다는 빈 공간을 메우는 독창적 주제를 발굴하는 데 관심이 있었다는 뜻이다. 그는 또 “성공 자체에 집착하기보다 꾸준한 노력과 인내를 통해 성과를 이루어 왔다”고 덧붙였다.

AI 연구와 관련해 최 교수는 커뮤니케이션과 협업도 강조했다. 그는 “혁신적 아이디어는 개인의 고립된 노력보다 커뮤니티와의 협업을 통해 확산되고 발전한다”고 강조하면서 “널리 소통하고 동료 연구자들과 협업하는 것이 중요하다”고 밝혔다. 그는 커뮤니티와 함께 방향을 만들어가는 협력적 리더십을 중요시하면서 혼자선 어렵지만 남들과 함께하면 성과가 더 빨라질 수 있다는 인식도 내보였다.

수평적 문화와 개인의 잠재력 발현

이어 AI 인재 양성과 관련해 최 교수는 인터뷰에서 수평적 문화의 중요성과 개인의 잠재력도 언급했다. 그는 “권위적 환경보다 수평적이고 자유로운 토론이 가능한 문화가 인재 양성에 필수적”이라고 믿는다면서 “주입식 교육이 아니라 스스로 생각하고 도전할 수 있도록 돕는 환경을 조성해야 한다”고 말했다. 그는 또 “각 개인이 지닌 잠재력은 사회, 지역, 가정 문화 등 다양한 변수가 결합하여 발현되는데, 적합한 환경이 마련되면 개인의 역량은 크게 성장할 수 있다”고 밝혔다.

아울러 그는 한국 사회의 여성상에 대한 고정관념을 언급하면서 학생들에게도 자신의 개성을 유지하며 연구와 학습에 임할 것을 조언했다. 여성 과학자로서 그는 한국 여성들이 여전히 순종적이거나 연약해 보이는 이미지를 강요받는 것에 문제를 제기하면서 미국에서 더 자유롭게 자신의 스타일과 사고를 표현하며 삶의 만족도가 높아졌

다고 털어놓았다. 그는 평소에 검은 가죽 재킷을 즐겨 입는 것으로 알려져 있다.

‘스마트 스케일링’을 위한 획기적 알고리즘 제시, AI 패러다임 전환

특히 최 교수는 최신 AI 연구 동향을 전하며 ‘스마트 스케일링’을 언급했다. 스마트 스케일링은 그가 인터뷰에 앞서 강의한 내용과 관련이 깊다. 오랫동안 AI 분야에서는 ‘스케일링 법칙’이란 원칙이 지배적이었다. AI 모델에 데이터와 컴퓨팅 자원을 더 많이 투입할수록 성능이 비례해 향상된다는 법칙이다. 하지만 최 교수는 이렇게 ‘무작정 키우는’ 브루트포스 스케일링 방식이 이제 한계에 직면했다고 지적한다. 방대한 데이터를 수집하고 엄청난 컴퓨팅 자원을 동원하는 것이 환경적으로, 경제적으로 지속 가능하지 않다는 뜻이다. 이제는 단순히 ‘더 많이’가 아니라 ‘더 똑똑하게’ AI를 개발하는 방법을 찾아야 할 때라는 것이 최 교수가 강조하는 ‘스마트 스케일링’의 핵심이다.

강연에서 최 교수는 스마트 스케일링을 가능하게 하는 획기적인 알고리즘 혁신을 몇 가지 소개했다.

첫째, 합성 데이터를 활용한 스마트 학습이다. 현실 데이터 없이도 합성 데이터(synthetic data)로 똑똑하게 학습하기를 제안했다. AI 모델 학습에는 엄청난 양의 데이터가 필요하지만, 모든 데이터를 현실에서 얻기 힘들고 비용도 많이 들므로, 컴퓨터가 실제 데이터를 바탕으로 만들어낸 ‘가짜’ 데이터인 합성 데이터를 활용하자는 뜻이다. 최 교수는 ‘그래디언트 기반 방법론’을 이용해 합성 데이터의 품질을 극대화하는 연구를 진행하고 있다. 이는 단순히 데이터를 생성하는 것을 넘어 AI 학습에 가장 효율적인 데이터를 만들어냄으로써 적은 양의 데이터로도 모델 성능을 높일 수 있는 방법론이다.

