(국문/영문)이름: 김원화 / Kim, Won Hwa
(국문/영문)직위: 부교수 / Associate Professor
(국문/영문)소속: 포항공과대학교 / POSTECH

국문 강연제목: 강건한 의료영상 AI 를 위한 제어 가능한 적대적 공격

영문 강연제목: Controllable Adversarial Attacks for Robust Medical Imaging AI

Abstract(영문):

Deep learning has achieved remarkable success in medical image analysis, yet it suffers from lack of data for real-world deployment. In this talk, I will present a series of studies from our group that leverage adversarial perturbations —originally studied as a security threat— can be reinterpreted as a powerful tool for data augmentation and robustness enhancement to improve the reliability of medical imaging AI. These approaches demonstrate how controllable adversarial perturbations can transform small changes into significant gains in model performance, from classification to segmentation, paving the way toward robust AI in clinical imaging.

Brief Biosketch (간단한 이력, 연구/대외활동 소개,국문/영문)

Won Hwa Kim is an Associate Professor in Graduate School of Artificial Intelligence (GSAI) / Computer Science and Engineering (CSE) / Medical Science and Engineering (MED) at Pohang University of Science and Technology (POSTECH). Prior to joining POSTECH, he was an Assistant Professor in Computer Science and Engineering at the University of Texas at Arlington (2018 – 2023, last 2 years on leave-of-absence), and he was a Researcher in Data Science Team at NEC Labs., America (2017-2018). He obtained his Ph.D. in Computer Sciences from University of Wisconsin-Madison in 2017, M.S. in Robotics from KAIST (2010) and B.S. in Electrical Engineering from Sungkyunkwan University (2008). He developed Hybrid Vehicles at Hyundai Motors Company in 2010-2011 before he dived into Artificial Intelligence. He is a recipient of prestigious grants such as NSF CISE CRII (2020) in the U.S., NRF Mid-Career Researcher Program (2021) and Basic Research Lab (2025) in South Korea.