# "We predict where pedestrians will head by considering social relationships." GIST develops AI that predicts pedestrian paths that thinks like a human

- Through Large Language Model (LLM) technology, an algorithm with social reasoning capabilities similar to human thinking is introduced... Expected to be used in the fields of autonomous driving and service robotics where pedestrian safety is required
- LLM can break away from the limitations of 'characters' and directly predict human physical behavior dynamics... "Expected to accelerate the expansion and commercialization of technology into artificial general intelligence (AGI)"
- Professor Hae-Gon Jeon's team from the AI Graduate School is scheduled to present on June 19 at the world's most prestigious international computer vision conference, CVPR
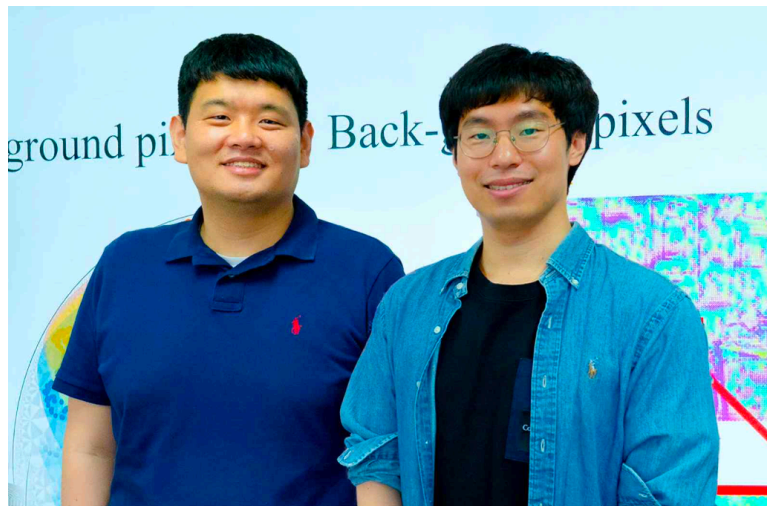
▲ (From left) Professor Hae-Gon Jeon and doctoral student Inhwan Bae

In the service robotics sector, such as autonomous driving and delivery robots, it is essential to understand pedestrian movements in advance to comply with traffic laws and ensure citizen safety. Recently, research on estimating walkable paths and final destination locations through video images is receiving great attention in the fields of computer vision and robotics.

The Gwangju Institute of Science and Technology (GIST, President Kichul Lim) announced that Professor Hae-Gon Jeon's research team at the AI Graduate School developed an algorithm that predicts accurate pedestrian paths through a process that mimics human thinking using Large Language Model (LLM) technology.

It is expected that the results of this research can be used in the field of pedestrian avoidance technology and service robotics in autonomous driving systems that must ensure the safety of pedestrians.

Until now, the methodology for predicting the future path of pedestrians using artificial intelligence has modeled the location of pedestrians using human behavioral dynamics using numerical regression techniques* to predict walkable paths and final destinations. This method uses only numbers to predict the most likely location, so it has limitations in representing human thinking.

* numerical regression: A method of mathematically modeling relationships between variables from given data by using mathematical or statistical models to predict relationships between variables.

To solve this problem, the research team applied the vast amount of knowledge possessed by the macrolanguage model to analyze the current state of pedestrians and their social relationships with people around them like a human, thereby predicting future walking plans much more similar to human thinking.

Large Language Model (LLM), such as ChatGPT, is a type of artificial intelligence model based on deep learning with the ability to understand and generate human language by learning massive amounts of text data.

In this study, the high-level language understanding and generation capabilities of the macrolanguage model were used to predict walking direction and destination, form groups of pedestrians, and avoid the possibility of collision. It was developed to enable human cognition and social reasoning, such as the preceding-following theorem.

In particular, the biggest advantage of this study is that, unlike existing methodologies that determine why artificial intelligence predicted behavior using only numbers, the language model can directly inform the results of social inference through conversation.

▲ The overall approach of this study. Unlike existing methodologies that only use location coordinates composed of numbers, the knowledge of a large language model can be used to simulate human-like reasoning methods to explain and accurately predict the future.

In addition, the results of this research enabled the macrolanguage model to break away from the limitations of text and directly predict the dynamics of human physical behavior. The language model, which recognizes the grammar and flow of writing as a pattern, recognizes each step of a pedestrian as a kind of pattern and predicts the next step.

The findings are expected to enable AI's understanding of dynamics, combined with its instantaneous social reasoning in every situation it encounters, to think more like humans and predict the future similar to human decisions.

▲ Visualization of the prediction results of this study. The macrolanguage model learns human behavior data and performs social reasoning such as predicting walking direction and destination, forming a group of pedestrians, avoiding the possibility of collision, and leading-lagging theorem, and uses this to accurately predict the actions that pedestrians will take in the future.

Professor Hae-Gon Jeon said, "This research outcome has great academic significance in that the macrolanguage model imitates human thinking to infer social relationships and learns human behavioral dynamics to predict future behavior. If the macrolanguage model goes beyond text and becomes capable of physical and mechanical reasoning, it is expected to accelerate the expansion and commercialization of technology to artificial general intelligence (AGI)."

This research, led by Professor Haegon Jeon of GIST AI Graduate School and conducted by doctoral student Inhwan Bae, will be presented on June 19, 2024, at the 'Computer Vision and Pattern Recognition Conference (CVPR)', the world's most prestigious international academic conference in the field of artificial intelligence.

Meanwhile, the Computer Vision and Pattern Recognition Society (CVPR), one of the world's top academic societies in the field of computer vision and pattern recognition, celebrates its 42nd anniversary this year and will be held in Seattle, Washington, USA. Global scholars and top authorities in various fields are expected to gather in one place to share the latest artificial intelligence technologies and research trends, such as machine learning and deep learning.

GIST
Since 1993