

ISSN1016-9288

제51권 12호

2024년 12월호

전자공학회지

The Magazine of the IEIE

vol.51. no.12

AI 반도체를 위한 하드웨어-소프트웨어 공동 경량화 기술

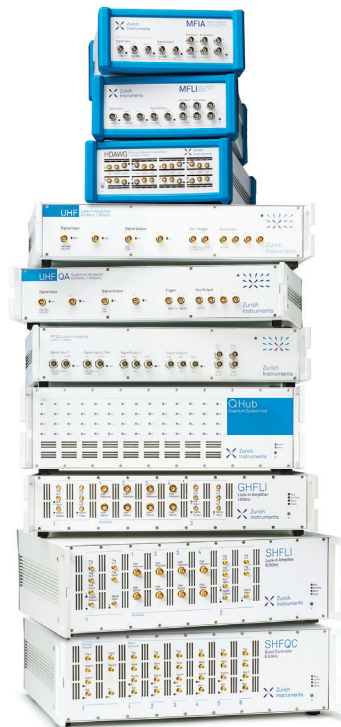
- AI 반도체를 위한 SW/HW 연구 동향
- 영상처리를 위한 경량화 뉴럴프로세서 설계
- AI 반도체를 위한 TinyML 기술 동향
- 플래시 메모리 기반의 인공지능 반도체
- 반도체 설계를 위한 AI/ML 기술



Boost Your Lab's Performance in Quantum, Photonics, Materials



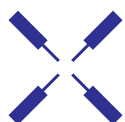
Software to fast-track
your research goals



Hardware for
highest fidelity



Support by our
application experts



**Zurich
Instruments**

한국담당자 ilnam.yeom@zhinst.com
82-10-6456-3463



ICEIC 2025

International Conference on Electronics, Information,
and Communication 2025



JAN. 19 ^(SUN) - 22 ^(WED) 2025 | Osaka International House, Japan

ICEIC 2025 Organizing Committee

General Chair

- Kwang-Hyun Baek Chung-Ang University

General Co-Chair

- Yoshitumi Nishio Tokushima University

Organizing Committee Chair

- Hyuck-In Kwon Chung-Ang University

Organizing Committee Co-Chair

- Jong-Ok Kim Korea University
- Mun-Sik Kang Gangneung-Wonju National University
- Myounggon Kang University of Seoul
- Yongfu Li Shanghai Jiao Tong University
- Z. Wesley Zun Cornell University

TPC Chair

- Minsuk Koo University of Seoul

TPC Co-Chair

- Yoko Uwate Tokushima University
- Ickhyun Song Hanyang University
- Tony Kim Nanyang Technological University
- Yong Shim Chung-Ang University
- Richard J. Cha George Washington University
- Lua Ngo VNU-HCM International University

Best Paper Track Chair

- Sangwan Kim Sogang University

Special Session Chair

- Jae Ho Han Korea University
- Suk-Ju Kang Sogang University

Special Session Co-Chair

- Byeong Ho Choi Korea Electronics Technology Institute
- Sung-Joon Jang Korea Electronics Technology Institute
- Manbok Park Korea National University of Transportation
- Tae-Heon Yang Konkuk University
- Sung Hyun Cho Hanyang University
- Hyun Kim Seoul National University of Science and Technology

Tutorial Chair

- Yoon Kim University of Seoul
- Sung-In Cho Dongguk University

Financial Chair

- Cheol-Ho Hong Chung-Ang University

Registration Chair

- Hyungjin Kim Hanyang University

Publicity Chair

- Ick-Joon Park Joongbu University
- Tae In Kim Inha University

Publication Chair

- Garam Kim Myongji University

Local Arrangement Chair

- Incheon Paik The University of Aizu
- Yousun Kang Tokyo Polytechnic University
- Hyung Do Kim Kyoto University

Conference Activity Chair

- Han Lim Lee Chung-Ang University

Conference Secretary Chair

- Min-Hwi Kim Chung-Ang University

Industry Chair

- Min-Seong Choo Hanyang University

CALL FOR PAPERS

The 24th International Conference on Electronics, Information, and Communication (ICEIC 2025) is a forum open to all the participants who are willing to broaden professional contacts and to discuss the state-of-the-art technical topics. Regular sessions of ICEIC 2025 will include more than 400 oral and poster presentations. In addition, the conference will offer special sessions, invited talks, keynote speeches, and tutorials to cover a broad spectrum of topics on electronics, information, and communication technologies.

Plenary Speakers

- Jong-Ho Lee (Professor / Former Minister of Science and ICT) Seoul National University
- Joo-Ho Lee (Professor) Ritsumeikan University

Tutorial Speakers

- TOMIOKA Yoichi (Professor) The University of Aizu
- Minkyu Je (Professor) KAIST
- Jian Zhao (Professor) School of Electronic Science and Engineering at Nanjing University

- Jinkyu Kim (Professor) Korea University
- Min-Seong Choo (Professor) Hanyang University
- Sung In Cho (Professor) Dongguk University

TOPIC

Communications

Communication & Information Theories, Communication Networks & Systems, Microwave & Optics, Switching and Routing, Microwave, Antennas and Propagation, Intelligent Transportation System (ITS), Wireless PAN/BAN, Future Networks

Semiconductor and Devices

Analog/Digital Circuits & Systems, RF Integrated Circuits, ComputerAided Design & Modeling, SoC Design & Applications, Semiconductors, Materials and Components, Lightwave and Quantum Electronics, PCB & Packaging, Solar Cell & Semiconductor Devices

Computer and Information

Computer Systems & Applications, Software for Smart Systems, Human Computer Interaction (HCI), Convergence Computing, Multimedia, Graphics, Ubiquitous System, Information Security, Artificial Intelligence, Neural Networks, Machine Learning

Signal Processing

Computer Vision, Digital Signal Processing, Digital Image/Video Processing, Audio, Speech & Acoustic Signal Processing

System and Control

Vehicular Electronics, Instrumentation and Control, Power Electronics & Circuits

Emerging Technologies

Biomedical Electronics and Bioengineering, Bio-electronics, ITConvergence, Renewable Energy, Car & Aviation IT

IMPORTANT DATES

- Submission of Paper : **October 13, 2024**
- Notification of Acceptance : **November 11, 2024**
- Submission of Camera-Ready Paper : **November 25, 2024**

SUBMISSION OF PAPERS

Prospective authors are invited to submit original papers (1~6 pages) of either MS Word or PDF format written in English.