둘째, ‘테스트 시점 추론(test-time reasoning)’이다. 일반적 AI 모델은 훈련이 끝나면 그 상태로 ‘추론(inference)’을 수행한다. 즉 학습된 지식만 활용해 주어진 문제를 해결한다는

뜻이다. 하지만 최 교수는 모델이 ‘테스트 시점에서도’ 더 깊이 추론할 수 있는 개념, 즉 ‘테스트 시점 추론’을 제시한다. 이는 ‘기호적 탐색(symbolic search)’이라는 알고리즘을 활용해, 사람이 문제를 만났을 때 잠시 고민하고 여러 가능성을 탐색하듯이, AI 모델도 테스트 시점에 새로운 정보를 탐색하고 이를 바탕으로 더 정확하고 심층적인 문제 해결 능력을 발휘할 수 있도록 돕는 것이다. 더 나아가 최 교수는 ‘테스트 시점 학습(test-time training)’이라는 더 혁신적인 개념도 소개한다. 기존에는 AI 모델의 학습이 완료되면 모델의 내부 가중치(weights)가 고정되어 더 이상 변하지 않았지만, 테스트 시점 학습은 모델이 테스트를 수행하는 동안에도 끊임없이 학습을 지속하게 하는 방식이다. 이를 통해 모델은 새로운 환경이나 예상치 못한 데이터에 직면했을 때 즉시 스스로를 업데이트하고 적응력을 높여 성능을 향상시킬 수 있다.

끝으로 최 교수는 ‘새로운 토큰화 알고리즘’을 제시한다. 토큰화(tokenization)는 문장을 AI가 이해할 수 있는 작은 단위(토큰)로 쪼개는 과정인데, 이는 대형언어모델(LLM)의 효율성과 성능에 지대한 영향을 미친다. 현재 널리 쓰이는 ‘바이트 페어 인코딩(BPE, Byte Pair Encoding)’과 같은 방식의 한계를 극복하고자, 최 교수는 더욱 정확하고 빠른 추론을 가능하게 하는 새로운 토큰화 알고리즘을 제안했다. 이 새로운 방식은 언어모델이 정보를 처리하는 속도를 획기적으로 높이는 동시에 이해도를 향상시켜 더욱 정확한 답변을 제공하게 될 것이라고 한다.

최 교수는 AI 개발이 단순히 자원을 ‘양적’으로 늘리는 것을 넘어 알고리즘의 ‘질적’ 혁신을 통해 더욱 지속 가능하고 효율적인 성장을 이루어야 한다는 강력한 메시지를 전했다. AI의 미래가 무한한 자원 경쟁이 아니라, 더욱 영리하고 정교한 기술 개발에 달려 있다는 뜻이다.

“독자적 모델 통해 기술 주권 확보, 비즈니스 모델 구축 가능”

최근 국내에서 ‘소버린 AI’에 대한 관심이 높아지고 있



지난 6월 5일 국가AI연구거점을 방문한 최예진 교수.

다. 이에 대해 최 교수는 자국 내 LLM 개발의 필요성을 강조했다. 즉 그는 “외국에서 제공하는 LLM에 의존하기보다 국내에서도 독자적인 모델을 개발해 보유해야 한다”며 “그래야만 기술 주권을 확보하고 데이터 프라이버시를 지키며 비즈니스 모델을 구축할 수 있다”고 설명했다.

또 한국이 대량의 그래픽처리장치(GPU)를 확보하려는 계획에 대해서도 최 교수는 다음과 같이 조언했다. 그는 “GPU 자원이 한정적이라든가 방법을 축소해 문제를 해결하거나 클라우드나 공용 인프라를 활용해 실제 연구를 수행할 수 있다”고 제안하면서 “작은 규모라도 직접 시도해 봐야 기술적 직관과 전문성을 키울 수 있다”고 말했다.

최 교수가 거대언어모델을 따라갈 필요 없이 한국 맞춤형 소형 모델로 시작해도 의미가 있다고 밝힌 셈이다. 그는 국내 AI 생태계 강화에 대해서도 의견을 피력했다. 즉 국내 기업, 정부, 학계가 협력해 자국 기술을 개발하고 육성함으로써 글로벌 경쟁력과 함께 데이터 주권을 확보해

야 한다는 생각을 드러냈다.

한국 AI 생태계의 미래 전망

최예진 교수의 이번 강연과 인터뷰는 한국의 AI 발전 방향에 중요한 시사점을 제공한다. 단순히 해외 기술에 의존하기보다는 독자적인 기술 개발과 혁신적인 알고리즘 연구를 통해 지속 가능한 AI 생태계를 구축해야 한다는 메시지가 핵심이다.

특히, 국가AI연구거점(NAIRL)과 같은 연구 인프라를 통해 세계적 수준의 연구를 수행하고, 소버린 AI 개발을 통해 기술 주권을 확보하는 것이 한국이 AI 3대 강국으로 도약하기 위한 필수 과제임을 시사한다. 최 교수가 제시한 스마트 스케일링 개념은 자원 제약이 있는 한국 상황에서 특히 유용한 접근법이 될 수 있을 것으로 평가된다.



디지털 헬스케어 혁신 기업,
에버엑스 EverEx

AI로 재활의 미래를
바꾼다



지난 7월 16일 국가AI연구거점의 파트너 기업 중 하나인 에버엑스 EverEx를 방문했다. 서울 강남대로에 자리한 에버엑스에서 윤찬 대표, 김병훈 이사, 정형진 박사를 만나 AI를 기반으로 한 회사의 기술력과 디지털 헬스케어의 미래에 대해 들었다.

