PH.D. JEMIN LEE

leejaymin@etri.re.kr \diamond Homepage Electronic Telecommunication Research Institute (ETRI) \diamond Gajeong-ro 218, Yuseong-gu, Daejeon, 34130, Republic of Korea

RESEARCH INTEREST

Systems for Machine Learning and AI

AI Compiler to generate optimized kernel codes for Neural Processing Units(NPUs) DNN Compression: Quantization, Pruning

Applied Machine Learning

Interruptibility Management in Multiple Mobile Devices Optimizing Deep Neural Networks for Embedded Systems

Mobile Systems

Automated Power Model Generation for Smartphones **Energy Optimization for Smartwatches**

WORK EXPERIENCES

AI Research Laboratory, ETRI Daejeon, South Korea Senior Researcher Dec. 2018 – Present **Interactive Computing Laboratory, KAIST** Daejeon, South Korea Post-doctoral Researcher (including compulsory military service period) Nov. 2017 - Aug. 2018 Superviosr: Uichin Lee

EDUCATION

Chungnam National University Ph.D. in Department of Computer Science and Engineering Embedded System Laboratory Thesis: Power Modeling, Analysis, and Optimization for Mobile Devices Advisor: Hyungshin Kim Outstanding Ph.D. Thesis Award (top 1 out of 115)

Chungnam National University

B.S. in Department of Computer Science and Engineering

RESEARCH EXPERIENCE

Period: Jun. 2023 -Development of open edge AI SoC hardware and software platform Dec. 2027

Research Staff

· Develop a modular NPU and open-edge AI SoC prototype auto-generator, and develop an open-edge AI SoC platform that can optimize hardware and software according to new edge AI application requirements. The results of this research and development project will be used for the training of hardware and software personnel required in the field, with the aim of creating an open ecosystem through public disclosure.

Technology development of smart edge device SW development platform Period: Apr. 2022 – Dec. 2026 Research Staff

Last Updated September 2, 2023

Grant: \$1M for 5 years

Sept. 2011 - Aug. 2017

Mar. 2006 - Aug. 2011

Grant: \$14M for 5 years

• In this project, the goal is to develop a compiler that supports four types of NPU developed in Korea and a SaaS-based platform. Specifically, I am responsible for developing quantization and compiler support for various NPUs, including 4-bit, 8-bit, and FP16 quantization, as well as state-of-the-art models like Transformers.

Neuromorphic Computing Software Platform for Artificial Intelligence Systems Period: Dec. 2018 – Dec. 2022

Research Staff

 \cdot In the project, my role is to implement a deep learning compiler called NEST-C for a hardware accelerator. The NEST Compiler (NEST-C) is an open source project led by ETRI, which is based on GLOW project. Glow is a machine learning compiler and execution engine for hardware accelerators. It is designed to be used as a backend for high-level machine learning frameworks. The compiler is designed to allow state of the art compiler optimizations and code generation of neural network graphs. The objective of NEST-C is to generate optimized code for various kinds of Neural-network Processing Uints (NPUs). Therefore, NEST-C provides automatic tuning functionalities and tools for each optimization step.

SuggestBot: Development of a context-based smart interaction service platform Period: Sept. 2017 – Oct. 2018t

Research Assistant

• To train SuggestBot core engines and enabling context-based association/suggestion applications, this sub-project aims at collecting conversation-based interaction and context big data as well as mobile/wearable sensor and interaction big data. We develop (1) novel crowdsourcing techniques and open crowdsourcing platforms for conversational interaction and context data annotation; and (2) mobile/wearable sensor and interaction data collection SW (e.g., bio signals, wearable/mobile interaction data, speech data, image/environment/context information).

Mobile context sensing platform study using smart gadgets Period: Nov. 2014 – Apr. 2017 Research Assistant Grant: 39,000,000 KRW (per year)

• This research project aims to develop a software platform that supports continuous sensing of mobile context information from a smartphone and a wearable device. It is possible to extract more accurate and complicated context information by using a wearable device such as a smart watch recently emerged. The continuous sensing for extracting context information can become energy efficient through data fusion and inference. With the extracted context information, we can optimize conventional OS services like the scheduler and memory manager.