Paper submission procedures are available at <https://iceic.org/>

※ You may submit your paper in either Full Paper format or Abstract format.

CONTACT POINT

- E-mail : inter@theieie.org
- Tel : +82-2-553-0255(Ext. 4)
- <https://iceic.org/>



대한전자공학회

2024

학회상 수상자

우리 학회는 1968년부터 매년 전자·정보·통신 분야에 탁월한 업적을 이루어 전자공학 학문과 기술의 발전에 공헌한 회원들에게 학회상을 시상하고 있으며, 금년에는 11월 22일(금) 하이원리조트 컨벤션타워 5층 컨벤션홀에서 시상식이 개최되었다.

[대한전자공학대상]



[기술혁신상]



[2024 IEIE Research Pioneer Award]



[논문상 단체]



[공로상 단체 1]



[공로상 단체 2]





대한전자공학대상



한 종 희 대표이사 부회장 (삼성전자주)

| 주요 학력/이력 |

- 1981년~1988년 인하대학교 전자공학 학사
- 2007년~2011년 삼성전자 영상디스플레이사업부 개발그룹 담당(상무)
- 2011년~2013년 삼성전자 영상디스플레이사업부 상품개발팀장(전무)
- 2013년~2017년 삼성전자 영상디스플레이사업부 개발팀장(부사장)
- 2017년~2021년 삼성전자 영상디스플레이사업부 사업부장(사장)
- 2021년~현재 삼성전자 대표이사, DX부문장(대표부회장)

| 대외이력 |

- 2018년~2019년 한국스마트홈산업협회 회장
- 2022년~현재 한국공학한림원 정회원
- 2022년~현재 한국전자정보통신산업진흥회 회장

| 주요업적 |

한종희 대표이사는 1988년 삼성전자 입사 후 36여년간 TV, 오디오, 생활가전, 스마트폰 등 다양한 제품 분야에서 지속적인 기술혁신을 통해 QLED TV, Micro LED TV, Galaxy S/Fold 시리즈 등 최고 사양/신기술이 적용된 제품을 지속 런칭하여 18년 연속 전 세계 TV시장 1위, 13년 연속 글로벌 스마트폰 시장 1위를 달성하는 등 국내외 전자산업 발전을 주도하였다.

또한, 로봇, 빅데이터, AI 등 미래 산업 핵심 신기술에 대한 과감하고 지속적인 투자와 함께 Micro LED, 광색역 디스플레이, 자발광 QLED 소재 관련 기술 등 국내 대학들과의 지속적인 '산학협력 과제'들을 발굴, 수행하여 한국 전자기술 발전과 IT강국으로서의 대한민국 위상제고에도 크게 공헌하였다. 그리고 한국전자정보통신산업진흥회장, 한국스마트홈산업협회장, 한국공학한림원 회원 등 여러 대외 활동을 통해 국가 산업 혁신과 미래전략 정책을 제안하고 유관기관과의 협력 확대를 추진하였다.

더불어, 국내 5대 그룹 중 유일한 공개채용 제도 운영 및 지속적인 고용확대(삼성전자 인력규모, 2023년말 기준 약 12.5천 명), 청년 S/W 아카데미 교육 등을 통해 국가 청년 실업 해소에 앞장서고 있으며 사내외 벤처지원 제도인 C-Lab Inside / Outside와 스마트공장 구축지원 등을 통한 스타트업 생태계 활성화 및 상생 협력에도 이바지하였다. 또한, 2019년 새로운 사회공헌 비전인 '함께 가요 미래로! Enabling People'을 선포하고 기업의 사회적 책임을 다하기 위해 청소년 교육을 중점 테마로 드림클래스, 희망디딤돌 등의 다양한 사회공헌 활동을 주도하고 있다.



기술혁신상



이 장 규 대표이사 (주)텔레칩스

| 주요 학력/이력 |

- 1982년~1986년 서강대학교 전자공학과 공학사
- 1986년~1988년 서강대학교 전자공학과 공학석사
- 1988년~1993년 삼성전자 반도체 부문 대리연구원
- 1993년~1999년 (주)씨앤에스테크놀로지 연구소장
- 1999년~2014년 (주)텔레칩스 공동설립 부사장
- 2014년~현재 (주)텔레칩스 대표이사
- 2018년~현재 (주)오토실리콘 대표이사

| 주요업적 |

이장규 대표는 시스템반도체 분야의 전문가로서 Caller ID, MP3 인코더 등 시장 선도적인 혁신적 기술 개발을 통해 국내 반도체 업계의 성장에 기여했다. 이러한 기술력을 바탕으로 2007년 국내 최초로 현대기아차에 카오디오 프로세서를 공급했으며, 2015년에는 차량용 AVN 프로세서, 2018년에는 HUD 및 디지털 클러스터 솔루션까지 그동안 전량 수입에 의존하던 차량용 인포테인먼트 반도체를 국산화하는 데 큰 역할을 했다. 또한 차세대 전기/스마트카 아키텍처가 Zonal 아키텍처로 진화하고, SDV 솔루션에 적합하게 변화해 감에 따라 고성능 인포테인먼트뿐만 아니라 ADAS/자율주행용 AI 가속기, 네트워크 게이트웨이, 차량용 고성능 MCU 등 미래지향적 차량용 핵심 반도체를 지속해서 개발하여 차량용 반도체의 국산화와 신성장 기술을 선도 하였으며 국내 시스템 반도체 생태계 조성과 글로벌 경쟁력 강화에 크게 기여하였다.