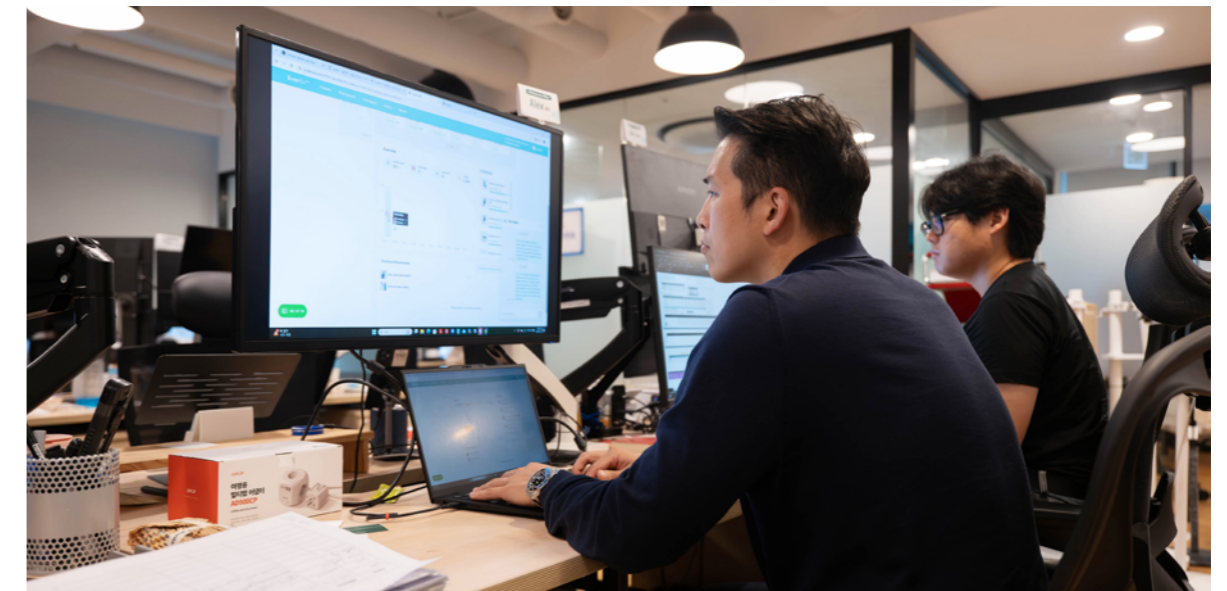
왼쪽부터 에버엑스의 정형진 박사, 김병훈 이사, 윤찬 대표.

'가치 중심 의료'에서 떠오르는 AI 기반 디지털 재활 솔루션

병원에서 받는 재활 치료를 집에서도 정밀하게 받을 수 있다면 어떨까? AI를 활용해 바로 이 질문에 답하는 혁신 스타트업이 바로 에버엑스다. 정형외과 전문의 출신인 윤찬 대표가 2019년에 창업한 에버엑스는 근골격계 질환을 위한 AI 기반 디지털 재활 솔루션을 개발하며 국내외 의료 현장에서 주목받고 있다.

최근 전 세계 의료 시스템이 주사, 수술 같은 의료 서비스 하나하나에 비용을 청구하는 방식인 '행위 중심 의료'

Everyone'인데, 여기서 개인화된 부분에 AI가 중요한 전략적인 기술 코어를 이룬다"면서 "개인화된 치료는 미래 의학에서 중요한 부분이고 AI가 중요한 역할을 할 것"이라고 설명했다. 에버엑스는 병원 치료뿐만 아니라 일상에서도 언제 어디서나 맞춤형 재활 운동이 가능하도록 돕는 것을 목표로 하고 있다. 현재 에버엑스는 삼성전자, LG전자, 노바티스와 전략적 파트너십을 맺고 있으며, 서울아산병원, 삼성서울병원, 가톨릭대 서울성모병원 등의 여러 의료기관과 함께하고 있다. 또 미국 캘리포니아주 산호세에 미국



윤찬 대표의 자리는 직원들과 함께 쓰는 사무실 한쪽에 있다.

에서 환자의 건강 결과에 기반해 보상하는 방식인 '가치 중심 의료'로 전환되면서 개인 맞춤형 재활 솔루션, AI 기반 디지털 치료제 등이 부상하고 있다. 윤 대표는 "정형외과에서 10년 정도 진료할 때만 해도 근골격계 질환 환자가 많지만 재활 운동 치료를 대면으로 제공하기에 의료인이 적어 해결책이 없었다"면서 "이제 생성형 AI 모델이 고도화되면서 많은 환자에게 큰 가치를 줄 수 있는 디지털 솔루션을 제공할 수 있게 됐다"고 설명했다.

