

SUMYEONG AHN

9th Building Room 9514, 85, Hoegi-ro, Dongdaemun-gu, Seoul, Republic of Korea (02455)

(+82)10-9236-0343, sumyeongahn@kaist.ac.kr

EDUCATION

Ph. D. Candidate

Sep. 2017 - Present (Expected Aug. 2023)

Korea Advanced Institute of Science and Technology, S.Korea

Kim Jaechul Graduate School of Artificial Intelligence

Advisor: Prof. Seyoung Yun (Co-advisor: Prof. Yung Yi @ KAIST EE)

Master of Science

Mar. 2015 - Aug. 2017

Korea Advanced Institute of Science and Technology, S.Korea

Department of Electrical Engineering

Thesis: Multi-armed Bandit Problem with Intra- and Inter- Correlations

Advisor: Prof. Yung Yi

Bachelor of Science

Mar. 2011 - Feb. 2015

Korea University, S.Korea

School of Electrical Engineering

RESEARCH INTERESTS

- Noisy label dataset [Submitted to ICML 2023]
- Class imbalance problem [ICLR 2023]
- Dataset Bias [ICLR 2023, AAI 2023, Preparing ICCV 2023]
- Unsupervised Domain Adaptation with GANs. [ICPR 2020]
- Stochastic Multi-armed Bandits with Additional Conditions [SIGMETRICS 2018]
- AI for Computer Networking Systems [SIGMETRICS 2018, MobHoc 2021]

PUBLICATIONS (RECENT PAPER FIRST)

Preprint

- (P1) **Sumyeong Ahn** and Se-Young Yun. Orbis: Open dataset can rescue you from dataset bias problem. Preparing.

Conference

- (C1) **Sumyeong Ahn***, Seongyeon Kim*, and Se-Young Yun. Mitigating dataset bias by using per-sample gradient. 2023a. ICLR'23 To appear.
- (C2) **Sumyeong Ahn***, Jongwoo Ko*, and Se-Young Yun. CUDA: Curriculum of data augmentation for long-tailed recognition, 2023b. ICLR'23 Spotlight, To appear.
- (C3) **Sumyeong Ahn** and Se-Young Yun. Denoising after entropy-based debiasing a robust training method for dataset bias with noisy labels. 2023.
- (C4) Sangwoo Moon, **Sumyeong Ahn**, Kyunghwan Son, Jinwoo Park, and Yung Yi. Neuro-dcf: Design of wireless mac via multi-agent reinforcement learning approach. In *22nd ACM International Symposium on Theory, Algorithmic Foundations, and Protocol Design for Mobile Networks and Mobile Computing (MobiHoc'21)*, 2021.

- (C5) Tran Hai H, **Sumyeong Ahn**, Taeyoung Lee, and Yung Yi. Enlarging discriminative power by adding an extra class in unsupervised domain adaptation. In *2021 25th International Conference on Pattern Recognition (ICPR)*, 2021.
- (C6) Donggyu Yun, **Sumyeong Ahn**, Alexandre Proutiere, Jinwoo Shin, and Yung Yi. Multi-armed bandit with additional observations. In *2018 ACM International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS 2018, Beckman CenterIrvine, United States, 18 June 2018 through 22 June 2018*, pages 53–55. Association for Computing Machinery (ACM), 2018a.

Journal

- (J1) Donggyu Yun, **Sumyeong Ahn**, Alexandre Proutiere, Jinwoo Shin, and Yung Yi. Multi-armed bandit with additional observations. *Proceedings of the ACM on Measurement and Analysis of Computing Systems*, 2(1):1–22, 2018b.

Workshop

- (W1) Jongwoo Ko*, **Sumyeong Ahn***, and Se-Young Yun. Efficient utilization of pre-trained model for learning with nosiy labels. 2023. ICLR’23 Trustworthy Workshop.
- (W2) **Sumyeong Ahn***, Jongwoo Ko*, and Se-Young Yun. CUDA: Curriculum of data augmentation for long-tailed recognition. In *NeurIPS ML Safety Workshop and NeurIPS 2022 Workshop on Distribution Shifts: Connecting Methods and Applications*, 2022a. URL <https://openreview.net/forum?id=5yrzySCf1M1>.
- (W3) **Sumyeong Ahn***, Seongyeon Kim*, and Se-Young Yun. Mitigating dataset bias by using per-sample gradient. In *NeurIPS ML Safety Workshop and NeurIPS 2022 Workshop on Distribution Shifts: Connecting Methods and Applications*, 2022b. URL <https://openreview.net/forum?id=wMCCObWq-FI>.

Patent

1. Nansol Seo, Jaemoon Lee, Jaeha Ahn, **Sumyeong Ahn**, Suho Shin, Yoonpyo Koo, and Yung Yi. Method for controlling multiple uavs based on multi-hop wireless mesh networks, September 3 2019. KOR patent number: 10-2019874-0000.

MISC (Korean conference and journal)

1. Sangmin Bae, Taehyeon Kim, **Ahn, Sumyeong**, Sangmook Kim, Jongwoo Ko, and Se-Young Yun. Client sampling algorithm in federated learning via combinatorial averaging and multi-armed bandits. *KIISE*, pages 1088–1090, 2022.
2. Yoonpyo Koo, **Sumyeong Ahn**, Kyoungwhan Son, Suho Shin, Jeonghun Yu, Jaesin Kim, and Yung Yi. An implementation of multi-hop voice communication system using drones. *KICS*, pages 220–221, 2017.