IEIE Research Pioneer Award



한 재 호 교수 (고려대학교)

| 주요 학력/이력 |

- 1994년~1998년 고려대학교 전자공학과 공학사
- 1998년~2000년 고려대학교 전파공학과 공학석사
- 2000년~2005년 LG전선(주) 연구소 주임연구원
- 2005년~2010년 美 존스홉킨스대 전기·컴퓨터공학과 공학박사
- 2010년~2011년 美 국립표준기술연구소(NIST) 박사후연구원
- 2011년~현재 고려대학교 교수

| 주요업적 |

한재호 교수는 광신호처리 분야에서 혁신적인 연구를 수행하여 광학 기반 바이오 이미징 및 의료 영상처리 분야에서 신경 생체 조직의 신호를 효과적으로 처리할 수 있는 새로운 기술의 개발과 문제 해결에 기여하였다. 기존 광학 이미징 기술의 한계를 뛰어넘는 새로운 방식을 제시하여 서브 나이퀴스트 샘플링을 통한 이미지 획득, 공간주파수 다중화를 통해 단일 카메라로 물리적 공간을 초과하는 넓은 영역의 촬영과 라벨링 없이도 신경 세포의 특성을 측정하는 다양한 측면에서 기술 혁신으로 기존 샘플링 방법의 한계와 영상처리 시스템 등 기존 이미징 과정의 근본적인 문제를 해결하는 데 공헌하였다. 이러한 새로운 융합적인 연구 분야의 개척을 통해 이전에는 감지하기 어려웠던 미세한 뇌신경 섬유와 광학적 특성이 측정되었으며, 단일 카메라를 통해 실시간으로 여러 물체를 동시에 관찰하는 것이 가능해지고, 생물 세포의 구조부터 반도체 리소그래피 공정에 이르기까지 다양한 현상을 더 깊이 이해하는 데 기여하였다. 이와 더불어 인공지능 알고리즘을 개발하여 안과학 분야에서 망막 영상 분석 자동화를 통해 망막 질환의 정확하고 신속한 진단과 수술 후의 예측에 활용될 것으로 보인다. 궁극적으로는 광계측 제어시스템, 생체신호 정보처리, 의료 인공지능을 포함하여 다양한 융합 분야에 걸쳐서 폭넓게 영향력을 미치고 상용화에 공헌하였다. 한국연구재단, 정보통신기획평가원, 식품의약품안전평가원, 한국환경산업기술원, 한국보건산업진흥원, 한국과학기술연구원, 삼성미래기술육성센터, LG연암문화재단 등의 기관으로부터 연구 지원을 받아, 연구 결과들은 IEEE TIP, SPM, TIE, TMECH, TMI, JSTQE 등의 저명한 최상위권 국제학술지에 게재되어 그 우수성을 인정받고 있다.



대한전자공학회 논문상

| 반도체 분야 |



차 혁 규 교수 (서울과학기술대학교)

| 주요 학력/이력 |

- 1998년~2003년 KAIST 전기 및 전자공학과 공학사
- 2003년~2009년 KAIST 전기 및 전자공학과 공학박사
- 2009년~2012년 Institute of Microelectronics, A*STAR, Singapore, Scientist
- 2012년~현재년 서울과학기술대학교 전기정보공학과 교수

| 주요업적 |

차혁규 교수는 센서 인터페이스 저잡음 신호 처리 회로, 고효율 전력 변환 및 관리 회로, 저전력 무선 통신용 송수신기 회로를 포함한 아날로그/RF 집적회로 설계 연구를 수행하고 있다. 특히, 다양한 목적을 갖는 고전압 동작 회로의 집적도 개선과 시스템 내에서 저전압 회로와의 호환성을 위해 표준 CMOS 반도체 소자를 안정적으로 적용한 고전압 회로 구조 연구에 기여하였다. 그동안 수행한 아날로그/RF 집적회로 설계 연구에 대한 결과물로 다수의 국내외 저널 게재 및 학술대회 논문을 발표하는 등 관련 분야 발전에 공헌하였다.

| 컴퓨터 분야 |



오 희 석 교수 (한성대학교)

| 주요 학력/이력 |

- 2003년~2010년 연세대학교 전기전자공학부 공학사
- 2010년~2012년 연세대학교 전기전자공학부 공학석사
- 2012년~2017년 연세대학교 전기전자공학부 공학박사
- 2017년~2017년 삼성전자 DMC연구소 책임연구원
- 2017년~2020년 한국전자통신연구원 선임연구원
- 2020년~현재 한성대학교 조교수

| 주요업적 |

오희석 교수는 2차원 영상 복원 외, 컴퓨터비전, 3차원 영상 재구성 등 다양한 지능적 영상처리 분야에서 신기술 개발을 통해 학계에 공헌했다. 특히, 가상현실 콘텐츠의 품질을 정량화하고 사이버 멀미를 예측하는 연구를 통해 몰입형 콘텐츠의 품질 향상에 기여하였으며, 이와 같은 인간 시각 체계 기반의 연구 결과는 다수의 국내외 학회 및 학술지에 게재되었다. 특히 최근의 연구는 차세대 멀티모달 콘텐츠 처리 기술과 미래의 몰입형 경험을 위한 중요한 통찰을 제공하며, 관련 논문들이 다수 인용되는 등 해당 분야 연구자들에게 연구 방향을 제시하였다.