김병훈 이사는 "회사의 비전이 '재활이 필요한 모두에게 맞춤형 재활을 제공한다'(Bringing Personalized Rehabilitation to

지사도 두고 있다.

AI가 실시간으로 자세를 분석한다

윤 대표가 스마트폰으로 직접 시연하며 에버엑스의 대표 플랫폼인 모라 MORA와 자체 개발한 AI 자세 추정 모델인 그리핀 Griffin의 성능을 실시간으로 보여준다. 모라 Mobile Orthopedic Rehabilitation Assistant는 환자의 자세를 AI가 실시간으로 분석하고 피드백을 제공하는 스마트 재활 시스템이고, 그리핀은 24개 핵심 관절을 기반으로 수만 건의 임상 데이터를 학습해 정확한 움직임 평가를 지원한다. 이를 통

해 병원 밖에서도, 환자 스스로 효과적인 맞춤형 운동이 가능해졌다.

구체적으로 살펴보면, MORA Vu는 AI 자세 추정 기술을 활용해 근골격계 동작을 분석하는 의료기기로 식품의약품안전처 2등급 인증을 받은 소프트웨어 의료기기다. 초기 도입 비용과 공간 부담이 없으며 의학적 근거를 기반으로 해 정확한 근골격계 상태를 분석해준다. 또 MORA Ex는 환자 만족도를 높이는 맞춤형 근골격계 재활운동 솔루션이며, MORA Care는 기업 임직원을 위한 AI 근골격계 건강관리 서비스이다. 이로써 3,000개 이상의 운동 영상을 활용해 환자 맞춤형 재활 프로그램을 생성하며 운동 수행결과에 따른 데이터 기반 치료가 가능하고, 재활 전문가의 임직원 1 대 1 맞춤형 케어도 가능하다.

최근 재활운동과 인지행동치료를 병행하는 근골격계 질환 디지털 치료기기로 개발한 MORA Cure는 식품의약품안전처의 임상시험계획 승인 및 임상시험 진행 중에 있다. 아울러 재활운동 원격 치료 모니터링^{RTM} 솔루션인 EverEx Rehab은 미국식품의약국^{FDA} 2등급 등록 의료기기로 주목받고 있다. 1,600개 이상의 운동 영상을 활용해 환자 맞춤형 재활 프로그램을 생성하며, 원격 치료 모니터링 수가를 통해 의료진에게 수익을 창출해 준다. 글로벌 진출을 위한 발판 또한 빠르게 마련 중이다.

AI 기술과 임상 경험의 결합

에버엑스의 강점은 현장 중심의 데이터와 AI 기술의 결합이다. 병원에서 수집된 실제 운동 영상과 의료 전문가의 지식을 기반으로, AI 모델이 실제 환자의 재활 상황을 이해하고 대응할 수 있도록 이용된다. 이를 통해 단순 자세 인식이 아니라, 임상적으로 의미 있는 분석과 운동 가이드가 가능해진다.

그렇다면 에버엑스가 공개 API를 사용하는 대신 AI 자세 추정 모델을 자체 개발한 이유는 무엇일까. 정형진 박사는 “기존 자세 추정과 달리 관절 위주로 위치를 잡되,

17개 포인트 대신 24개 포인트를 잡았고 데이터 수집 프로세스를 고도화해 재활에 좀 더 맞는 모델을 만들었다”고 답했다. 정 박사는 또 “지금은 자세 추정 모델을 이용해서 규칙 베이스로 하고 있다면, 앞으로는 차세대 기술로 생각하는 비디오 이해 모델을 통해 자동화된 AI 에이전트로 만드는 방향으로 가려 한다”면서 “비디오 이해 모델을 통해 사람의 움직임을 이해하려면 비디오 LLM 관련 기술 개발이 필요한데, 이게 잘되면 다양한 분야에 이용할 수 있는 범용 기술이 될 수 있다”고 덧붙였다.

데이터 수집과 관련해 디지털 치료기기는 사용자의 순응도가 중요하다. 김병훈 이사는 “기존 제품 중에는 순응도가 떨어져 잘 안 되는 사례도 많은데, 우리는 이를 해결하기 위해 인지행동 치료 같은 부분을 접목하고 있다”면서 “순응도를 확보하기 위해 체계적인 사용성 평가도 중요하게 생각한다”고 밝혔다. 인지행동 치료는 통증에 대한 왜곡된 인지 수정, 행동적 개입 유도, 개인화된 AI 코칭 등을 하고 있다.

스마트폰에서 모라 플랫폼을 시연하는 모습.



에버엑스의 김병훈 이사(왼쪽), 정형진 박사(가운데), 윤찬 대표(오른쪽)는 입을 모아 버티컬 AI의 중요성을 강조했다.