The Development of Core Technologies for Zone-based Services using Beacons Period: Oct. 2014 – Sep. 2017

Research Assistant

• This research aims to develop room-level activity detection in business areas such as shop, office, and home using BLE-Beacon as a next-generation indoor location based service. The proposed method includes heterogeneous data communication, Beacon authentication, and occupant detection.

Energy Analysis and Optimization for Smartphone Research Assistant

• Since the smartphone users are dramatically increasing, efforts to reduce power consumption and extend the phone lifetime are becoming more important. To reduce energy consumption, it is necessary to analyze application's power consumption in fine grain. A power profiler is the tool for the purpose. Though, until now mobile applications have been developed without supports from the tool, in near future, applications with low energy efficiency will not be selected by users. In this research, we study methods to optimize power consumption of smartphone applications.

PUBLICATIONS

Grant: 150,000,000 KRW (per year)

Grant: 750.000.000 KRW (per year)

Grant: 3.000.000.000 KRW (per year)

Period: April 2011 – May 2014

Grant: 40,000,000 KRW (per year)

Peer-reviewed Journals and Proceedings

 $\textbf{C.11} \ \ \textbf{Q-HyViT: Post-Training Quantization for Hybrid Vision Transformer with Bridge Block Reconstruction}$

Jemin Lee, Yongin Kwon, Jeman Park, Misun Yu, Hwanjun Song preprint at Arxiv 2023

J.10 PartitionTuner: An operator scheduler for deep-learning compilers supporting multiple heterogeneous processing units

Misun Yu, Yongin Kwon, **Jemin Lee**, Jeman Park, Junmo Park, Taeho Kim ETRI Journal Vol 45 Issue 2 pp. 187-357, Apr 2023 (JCR21 IF: 1.622) ISSN: 1225-6463, doi: https://doi.org/10.4218/etrij.2021-0446

J.9 Software-level Memory Regulation to Reduce Execution Time Variation on Multi-core Real-time Systems

Sihyeong Park, Jemin Lee, Hyungshin Kim IEEE Access, Vol. 10, pp.93799-93811, Sept. 01, 2022 (JCR21 IF: 3.476) ISSN: 2169-3536

C.10 CPrune: Compiler-Informed Model Pruning for Efficient Target-Aware DNN Execution

Taeho Kim, Yongin Kwon, **Jemin Lee**, Taeho Kim, Sangtae Ha, European Conference on Computer Vision (ECCV'), pp.651–667, Oct 23-27, 2022. BK-IF 2, Acceptance Rate 28% (1,650 papers accepted out of 5,803 submitted)

J.8 Time-Invariant Features-Based Online Learning for Long-Term Notification Management: A Longitudinal Study

Jemin Lee_i/strong¿, Sihyeong Park, Taeho Kim, Hyungshin Kim Applied Sciences Vol. 12, No. 11 Article-Num. 5432, June 01, 2022 (JCR21 IF: 2.838, ISSN: 2076-3417

- J.7 Quantune: Post-training quantization of convolutional neural networks using extreme gradient boosting for fast deployment
 Jemin Lee*, Misun Yu, Yongin Kwon, Taeho Kim
 *corresponding author
 Future Generation Computer Systems 2022, Vol. 132, 2022, pp. 124-135 (impact factor: 7.187, JCR21: Top 5.9%, computer science, theory & method rank 7/110), ISBN: 0167-739X
- **J.6** PASS: Reducing Redundant Notifications between a Smartphone and a Smartwatch for Energy Saving

Jemin Lee, Uichin Lee Hyungshin Kim

IEEE Transactions on Mobile Computing, Vol 19, Issue 11, 1 Dec. 2020, (SCI, Top 9% (4.474) impact factor in JCR computer science & information systems category rank 14/155).