대한전자공학회 논문상

| 산업전자 분야 |



김 필 중 교수 (조선이공대학교)

| 주요 학력/이력 |

- 1988년~1994년 조선대학교 전자공학과 공학사
- 1994년~1996년 조선대학교 전자공학과 공학석사
- 1996년~2000년 조선대학교 전자공학과 공학박사
- 1996년~1999년 SK하이닉스(구, 현대전자) 메모리연구소 연구원
- 2001년~2012년 성화대학 IT계열 조교수
- 2015년~2024년 조선이공대학교 전자과 조교수
- 2024년~현재 조선이공대학교 전자과 부교수

| 주요업적 |

김필중 교수는 산업전자 분야에 상용화되고 있는 전자뇌관 회로에 대한 근본적인 분석을 통해 기존 제품보다 더 소형화, 안정화, 저전력화가 가능한 방안을 제안하여 관련 산업 종사자에게 연구 개발 방향과 통찰력을 주었으며, 또한 파워 온 리셋 회로와 저전압용 기준전압 발생기 등 회로 설계에 관한 연구를 통해 메모리 반도체 등 전자시스템의 성능 개선에 공헌하였다. 그리고 자동차 브레이크 등 상태 감지 시스템의 회로 설계와 구현에 관한 연구로 자동차 전장시스템의 성능 개선과 UV 젤 네일을 위한 LED 램프 회로 설계에 관한 연구로 미용 분야 전자시스템의 기능향상 등 산업전자 다양한 분야의 발전에 공헌하였다.



2024년도 학회 공로상 수상자 명단

성명	직위	소속	공적
강운병	마스터	삼성전자(주)	반도체(소) 활성화
강석주	교수	서강대학교	2024 AISP학술대회 조직위원장
고형호	교수	충남대학교	반도체(소) 활성화
공성호	교수	경북대학교	전임지부장(대구 · 경북)
김선용	교수	건국대학교	통신(소) 활성화
김세종	이사	(주)SJ정보통신	산업전자(소) 활성화
김용석	석좌교수	가천대학교	반도체(소) 활성화
김원준	교수	건국대학교	인공지능신호처리(소) 활성화
김 혁	교수	서울시립대학교	2024 ICCE-Asia 조직위원
나태희	교수	인천대학교	2024 ICCE-Asia 조직위원
박병섭	교수	인하공업전문대학	산업전자(소) 활성화
박영훈	교수	숙명여자대학교	컴퓨터(소) 활성화
서병석	교수	상지대학교	산업전자(소) 활성화
안성수	교수	명지전문대학	산업전자(소) 활성화
정성엽	교수	고려대학교	2024 하계 조직위원
정승원	교수	고려대학교	2024 하계 TPC위원장
조성인	교수	동국대학교	2024 AISP학술대회 TPC위원장
주영복	교수	한국기술교육대학교	2024 시스템및제어(소) 회장
추민성	교수	한양대학교	반도체(소) 활성화

대한전자공학회

2024 해동상

대한전자공학회가 주관하고 해동과학문화재단(김영재 이사장)이 후원하는 제34회 해동기술상 수상자로 김병훈 CTO 부사장을 그리고 해동젊은공학인상(학계)에는 김형진 한양대학교 교수가 각각 선정되었다.

김병훈 부사장은 ICT 산업계에서 지난 20여 년 동안 다양한 기술 분야를 연구 개발한 글로벌 기술리더로 GCT Semiconductor, Qualcomm, LG전자에서 반도체 기술 개발 및 통신 분야 선행기술 연구와 활발한 표준 활동을 통해 다수의 표준 핵심 특허를 확보하고 1만 5천 건 이상의 연구 피인용 횟수를 기록하는 등 국내외 산학연 협력을 통해 대한민국 기술의 국제적 위상과 영향력을 높였다.

특히 2021년 12월 LG전자 CTO로 부임한 이후 5G/6G 통신, 인공지능, 로봇, SoC, SW플랫폼, 모빌리티, 미디어, IoT, 소재, 광학, 차세대 컴퓨팅 관련 기술들을 다수 개발하고 IP 확보 및 사업화에 기여하는 등 모든 R&D 혁신을 이끌고 있으며, 그 이외에도 차량용 반도체 과제 등 국책사업 참여, 유망 스타트업 발굴과 육성 등을 통해 국가 경쟁력 확보에도 기여했다.

김형진 교수는 반도체 소자 공정 및 특성평가와 메모리 소자를 활용한 컴퓨팅 시스템 분야에 창의적인 연구개발을 통해 103편(주저자 62편)의 논문을 SCI급 학술지에 게재하고 반도체소자 최고 학술대회인 IEEE IEDM에서 발표하였으며, 이러한 연구결과는 기존 메모리 기술들이 해결해야 했던 에너지 효율, 속도 등의 개선을 통해 인공지능 반도체 기술에 활용될 것으로 기대된다.

해동상은 1965년 대덕전자(주)를 설립하여 50여년간 PCB사업에 전념하여 오신 창업주 (故)김정식 회장님께서 우리나라 전자공학 분야의 학문 및 기술발전을 위하여 크게 업적을 쌓은 분들의 노고를 치하하기 위해 제정한 상이다.

해동상은 1991년 설립된 해동과학문화재단에서 <대한전자공학회 해동상>을 제정한 이래 현재까지 총 34회에 걸쳐 시행하여 왔다.