국가AI연구거점과의 협력

에버엑스는 현재 국가AI연구거점의 파트너 기업으로 참여 중이다. 김병훈 이사는 “지난해 KAIST에서 국가AI연구거점 사업을 한다는 소식을 접하고 기회가 되면 꼭 참여하고 싶다고 김기웅 센터장에게 말씀드렸다”면서 “과제 제안서를 준비할 때부터 참여해 국가AI연구거점이 수립된 이후에도 함께 협업을 이어나가고 있다”고 말했다.

특히 국가AI연구거점의 첨단 연구와 에버엑스의 제품 개발이 시너지를 낼 수 있는 기회가 될 것으로 기대하고 있다. 김 이사는 “뉴럴 스케일링 법칙과 연관된 파운데이션 모델의 고도화에 큰 관심을 가지고 있으며, 이를 통해 인체에 대한 이해도를 높인 차세대 모델을 구축하고자 한다”면서 “이런 모델은 향후 자사 헬스케어 제품군에 직접 적용되어 AI 기반 의료 서비스의 정밀성과 개인화 수준을 한층 더 향상시킬 수 있을 것으로 기대된다”고 밝혔다. 그는 또한 “로보틱스 분야에도 관심을 갖고 있으며, 향후 AI 로보틱스 기술과의 융합을 통한 공동 연구 및 제품 연계 가능성도 열려 있다고 보고 있다”고 덧붙였다.

김 이사는 “국가AI연구거점의 훌륭한 연구자들과 소통할 기회가 생긴 것이 너무 감사하다”면서 “미래에는 차세대 제품에 생성형 AI와 같은 최첨단 기술을 빠르게 도입하고 기술적 차별화를 이루는 방향으로 나아가고 싶다”고 강조했다.

글로벌 시장으로의 확장, 버티컬 AI 강조

현재 미국을 중심으로 디지털 헬스케어 시장 자체가 굉장히 빠르게 확대되고 있다. 윤 대표에 따르면, 근골격계 질환에 대한 디지털 솔루션 시장의 경우 미국에서 디지털 솔루션을 사용하는 처방 건수가 매년 4.5배 정도씩 증가하고 있다. 윤 대표는 “회사의 기술 역량을 바탕으로 글로벌 시장으로의 확장을 적극 모색하고 있다”면서 “미국에서는 유사한 솔루션이 빠르게 수용되고 있으며 아시아 시장은 블루 오션에 가깝다는 점에서 에버엑스의 기술과 제품이 국제 시장에서 중요한 역할을 할 수 있을 것”이라고 강조했다.

에버엑스는 이미 CES 2024 혁신상, iF 디자인 어워드 2024, Red Dot 수상 등으로 기술성과 디자인을 모두 입증했으며, 식약처 임상시험 승인, TIPS 선정, 미국 현지 실증 계약 등 실질적인 사업화 성과도 이어가고 있다.

국내 AI 생태계와 관련해 에버엑스의 세 사람은 버티컬 vertical AI를 강조했다. 윤 대표는 “소버린 AI 역량을 갖추는 것도 중요하지만 우리나라가 진짜 앞서갈 수 있는 분야는 응용에 강한 버티컬 AI”라며 “특히 헬스케어, 재활 치료처럼 도메인 지식과 사용자 경험이 중요한 영역에서는 우리가 더 잘할 수 있다”고 강조했다. 그는 또 버티컬 AI에 GPU 같은 자원이 좀 더 풀리면, 우리나라 AI 산업 전체가 함께 성장할 수 있다는 뜻도 내비쳤다. [NAIRL](#)

‘AI 석학세미나 및 기술협력 네트워킹 스페셜 토크 시리즈’ 개최

국가AI연구거점은 KAIST AI 대학원 및 파트너십 기업을 대상으로 2025년 2월부터 ‘AI 석학세미나 및 기술협력 네트워킹 스페셜 토크시리즈’를 정기적으로 개최하며, 글로벌 AI 연구의 최전선에서 활발한 교류와 협력을 이끌고 있다.

본 프로그램은 미국, 캐나다, 프랑스, 홍콩, 아랍에미리트 등에 있는 세계 유수 대학 및 연구기관의 저명한 교수진이 직접 방한하여, 최신 연구 성과와 기술 트렌드를 국내외 연구자 및 산업계와 심도 있게 논의하는 장을 제공한다. 매월 진행되는 세미나와 스페셜 토크시리즈에서는 생성형 AI와 저작권, 이벤트 기반 비전, 신뢰성과 안전성을 갖춘 AI, 베이지안 추론의 최신 동향, 대규모 언어모델LLM과 강화학습, 멀티모달 파운데이션 모델, 3D 재구성 및 생성형 시각 콘텐츠, 데이터 소유권 등 AI 분야의 핵심 이슈와 혁신 기술이 폭넓게 다루어진다. 각 세션은 실제 연구 현장에서 도출

된 인사이트와 글로벌 트렌드를 공유함으로써 참가자들이 실질적인 연구 역량을 강화하고 새로운 산학연 협력 기회를 모색할 수 있도록 설계되었다.

