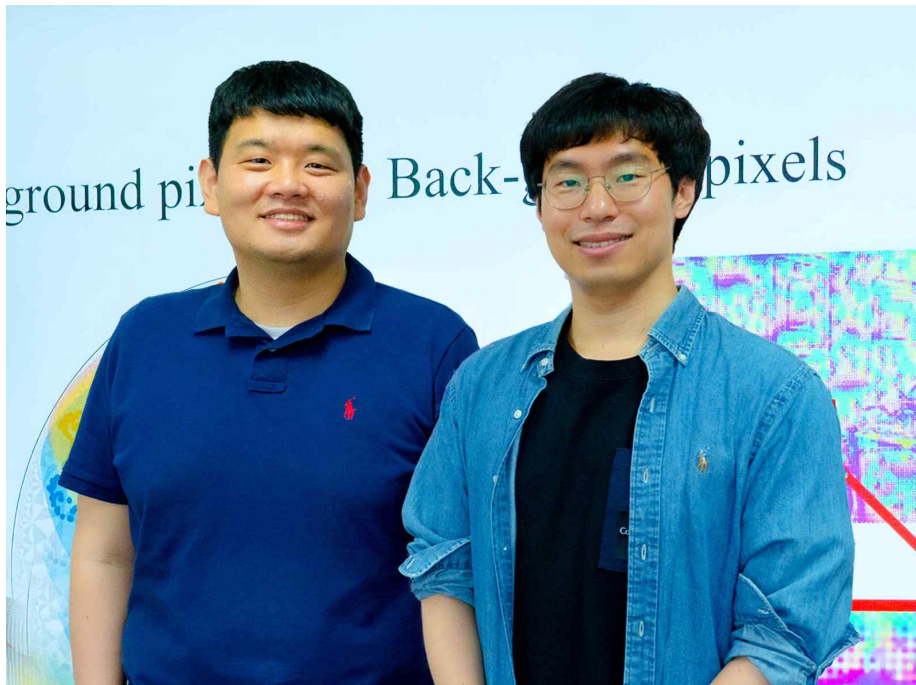


“사회적 관계까지 고려해 보행자가 어디로 향할지 예측한다”

GIST, 인간처럼 사고하는 보행자 경로 예측 AI 개발

- 거대언어모델(LLM) 기술 통해 인간의 사고방식과 유사한 사회적 추론 능력 갖춘 알고리즘 선보여... 보행자 안전 확보 필요한 자율주행 및 서비스 로봇틱스 분야에 활용 기대
- LLM이 '문자'라는 한계에서 벗어나 인간의 물리적 행동 역학을 직접적으로 예측할 수 있게 돼... “인공 일반 지능(AGI)으로의 기술 확장과 실용화 앞당길 것으로 기대”
- AI대학원 전해곤 교수팀, 세계 최고 권위의 컴퓨터비전 국제학술대회「CVPR」에서 6월 19일 발표 예정



▲ (왼쪽부터) 전해곤 교수, 배인환 박사과정생

자율주행과 배달로봇 등 서비스 로봇틱스 부문에서는 교통 법규 준수와 시민 안전 보장을 위해 보행자의 동선을 미리 파악하는 것이 필수적이다. 최근에는 **비디오 영상을 통해 보행 가능 경로와 최종 도착 위치를 추정하는 연구가 컴퓨터 비전과 로봇틱스 분야에서 큰 주목을 받고 있다.**

광주과학기술원(GIST, 총장 임기철)은 AI대학원 전해곤 교수 연구팀이 거대언어모델(LLM) 기술을 활용해 인간의 사고를 모방한 프로세스로 **정확한 보행자 경로를 예측하는 알고리즘을 개발했다고 밝혔다.**

이번 연구 성과는 보행자의 안전을 확보해야 하는 자율주행 시스템의 **보행자 회피 기술과 서비스 로봇틱스 분야에 활용될 수 있을 것으로 기대된다.**

지금까지 인공지능을 활용해 보행자의 미래 경로를 예측하는 방법론에서는 인간의 행동 역학을 수치 회귀 기법*을 통해 보행자의 위치를 모델링하여 **보행 가능 경로와 최종 도착지를 예측했다.** 이 방식은 오직 숫자만을 이용해 가장 가능성 있는 위치를 예측하므로 인간의 사고를 대변하기에는 한계가 있다.

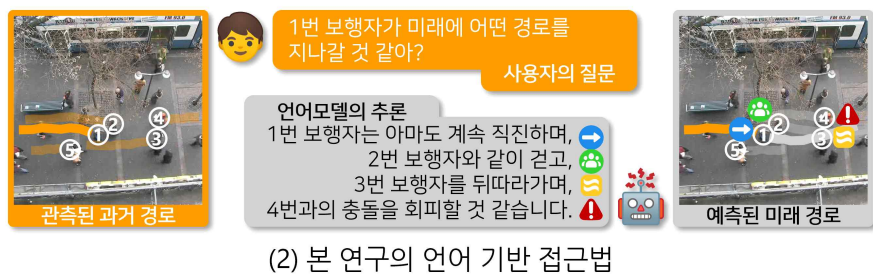
* **수치 회귀 기법(Numerical Regression)**: 주어진 데이터로부터 변수 간의 관계를 수학적으로 모델링하는 방법이다. 수학적 또는 통계적 모델을 사용하여 변수 간의 관계를 예측한다.