<대한전자공학회 해동상>은 해동상 설립 초기부터 시상하여 현재 학술상, 기술상, 젊은공학인상(학술상), 젊은공학인상(기술상) 4개 부문으로 구분하여 업적이 탁월한 분들에게 시상하고 있다.

한편, 이번 제34회 해동상 수상자 선정을 위해 각 대학, 연구소, 기업체에 후보자 추천을 의뢰하고 전 회원들에게 이메일 발송, 홈페이지 게시 및 각종 행사에서 홍보 등 다방면으로 홍보하여 8월 18일까지 후보자에 대한 추천서를 접수하였다.

해동학술상은 1인, 해동기술상은 2인, 젊은공학인상(학술상)은 2인, 젊은공학인상(기술상)은 1인이 각각 응모하였으며, 포상위원회에서(2024.10.08.)에서는 해동상 평가 세부 지침에 의거하여 심도 있는 평가를 거쳐 최종 수상후보자를 선정하였으며, 학회 6차 상임이사회(2024.10.11.) 의결을 거쳐 후보자를 선정하고 해동과학문화재단의 승인 후 확정하였다.

[해동상 수상자]





해동기술상



김 병 훈 CTO 부사장 (LG전자㈜)

| 주요 양력 및 이력 |

- 1990년~1994년 서울대학교 전자공학 학사
- 1994년~1996년 서울대학교 전자공학 석사
- 1996년~2000년 서울대학교 전기컴퓨터공학 박사
- 2011년~2012년 Wi-Fi Alliance Board of Directors (WFA)
- 2015년~현재 5G포럼 집행위원/운영위원, 공동의장(2024년~)
- 2018년~현재 Open Connectivity Foundation Korea Forum 부의장 (OCF Korea)
- 2019년~2020년 LG전자 C&M표준연구소 연구소장
- 2020년 국제전기전자공학회 (IEEE) Fellow
- 2021년 LG전자 미래기술센터 센터장
- 2021년 과기정통부 장/차관 전략 회의 (6G R&D 전략위원회) 전문가위원
- 2021년~현재 LG전자 CTO (부사장 겸)인공지능연구소장
- 2024년 제59회 발명의날 금탑산업훈장 수상

| 주요업적 |

김병훈 부사장은 지난 20여년 동안 다양한 기술 분야를 연구 개발한 글로벌 기업인이다. GCT Semiconductor에서 W-CDMA 및 WLAN 칩셋을 개발했고, Qualcomm에서는 MIMO기술 개발과 3GPP LTE 표준 및 설계를 담당하는 등 반도체 기술 개발 총괄을 역임했다. 이후 LG전자 C&M 표준연구소장 시절에는 Connected Car, IoT Connectivity 및 차세대 미디어/방송 분야에서도 글로벌 리더십을 발휘해 표준화와 원천기술 개발을 주도하였다.

2020년에는 통신 및 Connected Car 기술(V2X) 표준화와 혁신적인 연구 활동을 높이 평가받아 국제전기공학회(IEEE) Fellow로 선정되었다. 2000년부터 현재까지 통신분야 선행 기술 연구와 표준 활동을 통해 다수의 표준핵심 특허를 확보하고, 14,000여건 이상의 연구 피인용 회수를 기록하는 등 차세대 이동통신 분야의 기술 발전에 기여해 왔다. 또한 3GPP, IEEE802.11 및 5GAA 표준에 광범위하게 기여했으며, Wi-Fi Alliance의 이사회 구성원 역할을 맡는 등 국제 표준을 주도하였다.

2021년 12월에는 LG전자 최고기술책임자(CTO)로 부임하여 6G 핵심 통신, 인공지능, 스마트홈, 모빌리티, 물류로봇, Private 5G, 디지털 헬스케어, 메타버스, IoT, Block Chain 관련 다수의 기술들을 개발하고, IP 확보 및 사업화에 기여하는 등 LG전자의 모든 R&D 혁신을 이끌고 있으며, 특히 2022년에 자동차 SW기능안전 인증과 2023년에 가전 SW기능안전 분야의 지정시험기관 인증을 세계 최초로 획득하여 품질분야 국제적 전문성과 공신력을 가지게 되었다.

또한 산학연 협력을 통한 기술 생태계 구축 및 인재 양성에도 적극 참여하여 국가 과학기술 전략에 맞춰 차량용 반도체, 첨단로봇 등의 국가 경쟁력 확보를 위해 정부가 시행하는 주요 국책사업에도 참여하여 다양한 기업, 연구소 및 대학과 협력 네트워크를 강화하고 있으며, 미래 기술 특강을 통해 세상의 변화와 연구개발의 발명가 DNA를 이공계 대학생들에게 전파하여 미래 국가 기술발전에 기여할 인재 양성에도 힘쓰고 있다.

한편 2024년 5월에는 차세대 기술표준을 개발하여 관련 특허권 확보를 주도하고 R&D 혁신을 통해 글로벌 기술경쟁력을 크게 높인 공로로 제59회 발명의날 금탑산업훈장을 수상했으며, 2017년에 LG전자 R&D Hall of Fame(명예의 전당) 헌정대상자로, 2016년에 서울대학교 70주년 “한국을 빛내는 70인의 서울공대 박사”로 선정된 바 있다.