특히, 국가AI연구거점의 해외 공동연구진으로 참여 중인 세계적 석학들이 2주간 국내에 체류하며 세미나 외에도 기술협력 관계자와의 집중 네트워킹 미팅을 진행하고 있다. 이로써 산학연 공동 연구, 혁신 프로젝트 발굴, 국제 협력 네트워크 확대 등과 같은 구체적인 성과 창출로 이어지고 있다.

국가AI연구거점은 앞으로도 세계적 석학들과 함께 최신 AI 연구와 산업 트렌드를 심도 있게 조망하는 세미나 및 스페셜 토크시리즈를 지속적으로 개최할 계획이다. 연구자와 실무자 모두가 성장할 수 있는 열린 지식 생태계 조성을 목표로 적극적인 관심을 갖고 참여하기를 기대한다. [NAIRL](#)



| 일정 | 연사 | 강연 제목 |
|-----------|---|--|
| 2월 25일 | 김윤명 소장 (디지털정책 연구소) | 생성형 AI와 저작권법 |
| 4월 1일 | 윤국진 교수 (KAIST 기계공학과) | Event-based vision: 고속 프레임 대체 기술 |
| 4월 11일 | 허준화 Research Scientist (Google DeepMind) | Scalable approaches for visual scene understanding |
| 4월 29일 | Edwin Fong 교수 (홍콩대학교 통계·보험수리학과) | Recent Advances in Predictive Bayesian Inference |
| 5월 27일 | 김명주 소장 (AI안전연구소) | 신뢰할 수 있으며 안전한 AI |
| 5월 26~29일 | Scott Sanner 교수 (토론토대학교 산업공학·컴퓨터공학과) | 스페셜 토크 시리즈: LLMs, Embodied AI, Hybrid Planning 등 |
| 6월 5일 | 최예진 교수 (스탠퍼드대학교 컴퓨터공학과) | 스페셜 토크 Bending Scaling Laws with Brighter Algorithms |
| 6월 11~17일 | Pascal Poupart 교수 (위털루대학교 컴퓨터학과) | 스페셜 토크 시리즈: RL for LLM Alignment, Risk-sensitive RL 등 |
| 6월 20일 | Molei Tao 교수 (조지아공과대학교 수학과) | Implicit Biases of Large Learning Rates in ML |
| 6월 27일 | Ming-Hsuan Yang 교수 (캘리포니아대 머시드캠퍼스 전기공학/컴퓨터과학과) | Recent Results on Multimodal Foundation Models |
| 7월 1~2일 | Shenghua Gao 교수 (홍콩대학교 컴퓨터공학·데이터과학과) | Geometric-prior기반 3D 재구성 및 시각 콘텐츠 생성 |
| 7월 8일 | Jun-Yan Zhu 교수 (카네기멜론대학교 컴퓨터공학과) | Ensuring Data Ownership in Generative Visual Models |

국가AI연구거점, ‘연구성과 공유 및 참여기업 네트워킹 데이’ 개최



‘국가AI연구거점 연구성과 공유 및 참여기업 네트워킹 데이’에 참석한 주요 인사. 전성수 서초구청장, 장은정 정보통신기획평가원 AI디지털융합단장, 정승 KAIST AI대학원장 등이 자리를 빛냈다.



(좌측 상단부터) 환영사에 나선 김기웅 센터장, 첫 번째 세션에서 발표한 양은호 KAIST 교수, 이병준 고려대 교수, 조성현 POSTECH 교수, 노알버트 연세대 교수.

그 중심에 AI가 있다”면서 “AI는 기술 혁신을 넘어 사회, 문화 전반의 변화를 이끄는 힘인데, 이 자리가 그 변화의 역사적 출발점이 되길 바란다”고 강조했다.

이번 행사에서는 국가AI연구거점이 수행하고 있는 3가지 핵심 연구 주제인 뉴럴 스케일링 법칙 초월 연구(1세부과제), 로봇 파운데이션 모델 연구(2세부과제), 초고차원 멀티모달 파운데이션 모델 연구(3세부과제)와 관련된 연구성과들이 공개됐다. 첫 번째 세션에는 ‘뉴럴 스케일링 법칙 초월 연구’라는 주제로 AI 모델의 성능, 효율, 안정성을 향상하는 4건의 대표연구가 발표됐다. 먼저 1세부의 연구책임자인 양은호 KAIST 교수가 추측적 디코딩(speculative decoding)을 통한 시각적 자기회귀 모델(visual auto-regressive model)의 생성 가속으로 시각 이미지 생성 모델의 속도 한계를 개선한 연구를 소개했다. 이어 이병준 고려대 교수는 모델 학습 과정에 강화학습 기반의 실시간 최적화 전략을 도입해 학습 효율을 높이고 자원 소모를 줄이는 연구성과를 전했다. 조성현 POSTECH 교수는 생성형 AI를 활용해 3D 모델의 품질을 자동으로 높이는 기술을 선보였다. 끝으로 노알버트

