

Generative AI for Multimodal Visual Intelligence

2026.08.21. FRI 10:00 ~ 13:00

신공학과 161호



Yao Wei

Current

Postdoctoral Researcher EECS, QMUL

Education

University of Twente (Ph.D.)

Research Interests

Multimodal Learning, Generative Models
Vision-Language Models, Scene graphs
Embodied visual intelligence

Experience

AWS AI Labs (Intern)
IIT & Univ.of Bath (Visiting Researcher)

Abstract

Generative AI is still reshaping visual content. However can it serve as a paradigm for multimodal intelligence? Yao explores this in two directions:

Open-vocabulary perception with flow matching, using generative dynamics to improve vision-language alignment in object detection.

Diffusion-based 3D indoor scene generation, while utilizing structured semantic reasoning and physical constraints.

Ultimately, these advance a broader vision of multimodal intelligence, extending toward future embodied perception and interaction.