해동젊은공학인상(학술상)



김형진 교수 (한양대학교)

| 주요 양력 및 이력 |

- 2006년~2010년 서울대학교 전기공학 학사
- 2010년~2012년 서울대학교 전기컴퓨터공학 석사
- 2012년~2017년 서울대학교 전기컴퓨터공학 박사
- 2018년~2019년 University of California, Santa Barbara, ECE 박사후연구원
- 2020년~2024년 인하대학교 전자공학과 조교수, 부교수
- 2024년~현재 한양대학교 신소재공학부/반도체공학과 부교수

| 주요업적 |

김형진 교수는 차세대 메모리 소자를 개발하고, 이를 효율적인 정보 처리에 활용하는 인메모리(in-memory) 컴퓨팅 분야에서 활발한 연구를 수행해 오고 있다. 특히 실리콘 플래시 소자와 멤리스터(memristor) 소자와 같은 첨단 메모리 소자로 구성된 대규모 메모리 어레이를 직접 설계 및 제작하여, 이를 기반으로 뉴로모픽(neuromorphic) 컴퓨팅, 하드웨어 보안, 논리 연산과 같은 다양한 컴퓨팅 응용 분야에 대한 가능성을 검증해왔다. 이로써 메모리 반도체를 기반으로 한 혁신적인 컴퓨팅 기술을 제안하고, 이 기술을 통해 다양한 정보 처리 및 데이터 분석 시스템의 성능을 높이기 위한 연구를 진행 중이다.

김형진 교수는 이러한 연구를 위해 재료, 소자, 회로, 시스템을 아우르는 융복합적 접근을 통해 인공지능과 반도체 소자의 경계를 넘나드는 연구를 지향하고 있으며, 반도체 소자의 독특한 물리적 특성을 최대한 활용하여 새로운 기능을 구현하는 데 집중하고 있다. 예를 들어, 터널링과 같은 반도체 소자의 고유한 전류 특성을 기반으로 물리적 복제 불가능 함수(physically unclonable function, PUF)의 신뢰성을 높이는 연구를 수행했다. 이를 통해 PUF 소자가 외부 온도 변화에 강건한 특성을 가지도록 하여 보안성을 향상시키는 성과를 이루었다. 또한, 기존의 NAND 플래시 구조를 메모리와 커패시터의 병렬 어레이 구조로 활용하여 행렬곱 연산을 구현하는 새로운 접근을 시도하여 차별화된 연구 성과를 도출하였다. 이와 같은 연구 결과는 세계 최고 권위의 학술대회 중 하나인 IEEE IEDM(International Electron Devices Meeting)에 발표되었으며, 학계와 업계에서 주목받고 있다.

김형진 교수는 지금까지 SCI(E) 등재 국제저명학술지에 104편의 논문을 게재하였으며, 이 중 62편의 논문에서 주저자로서 연구를 이끌었다. 또한, 대한전자공학회뿐만 아니라 여러 학술단체에서 임원으로 활동하며 학문적 네트워크를 확장하고 있으며, 다수의 연구 과제에도 참여하여 기술 개발과 인재 양성에 기여하고 있다. 이러한 노력은 지능형 반도체 분야의 기술 발전에 기여함과 동시에, 차세대 반도체 분야의 우수한 연구 인력을 양성하는 데에도 중요한 역할을 하고 있다. 김형진 교수의 연구 활동과 성과는 반도체 기반의 차세대 컴퓨팅 기술 혁신에 중요한 초석이 되고 있으며, 그가 이끄는 연구는 미래 인공지능 및 정보 처리 기술의 발전에 핵심적인 기여를 할 것으로 기대된다.



SAE MOBILUS™에서 세계 최고 Automotive Engineering 정보를 만나보세요.

SAE International은 국제 자동차 공학자 협회 (Society of Automotive Engineers)로 자동차, 비행기 및 기타 내연기관 관련 산업의 과학과 기술을 선도하는 학회입니다.

SAE MOBILUS™는 207,000건 이상의 특허 기록 문서, 기술 자료집, 그리고 eBooks과 간행물 자료를 제공합니다. SAE MOBILUS™은 기업의 산업 현장, 학교, 그리고 연구소의 사용자들이 필요한 자료를 이용하실 수 있도록 다양한 구독 옵션을 함께 제공합니다.

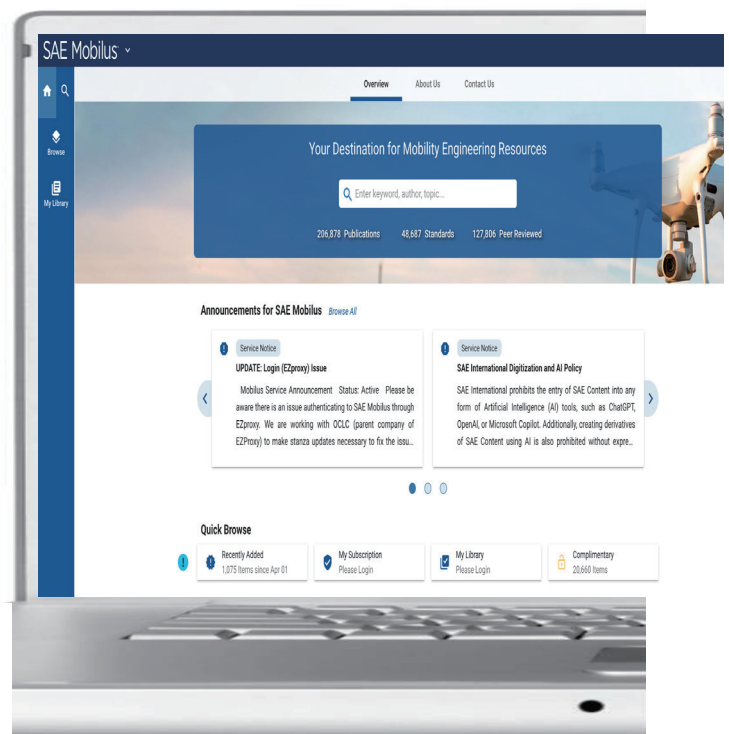
SAE MOBILUS™에서 다음과 같은 다양한 주제 분야에 대한 자료를 확인할 수 있습니다.