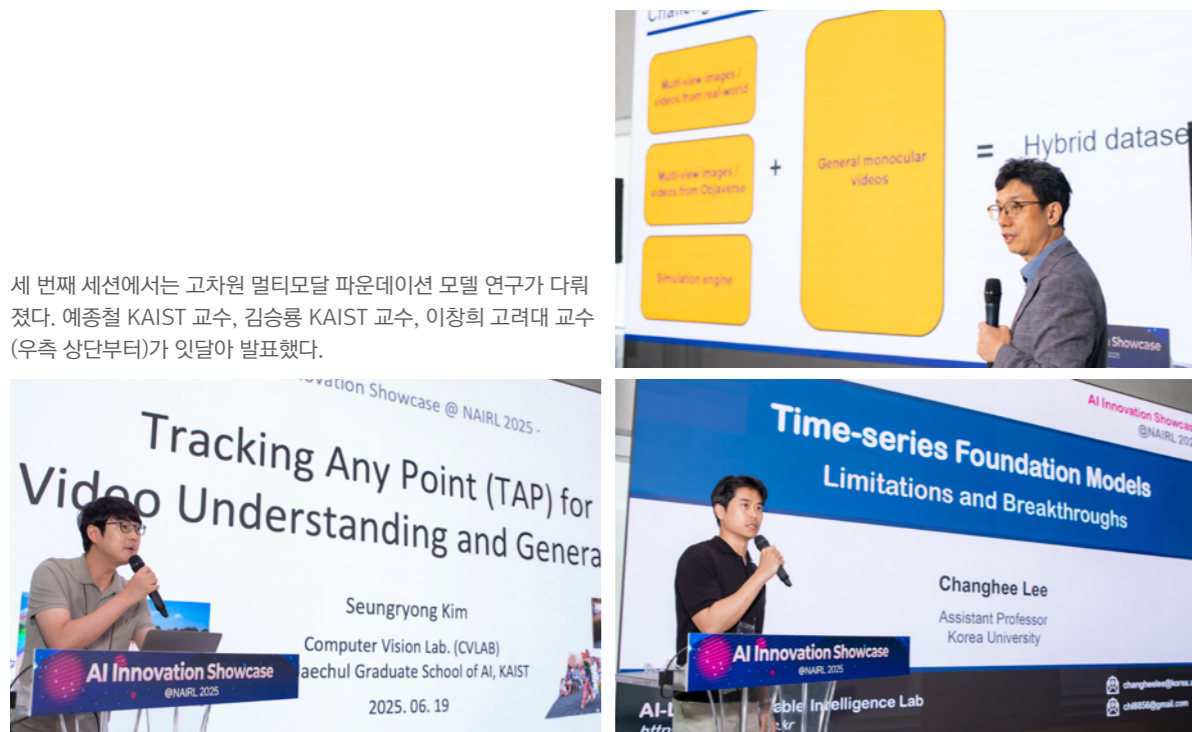
국가AI연구거점이 연구성과를 처음 공유하고 산학 협력 파트너 기업과 네트워킹을 하는 교류와 협업의 장을 마련했다. 지난 6월 19일 국가AI연구거점은 서울 서초구 서울AI허브에서 ‘국가AI연구거점 연구성과 공유 및 참여기업 네트워킹 데이’ AI Innovation Showcase’ 행사를 열고 그동안의 연구 현황을 발표했다.

이날 환영사에서 김기웅 국가AI연구거점 센터장은 “세계 최고 수준의 AI 연구로 산학 협력을 촉진하고, 국내 AI 생태계에 혁신을 일으키겠다”면서 “첨단 AI 연구성과를 산업계와 공유해 새로운 솔루션과 서비스를 창출하고 인재들이 혁신 네트워크 속에서 성장하도록 지원하는 것이 목표”라고 밝혔다. 축사에서 장은정 정보통신기획평가원 AI디지털융합단장은 “5차 산업혁명은 더 이상 선언이 아니라 현실이며



두 번째 세션은 ‘로봇 파운데이션 모델’을 주제로 진행됐다. 조민수 POSTECH 교수, 윤국진 KAIST 교수, 임재환 KAIST 교수, 최성준 고려대 교수(좌측 상단부터)가 차례대로 발표에 나섰다.

세 번째 세션에서는 고차원 멀티모달 파운데이션 모델 연구가 다뤄졌다. 예종철 KAIST 교수, 김승룡 KAIST 교수, 이창희 고려대 교수(우측 상단부터)가 잇달아 발표했다.