- J.5 Hardware Resource Analysis in Distributed Training with Edge Devices Sihyeong Park, Jemin Lee, Hyungshin Kim MDPI Electronics 2020, 9(1) 28, 26 Dec. 2019 (impact factor: 1.764)
- C.10 Fire in Your Hands: Understanding Thermal Behavior of Smartphones

Soowon Kang, Hyeonwoo Choi, Sooyoung Park, Chunjong Park, **Jemin Lee**, Uichin Lee, and Sung-Ju Lee,

In Proceedings of the ACM International Conference on Mobile Computing and Networking (Mobi-Com) 2019.

J.4 Reducing Smartwatch Users' Distraction with Convolutional Neural Network

Jemin Lee, Jinse Kwon, Hyungshin Kim (SCIE) Mobile Information Systems, vol. 2018, Article ID 7689549, 9 pages, 15 Mar. 2018 (special issue in Advances in Personalized Mobile Services). (SCIE, impact factor: 0.849)

- C.9 Analysis of Hardware Resources in Distributed Learning (poster) Sihyeong Park, Jemin Lee, Hyungshin Kim In Proceedings of International Workshop on Highly Efficient Neural Networks Design (co-located with EMSOFT), pp. 1-4, Seoul, South Korea, Oct. 2017.
- C.8 An Ultrasound-based Indoor Localization Using Gaussian ASK Modulation (WIP) Jinse Kwon, Jemin Lee, Hyungshin Kim In Proceedings of International Conference on Indoor Positioning and Indoor Navigation, pp. 1-4, Sapporo, Japan, 18-21 Sept. 2017.
- C.7 Deep Learning Training on Distributed Embedded Systems (poster) Sihyeong Park, Jemin Lee, Hyungshin Kim In Proceedings of the 12th IEMEK Symposium on Embedded Technology, Busan, South Korea, 18-19 May, 2017.
- C.6 Extending App Pre-Launch Service with Emotion Context (poster) Jinyoung Choi, Jemin Lee, Hyungshin Kim
 In Proceedings of the 2nd ACM/IEEE International Conference on Internet-of-Things Design and Implementation (IoTDI'17) Adjunct, pp. 1-2, Pittsburgh, USA, 18-21 Apr. 2017.
- J.3 QDroid: Mobile Application Quality Analyzer for App Market Curators
 Jemin Lee, Hyungshin Kim
 Mobile Information Systems, vol. 2016, Article ID 1740129, 11 pages, 10 Oct. 2016. (SCIE, impact factor: 1.462)
- C.5 Reducing Distraction of Smartwatch Users with Deep Learning

Jemin Lee, Jinse Kwon, Hyungshin Kim In Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI'16) Adjunct pp. 948-953, Florence, Italy, Sept. 2016.

- J.2 O-Sleep : Output-Oriented Power Saving Mode for Smartphones Hyunwoo Joe, Jungseok Kim, Jemin Lee, Hyungshin Kim Future Generation Computer Systems-The International Journal of eScience, 6 Jun. 2016, ISSN 0167-739X.
 (SCIE, Top 10% (2.430) impact factor in JCR Theory&Methods: Category Rank 11/150)
- J.1 Automated Power Model Generation Method for Smartphones
 Jemin Lee, Hyunwoo Joe, Hyungshin Kim
 IEEE Transactions on Consumer Electronics, Vol. 60(2), pp. 190-197, May, 2014. (SCI, impact factor: 1.045)
- C.4 Framework for automated power estimation of Android applications (poster) Jemin Lee, Hyungshin Kim International conference on Mobile systems, applications, and services (Mobisys'13), Taipei, Taiwan, pp. 541-542, Jun. 2013.