- Noise, Vibration, and Harshness (NVH)
- Parts and Components
- Vehicle and Performance
- Quality, Reliability, and Durability
- Maintenance and Aftermarket
- Design Engineering and Styling
- Chassis
- Bodies and Structures
- Safety
- Manufacturing
- Power and Propulsion
- Interiors, Cabins, and Cockpits
- Human Factors and Ergonomics
- Materials
- Fuels and Energy Sources
- Environment
- Electrical, Electronics, and Avionics
- Tests and Testing
- Management and Organization
- Transportation Systems

SAE MOBILUS™에서 항공 우주 분야 R&D에 필요할 표준 및 규격 그리고 다양하 자료를 검색 및 이용할 수 있습니다.

- +9,300 SAE Ground Vehicle Standards (J-reports)
- +15,500 SAE Aerospace Standards(AS,ARP, AIR, and 2D/3D Configurator Parts)
- +18,200 SAE Aerospace Material Specification (AMS)
- +3,600 SAE ITC Engine & Airframes Standard
- SAE Historical Standards
- Composite Materials Handbook(CMH-17)
- 복합 재료에서 최종 품목을 설계하고 제작하는 데 필요한 정보와 지침을 제공

<http://saemobilus.sae.org>



IEEE와 Rutgers 비즈니스 스쿨에서 엔지니어와 기술 전문가를 위한 IEEE | Rutgers Online Mini-MBA를 소개합니다. 엔지니어와 기술 전문가들의 전문성과 비즈니스 감각을 높일 수 있는 본 프로그램으로 여러분을 초대합니다.

엔지니어와 기술 전문가를 위한 IEEE | Rutgers 온라인 Mini-MBA 프로그램을 소개합니다. 이 프로그램은 전문적인 기술 지식과 비즈니스 역량을 결합하여, 커리어에 새로운 도약을 준비할 수 있는 기회를 제공합니다.

총 12주간 진행되는 본 프로그램은 1인당 등록비 US\$3,495이며, 한 기관에서 10명 이상 등록할 경우, 추가 할인 혜택도 드립니다.

특히, 캡스톤 프로젝트가 포함되어 있어 실제 업무에 프로그램 내용을 적용하고, 업무 효율과 생산성을 높일 수 있는 기회를 제공합니다. 또한, 여러 부서와 협력하여 현재 진행 중인 프로젝트를 프로그램과 연계해 진행할 수도 있습니다.

IEEE | Rutgers Mini-MBA 프로그램의 주요 장점

- 기술적 성과와 조직의 목표를 효과적으로 연결하는 방법을 배웁니다.
- 기술적 경험과 비즈니스 관리 능력을 균형 있게 성장시킬 수 있는 리더십을 강화합니다.
- 다양한 부서와의 긴밀한 협력을 통해 문제 해결 능력을 높이고, 종합적인 관점을 기릅니다.
- 비기술적인 이해관계자들과 원활하게 소통할 수 있도록, 기술과 엔지니어링의 가치를 비즈니스 언어로 표현하는 방법을 익힙니다.

프로그램 수료 시 제공되는 혜택

- IEEE | Rutgers Executive Education Certificate 발급
- 전체 MBA 과정에서 3학점 면제
- 35 PDH(Professional Development Hours) 제공
- 디지털 배지 제공

10명 이상의 그룹이 참여할 경우, 기관 맞춤형 프로그램으로 운영될 수 있으며, 일정 조정도 가능합니다. 학습 내용을 실제 업무에 적용할 수 있는 맞춤형 캡스톤 프로젝트를 통해 실력을 발휘하고, 기관 내에서 의미 있는 성과를 달성할 수 있습니다.

글로벌 비즈니스 시장에서 성공을 이루고 싶은 당신을 위한 기회
비즈니스와 기술의 경계를 넘나드는 실력을 갖추고, 더 큰 성공을 이루고자 한다면, 이 프로그램은 완벽한 선택입니다.

IEEE와 Rutgers Business School이 여러분의 성장을 지원합니다.

자세한 내용은 IEEE | Rutgers Mini-MBA 프로그램에서 확인하실 수 있습니다.

등록에 관심이 있거나 궁금한 점이 있으시면 언제든지 연락주시기 바랍니다:

Mini-MBA@ieee.org | innovate.ieee.org/rutgers-mini-mba-for-engineers

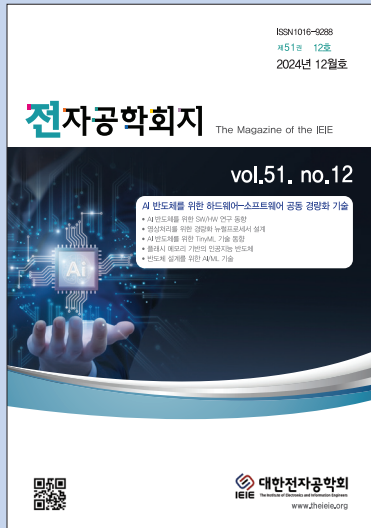


강의 주제 포함:

- 비즈니스 벤처 및 혁신
- 신제품 개발 관리
- 재무제표 분석
- 성공적인 협상의 열쇠
- 글로벌 경제의 영향
- 실무 마케팅
- 데이터 기반 문제 해결
- 리더십 전략
- 디지털 혁신의 이해
- 위기 관리 및 대응 전략
- 비즈니스 윤리와 사회적 책임
- 효율적인 팀 관리와 협업
- 창의적 문제 해결 능력 개발
- 자원 관리 및 최적화 전략
- 전략적 의사 결정 모델 구축
- 지속 가능 경영 전략