연구성과 발표에 이어서는 포스코홀딩스, 네이버클라우드, 코르카 등의 파트너 기업 관계자, 캡스톤파트너를 비롯한 벤처캐피털의 대표(좌측 상단부터)가 발표에 나섰다.

연세대 교수는 생성형 AI의 취약점인 과적합 문제를 다뤘는데, 디퓨전 모델의 기하학적 특성을 이용해 메모라이제이션 memorization 현상을 분석하고 이를 빠르게 감지할 수 있는 새로운 성과 지표를 제안했다.

두 번째 세션에서는 ‘로봇 파운데이션 모델’을 주제로 다양한 로봇에 언어, 영상, 행동 지능을 통합해 적용할 수 있는 범용 AI 기술이 주로 소개됐다. 먼저 2세부의 연구책임자인 조민수 POSTECH 교수가 여러 가지 상호작용 시나리오를 결합한 새로운 멀티태스크 시뮬레이션 환경을 개발하고 그 안에서 시각·언어·행동 모델이 연합되어 학습할 수 있는 모델 Mosaic3D instance segmentation model을 공개했다. 이어 윤국진 KAIST 교수는 6자유도^{6D} 프레임워크를 통해 외부 3D 모델 없이 낯선 물체의 위치와 회전을 실시간으로 예측하는 기술을 시연했고, 임재환 KAIST 교수는 로봇의 행동 지능을 강화하기 위한 시뮬레이션 기반 학습 시스템을 소개했으며, 최성준 고려대 교수는 VLA^{Vision-Language-Action} 모델에서 실패한 시연 데이터를 활용해 모델의 안정성과 능력을 향상시키는 기법을 제시했다.

세 번째 세션에서는 고차원 멀티모달 데이터를 처리하는 생성형 AI 모델 연구들이 소개됐다. 먼저 3세부의 연구책임자인 예종철 KAIST 교수는 현실 세계를 반영한 고차원 생성형 파운데이션 모델 연구에 대해 발표했다. 이어 김승룡 KAIST 교수는 비디오 생성과 이해에서 핵심 요소로 모션 기술을 지목하며 개별 프레임 속 포인트의 시공간적 움직임을 정밀하게 추적하는 모델을 제안했으며, 이창희 고려대 교수는 시계열 데이터를 위한 파운데이션 모델^{time-series foundation models}의 구축 방향을 공유했다.

연구성과 발표에 이어 파트너 기업의 관계자가 자사의 AI 기술을 소개하고 연구진과의 협업 가능성을 논의했다. HD현대 조선험양, 포스코홀딩스, LG전자, 네이버클라우드, 42MARU, 코르카^{CORCA}, (주)플레이오니, 에버엑스^{EverEx}, VESSL AI, (주)인이지, LG에너지솔루션 등의 관계자가 발표에 나섰다. 또 캡스톤파트너 송은강 대표를 비롯한 여러 벤처캐피털의 대표도 참여해 자신의 의견을 나누었다.

NAIRL

AI 프론티어 국제 컨퍼런스 2025

International Conference on AI Frontiers 2025

국가AI연구거점이 AI 기술 연구에 대한
국내외 논의부터 정책·산업 현장에 미치는 영향, 사회적 책임까지
포괄하는 **국제 공동연구 교류의 장**을 마련합니다.

일 시 2025년 10월 27일(월) 10:00~17:30

장 소 서울 용산구 서울드래곤시티 3층 한라홀, 신라홀

주 최 국가AI연구거점, 글로벌AI프론티어랩

주 관 과학기술정보통신부, 정보통신기획평가원

참석 대상 국내외 산학연 전문가, 산업계 종사자 등 500여 명

기조 연설 양 르쿤(미국 뉴욕대 쿠란트수학연구소 교수, 메타 수석AI과학자 겸 부사장),
최 예진(미국 스탠퍼드대 교수, 엔비디아 AI연구 선임 디렉터)
※ AI R&D 사업성과 홍보(포스터)도 드래곤시티 신라홀에서 진행됩니다.

대한민국 AI, 세계와 연결됩니다. 미래 기술의 심장, NAIRL



국가AI연구거점은 대한민국을 대표하는 세계적 수준의 인공지능 연구 구심점입니다.
2024년 과학기술정보통신부 및 정보통신기획평가원(IITP)의 지원으로 설립되어,
국내외 유수의 연구진이 활발히 교류하는 산·학·연·관 R&D 협력 생태계를 구축합니다.
차세대 AI 연구를 통해 새로운 기술 패러다임을 선도하고, 최첨단 AI 연구를 이끌
미래의 리더 연구자를 배출하며, 산업계 및 공공 협력을 통해 사회적 가치를 창출하는
우리나라 AI G3 기술 리더십의 핵심 플랫폼 역할을 수행합니다.

이 시대를 이끄는 기술 리더들과 함께합니다.
국내외 최고 기관·기업들과 함께 만들어갑니다.
Leading innovators. One shared mission.

