JAEHUI HWANG

jaehui.hwang@yonsei.ac.kr LinkedIn
https://J-H-Hwang.github.io Github

RESEARCH INTERESTS

I am interested in exploring the properties of various deep learning models to understand neural networks (NN) better. In particular, I have focused on analyzing the vulnerability of models and what elements including architectural components and input data contribute to the diverse behavior of models. Based on these interests, I aim to contribute to robust and reliable AI systems for diverse tasks and real-world scenarios. To be specific, I have considered the following research areas:

- **Reliability.** The development of neural networks has led to various technical improvements, especially in vision-related areas. However, several vulnerabilities of NNs have been pointed out, such as shortcut learning problems and non-robustness against adversarial attacks. I am interested in the vulnerability of recent NNs and design strategies that could improve their robustness. I have analyzed the weaknesses of action recognition models from the perspective of temporal modeling.
- **Model understanding.** Neural architectures trained for the same task can have different properties, such as distinct representation spaces, due to variations in their architectural components and input data. I have developed a similarity measurement between different neural architectures for a better understanding of them. On the other hand, by exploiting models' properties, I have designed an evaluation metric of generative models.
- **Adversarial Attack/Defense.** Adversarial attack is one of the main problems when we use AI in practical scenarios. It causes the unreliability of the AI system. I have developed adversarial attack and defense methods for human action recognition models.

EDUCATION

Ph.D. candidate in Integrated Technology from **Yonsei University** Mar 2019 - Present
Advisor: Prof. Jong-Seok Lee

B.S. in Integrated Technology from **Yonsei University** Mar 2016 - Feb 2019

PUBLICATIONS (SELECTED)

1. **Jaehui Hwang**, Huan Zhang, Jun-Ho Choi, Cho-Jui Hsieh, Jong-Seok Lee, "*Temporal Shuffling for Defending Deep Action Recognition Models against Adversarial Attacks*," Neural Networks, 2024 (accepted).
2. Hojung Lee, **Jaehui Hwang**, Hwin Dol Park, Jaehun Choi, Jong-Seok Lee, "*Classifying Gas Data Measured Under Multiple Conditions Using Deep Learning*," IEEE Access, 2022.
3. **Jaehui Hwang**, Jun-Hyuk Kim, Jun-Ho Choi, Jong-Seok Lee, "*Just One Moment: Structural Vulnerability of Deep Action Recognition against One Frame Attack*," IEEE/CVF International Conference on Computer Vision (**ICCV**), 2021.
4. **Jaehui Hwang**, Seong-Eun Moon, Jong-Seok Lee, "*On the Repeatability of EEG-based Image Quality Assessment*," IEEE International Conference Conference on Systems, Man, and Cybernetics (SMC), 2018. `https://github.com/clovaai/lffont`

Under Review

1. **Jaehui Hwang**, Dongyoon Han, Byeongho Heo, Song Park, Sanghyuk Chun*, Jong-Seok Lee*, "*Similarity of Neural Architectures using Adversarial Attack Transferability.*"

2. **Jaehui Hwang***, Junghyuk Lee*, Jong-Seok Lee, "*Anomaly Score: Evaluating Generative Models and Individual Generated Images based on Complexity and Vulnerability.*"

RESEARCH EXPERIENCES

Research Intern Mar 2022 - Sep 2022

Naver AI Lab, ML Research team (*Mentor: Sanghyuk Chun*)

- Exploring what architectural components contribute to the diversity of deep learning models
- Designing the metric that computes the architectural similarity

Research Intern Sep 2018 - Mar 2019

Multimedia Computing and Machine Learning Group, Yonsei University (*Professor: Jong-Seok Lee*)

- Analyzing the repeatability of EEG signals on image quality assessment

Research Intern Mar 2018 - Sep 2018

Intelligent Unmanned Systems Group, Yonsei University (*Professor: Jiwon Seo*)

ACADEMIC ACTIVITIES

Award

- Merit Academic Paper Award (2022-2 Yonsei Superior Paper Awards) Dec 2022

Reviewer

- CVPR 2022-2023, ICCV 2022-2023, TMLR 2023

Talks

- "Just One Moment: Structural Vulnerability of Deep Action Recognition Against One Frame Attack", Korean Conference on Computer Vision (KCCV), 2022.

PROJECTS

Korea Scholar's Conference for Youth Sep 2017 - Mar 2020

Position: Deputy Secretary General and Database Coordinator

- Organized and directed Asia's largest youth conference, which drew **over 6000 people**
- Planned and launched a **new dissertations' database system** that facilitated communication between mentors and facilitators for the dissertation review of 800+ applicants

SKILLS

Programming Languages & Frameworks (Selected)

- Programming Language: Python, Matlab, C
- Machine learning tools: PyTorch, Tensorflow, OpenCV, NumPy, Scikit-learn

SCHOLARSHIPS

Full scholarship for Graduate School Mar 2019 - Feb 2020

Institute for Information and Communications Technology Promotion (IITP)

Full scholarship for Undergraduate School Mar 2016 - Feb 2019

Institute for Information and Communications Technology Promotion (IITP)

TEACHING EXPERIENCES

Teaching Assistant (Yonsei University)

- Signals Sep 2019 - Dec 2019
- Understanding of integrated technology Sep 2023 - Dec 2023