PROJECT

Student Leader

- Developing NAS algorithm for layer-wise communication network failure detection utilizing multi-layer datasets.

KT (08/2021-08/2022)

- **Keyword:** Neural Architecture Search (NAS) , Comm. Net. failure detection, Tabular dataset
- **Abstract:** Although the communication networking hardware is geographically dispersed across a large area, the failure information it spews calls for a quick fix or a negligible failure due to a

temporary malfunction. Additionally, the information in the networking system is continuously changing (such as the occurrence of a new log type) and accumulating. In this project, utilizing networking log data (tabular and topological data), we create a Neural Architecture Search (NAS) model to reduce the burden of AI engineer when creating a model that determines if a breakdown needs to be resolved directly by employees.

- Application service implementation (Store open/close prediction) based on data from commercial stores' power consumption

KEPRI (04/2018-12/2019)

- **Keyword:** Time-series data, 1D-CNN, RNN, Power consumption data, Open/Close prediction
- **Abstract:** To promote customer convenience, a number of platform providers, including Google, Naver, offer Open/Close information alongside information from each commercial stores. However, because of unusual actions (such as vacation), information inaccuracies can make users less comfortable. We created a method to predict Open/Close using time series data from smart meters in order to avoid this inconvenience. The like provided below will take you to this service **POWERON**.

- Versatile Network System Architecture for Multi-dimensional Diversity

IITP (04/2016-08/2019)

- **Keyword:** Edge Networking, Fog Computing, FogOS, Matching between resources and requests
- **Abstract:** To fully take advantage of the proliferation of high-speed terminal devices and enhance the performance of cloud computing, edge networking (also known as Fog Computing) is being developed. Performance can be improved by using a module that effectively matches request and resources because resources are not as high performance computing equipment as cloud servers. To facilitate adequate cooperation among edge devices, we created FogOS, a platform that offers services by managing resources and requests.

- Designing Matching Algorithm for Distributed Edge Devices

KISA (06/2015-06/2016)

- **Keyword:** Matching algorithm, FogOS, Matcher
- **Abstract:** A sub-project linking resources and requests on the edge devices was performed.

- Modeling of Multi-hop Wireless Networks for Military Applications

ADD (09/2015-12/2016)

- **Keyword:** Flying Ad-Hoc Network (FANET), Ground Control Station (GCE), Voice Communication, UAVs, Drone
- **Abstract:** The Line of Sight (LoS) signals from flying objects (UAVs, drones, etc.) are strong, but they have drawbacks such short stay times and high shaking. In this project, we aimed to improve the success rate of the military or disaster situation mission by employing these drones to generate an ad-hoc voice call system. To do this, we create ground control station to manage a huge number of drones and voice call system which covers of 1Km.

Participating Student

- Video Transmission over Vehicular Networks

Bosch (01/2015-06/2015)

- **Keyword:** Vehicular Ad-Hoc Network (VANET), 802.11, WiFi, Video codec, real-time video communication, MPEG, JMPEG
- **Abstract:** Strong linearity characterizes vehicular communication, but significant mobility is also present. In this project, we looked at numerous codecs and conducted a study to build a

networking protocol that transfers images from the front car to the rear car. We discovered that it is most effective to communicate between cars to limit the re-transmission as much as possible since dropped frame becomes meaningless in real-time video communication and uniformizing the importance of each frame (e.g., do not use i-frame in MPEG) because one informative dropped frame can affect other frames.

AWARDS

- KAIST EE Best Teaching Assistant Award (2019 Fall - Embedded System)

SKILLS

- Programming Languages & Frameworks (Selected)

- Machine learning tools: PyTorch, Tensorflow, Scikit-learn
- Database: MySQL, Oracle Database
- System Programming: Bash, C/C++
- Others: OpenCV, Numpy, Matlab

TEACHING

- Teaching Assistance (@ KAIST)

- EE209 Programming Structures for Electrical Engineering (2015 Spring, 2015 Fall, 2020 Spring)
- EE323 Computer Networks (2016 Spring)
- EE205 Data Structures and Algorithms for Electrical Engineering (2017 Fall)
- HSS190(J) Freshman Seminar 1 (2018 Spring)
- EE414 Embedded Systems (2016 Fall, 2019 Fall)
- AI505 Optimization for AI (2020 Fall)
- AI603 Machine Learning Theory (2021 Spring)

COURSES

- Machine Learning

- (EE837) Deep Learning for Computer Vision (Prof. Junmo Kim)
- (AI604) Deep Learning for Computer vision (Prof. Jaegul Choo)
- (AI612) Machine Learning for Healthcare

- Networking

- (EE655) Economics in Communication Networks
- (EE827) Epidemics and Information Diffusion in Complex Networks
- (EE827) 3G/4G/5G Cellular Network Protocols
- (EE650) Optimization in Communication Networks
- (EE627) Performance Analysis of Communication Networks
- (EE827) Network Science
- (EE528) Engineering Random Processes

- Math

- (MAS241) Analysis 1
- (IE539) Convex Optimization

- System

- (EE516) Embedded Software
- (EE324) Network Programming

Last update: 2023/03/06 06:37:33