이를 해결하기 위해 연구팀은 거대언어모델이 가진 방대한 양의 지식을 접목해 보행자의 현재 상태와 주변 사람과의 사회적 관계를 인간처럼 분석함으로써 **훨씬 더 인간의 사고와 유사하게 미래 보행 계획을 예측하는 기술을 개발했다.**

챗GPT(ChatGPT)로 잘 알려진 거대언어모델(Large Language Model, LLM)은 인공지능의 한 종류로, 방대한 양의 텍스트 데이터를 학습하여 인간의 언어를 이해하고 생성할 수 있는 능력을 갖춘 딥러닝 기반의 모델이다.

본 연구에서는 거대언어모델이 갖는 높은 수준의 언어 이해와 생성 능력을 활용해 **인공지능이 보행 방향 및 도착지 예상, 보행자들의 집단 형성, 충돌 가능성 회피, 선행-후행 정리 등 인간의 인지**와 사회적 추론(Social Reasoning)을 할 수 있게끔 개발했다.

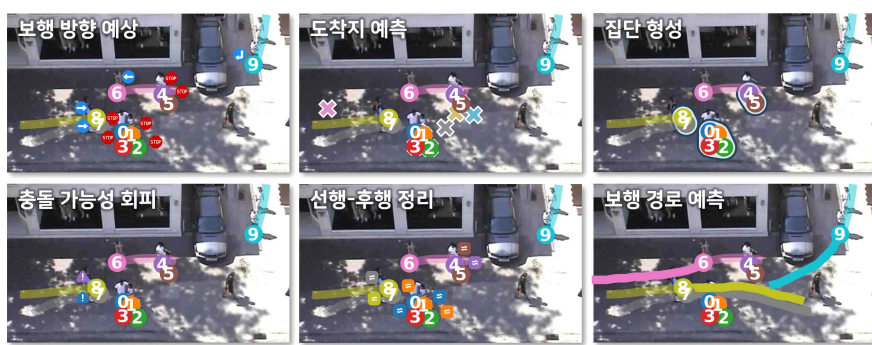
특히, 숫자만으로 인공지능이 어떠한 사유로 행동을 예측했는지 판단하는 기존 방법론과 달리 **언어모델이 직접적으로 사회적 추론 결과를 대화로 알려줄 수 있다**는 점이 이번 연구의 가장 큰 장점이다.



▲ **본 연구의 전체적인 접근법.** 숫자로 구성된 위치 좌표만을 사용하는 기존 방법론과는 달리, 거대언어모델의 지식을 활용해 인간과 같은 추론 방식을 모사하여 설명할 수 있고 정확한 미래를 예측한다.

또한 이번 연구 성과로 거대언어모델이 문자라는 한계에서 벗어나 인간의 물리적인 행동 역학을 직접적으로 예측할 수 있게 되었는데 글의 문법과 흐름을 패턴으로 인식하는 언어모델은 보행자의 한 걸음 한 걸음을 일종의 패턴으로 인지하여 다음에 발 디딜 위치를 예측한다.

이번 연구 성과로 인해 동역학에 대한 인공지능의 이해 능력은 인간이 마주하는 매 상황에서의 순간적인 사회적 추론과 결합하여 보다 더 인간처럼 사고하여 **인간의 결정과 유사하게 미래를 예측할 수 있을 것으로 기대된다.**



▲ **본 연구의 예측 결과 시각화.** 거대언어모델이 인간의 행동 데이터를 학습하여 보행 방향 및 도착지 예상, 보행자들의 집단 형성, 충돌 가능성 회피, 선행-후행 정리 등의 사회적 추론(Social Reasoning)하고, 이를 이용해 보행자가 미래에 택할 행동을 사람처럼 정확하게 예측한다.

전해곤 교수는 “이번 연구 성과는 거대언어모델이 인간의 사고방식을 모사해 사회적 관계성을 추론하며, **인간의 행동 역학을 배워 미래 동작을 예측했다는 데 학술적 의의가 크다**”며, “거대언어모델이 문자에서 더 나아가 물리 역학적 추론까지 가능하게 되면 **인공 일반 지능(AGI)으로의 기술 확장과 실용화를 앞당길 것으로 기대된다**”고 말했다.

GIST AI대학원 전해곤 교수가 지도하고 배인환 박사과정생이 수행한 이번 연구는 인공지능 분야 세계 최고권위의 국제 학술대회인 ‘CVPR(Computer Vision and Pattern Recognition Conference)’에서 오는 2024년 6월 19일 발표될 예정이다.

한편, 컴퓨터 비전 및 패턴인식 분야에서 세계 최고 학회로 손꼽히는 컴퓨터비전 및 패턴인식학회(CVPR)는 올해로 42회를 맞으며, 미국 워싱턴주 시애틀에서 개최된다. 다양한 분야의 글로벌 석학과 최고 권위자들이 한자리에 모여 기계학습, 딥러닝 등 최신 인공지능 기술과 연구 동향을 공유할 것으로 기대된다.

논문의 주요 정보

1. 논문명, 저자정보

- 저널명 : The IEEE/CVF Conference on Computer Vision and Pattern Recognition 2024 (CVPR)
 - 한국정보과학회 및 BK21+ 기준 최우수 학술대회
- 논문명 : Can Language Beat Numerical Regression? Language-Based Multimodal Trajectory Prediction
- 저자 정보 : 배인환(제1저자, GIST AI대학원 박사과정), 이준오(제2저자, GIST 전기전자컴퓨터공학부 석박통합과정), 전해곤 (교신저자, GIST AI대학원, 부교수)