- C.3 Energy Reservation Service for Smart Phone Application (poster) Vincent Dupre, Jaymin Lee, Hyungshin Kim 3rd ACM/SIGOPS Asia-Pacific Workshop on Systems (ApSys'12) Seoul, South Korea 23-24th, July, 2012.
- C.2 Smart Phone Power Model Generation Using Use Pattern Analysis

Jaymin Lee, Hyunwoo Joe, Hyungshin Kim IEEE International Conference on Consumer Electronics(ICCE'12) Las Vegas, NV, USA 13th-16th Jan 2012.

C.1 Smartphone, where does the power go?

Jaymin Lee, Hyunwoo Joe, Hyungshin Kim EU Korea Conference on Science and Technology (EKC'11) Paris, France, 21-23th, July 2011.

ACADEMIC SERVICES

Chair

 \cdot Web co-chair for ACM MobiSys 2019

Board of Directors

 \cdot Institute of Embedded Engineering of Korea (IEMEK), 2022, 2023

External Reviewer

- · Pervasive and Mobile Computing 2014.
- · IEEE Transactions on Mobile Computing 2015.
- $\cdot\,$ Journal of Medical Internet Research 2018.
- · Sustainable Computing, Informatics and Systems 2018.
- · IMWUT (UbiComp), May 2018.
- · IMWUT (UbiComp), Sept. 2018.
- \cdot IEEE SCC 2019.
- $\cdot\,$ CHI 2021.
- · MDPI Applied Science 2022 (Feb. Mar.)
- \cdot MDPI Electronics 2022 (Jan. Feb.)

HONORS AND AWARDS

KSC Best Paper Award	2019
\cdot The Korean Institute of Information Scientists and Engineers.	
IEMEK 2017 Best Presentation Award	2017
\cdot Korean Embedded Engineering Conference 2017, Institute of Embedded Engineering of Korea	ì.
Outstanding Ph.D. Thesis Award (top 1 out of 115)	2017
· Chungnam National University.	
Embedded System Design Challenge Bronze Award (out of 28 teams)	2017
\cdot Faster R-CNN Optimization for Embedded System, ACM SIGDA KOREA Chapter 2017.	
IEMEK 2015 Best Presentation Award	2015
\cdot Korean Embedded Engineering Conference 2015, Institute of Embedded Engineering of Korea	ì.
KSCI 2015 Best Paper Award	2015
\cdot Korea Society of Computer Information 2015, The Korea Society of Computer Information.	

KCC 2015 Best Paper Award	2015
\cdot Korea Computer Congress 2015, The Korean Institute of Information Scientists and Engineers	3.
Participation award	2015
\cdot Graduation Contest 2015, Chungnam National University.	
Best Paper Award	2014
\cdot Korea Computer Congress 2014, The Korean Institute of Information Scientists and Engineers	3.
Best Presentation Award	2012
\cdot Korea Computer Congress 2012, The Korean Institute of Information Scientists and Engineers	3.
Participation Award	2011
 Creative Work Contest 2011, Department of Computer Science and Engineering, Chungnam Nat University. 	tional

ISSUED PATENTS

Method and system for expecting users' mood based on status information and biometric information acquired by using user equipment Granted 06/15/2017, Korea Patent number 10-1749706 Hyungshin Kim, Jemin Lee, Jinyoung Choi

Method for Detecting Indoor Zone with Bluetooth and Ultrasound of Smartphone Granted 05/29/2017, Korea Patent number 10-1742960 Hyungshin Kim, Jemin Lee, Jinse Kwon

System and Method for Detecting Beacon Granted 05/24/2017, Korea Patent number 10-1741406 Hyungshin Kim, Jemin Lee, Seula Hwang

Portable terminal and method for controlling a battery charging of the same Granted 08/16/2016, Korea Patent number 10-1650038000 Hyungshin Kim, **Jemin Lee**, Donggeon Han

Search system and method of executable GUI Granted 04/20/2015, Korea Patent number 10-1513662000 Hyungshin Kim, Jemin Lee, Donggeon Han

Collaborative Power Model Creation Method and Service Module With the Same Granted 02/26/2013, Korea Patent number 10-12669710000 Hyungshin Kim, **Jemin Lee**