CONTENTS

제51권 12호 (2024년 12월)



※ 학회지 12월호 표지 (vol 51. No 12)

회지편집위원회

- 위 원 장 황 원 준 (아주대학교 교수)
- 부위원장 김 영 로 (명지전문대학 교수)
- 위 원 김 상 범 (한국폴리텍대학 교수)
박 승 창 (주유오씨 대표)
박 진 선 (부산대학교 교수)
심 규 성 (한경국립대학교 교수)
안 성 수 (명지전문대학 교수)
윤 중 현 (조선이공대학교 교수)
이 정 석 (인하공업전문대학 교수)
이 철 (동국대학교 교수)
정 승 원 (고려대학교 교수)
최 욱 (인천대학교 교수)
- 사무국 편집담당
이 안 순 부장
TEL : (02)553-0255(내선 6)
FAX : (02)562-4753
- 학회 홈페이지
<http://www.theieie.org>

학회소식

12 학회소식 / 편집부

특집 : AI 반도체를 위한 하드웨어-소프트웨어 공동 경량화 기술

16 특집편집기 / 류성주

17 AI 반도체를 위한 SW/HW 연구 동향 / 박재범, 안민영, 박종선

29 영상처리를 위한 경량화 뉴럴프로세서 설계 / 류성주

36 AI 반도체를 위한 TinyML 기술 동향 / 임형철, 이주석, 이성수

44 플래시 메모리 기반의 인공지능 반도체 / 이성태

51 반도체설계를 위한 AI/ML 기술 / 최규명

회원광장

63 논문지 논문목차

정보교차로

65 국내외 학술행사 안내 / 편집부

78 특별회원사 및 후원사 명단

2024년도 임원 및 각 위원회 위원

회 장	이 충 용 (연세대학교 교수)	김 기 남 (삼성전자 종합기술원 회장)
	수 석 부 회 장	백 광 현 (중앙대학교 교수)
	고 문	권 오 경 (한양대학교 석좌교수)
		김 영 재 (해동과학문화재단 이사장)
		방 승 찬 (한국전자통신연구원 원장)
감 사 부 회 장	안 승 권 (연암공과대학교 총장)	전 영 현 (삼성전자 부회장)
		최 창 식 (DB하이텍 부회장)
		전 병 우 (성균관대학교 교수) - 감사
		구 본 태 (한국전자통신연구원 본부장) - 산학연
		김 종 옥 (고려대학교 교수) - 사업총괄/대외협력
소사이어티 회장	노 원 우 (연세대학교 교수) - 국제총괄/ICCE Asia/영문논문/신기술	류 수 정 (서울대학교 교수) - 회원 및 여성
		이 승 호 (한밭대학교 교수) - 대외협력/지부
		고 정 환 (인하공업전문대학 교수) - 산업전자
		장 성 진 (삼성전자 상근교문) - 반도체
		최 용 수 (신한대학교 교수) - 컴퓨터
협동부회장		김 강 태 (삼성전자 부사장)
		김 영 한 (UC San Diego / 가우스랩스 대표이사)
		김 형 준 (한국과학기술연구원 소장)
		송 문 섭 (삼텍 회장)
		엄 낙 웅 (한국전자통신연구원 책임연구원)
상 임 이 사		원 제 형 (도쿄엘렉트론코리아 대표이사)
		이 광 업 (서경대학교 교수)
		이 병 선 (김포대학교 교수)
		이 서 규 (한국팹리스산업협회 회장)
		이 재 관 (한국자동차연구원 소장)
산 업 체 이 사		전 선 익 (파이낸셜뉴스 부회장)
		정 준 (솔리드 대표이사)
		강 석 주 (서강대학교 교수) - 총무
		권 혁 인 (중앙대학교 교수) - 사업총괄2
		김 영 로 (명지전문대학 교수) - 회지편집
이 사		김 용 신 (고려대학교 교수) - 대외협력/산학연
		김 윤 (서울시립대학교 교수) - 사업
		김 재 준 (서울대학교 교수) - JSTS
		김 현 (서울과학기술대학교 교수) - ICBC/기획
		동 성 수 (온인예술과학대학교 교수) - 교육
이 사		변 대 석 (삼성전자 마스터) - 교육
		선 우 경 (서울대학교 교수) - 회원 및 여성
		송 진 호 (연세대학교 교수) - ICCE Asia
		유 찬 세 (한국전자기술연구원 수석연구원) - 사업총괄1
		이 덕 우 (계명대학교 교수) - 논문 편집
이 사		이 후 진 (한성대학교 교수) - ICCE Asia
		정 일 권 (한국전자통신연구원 본부장) - 사업
		조 성 현 (한양대학교 교수) - 논문편집 총괄
		황 원 준 (아주대학교 교수) - 회지편집
		고 용 남 (하나미이크론 전무)
이 사		김 현 수 (삼성전자 상무)
		서 경 열 (뷰웍스 부사장)
		우 정 호 (비전넥스트 대표이사)
		이 도 훈 (국가보안기술연구소 수석연구원)
		이 수 인 (텔레칩스 상무)
이 사		천 이 우 (넥스트칩 연구소장)
		한 은 혜 (에스에스앤씨 대표이사)
		홍 국 태 (LX세미콘 연구위원)
		공 재 하 (고려대학교 교수) - 대외협력
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		
이 사		

협동이사

권 구 덕 (강원대학교 교수) - 기획
 권 태 수 (서울과학기술대학교 교수) - 정보화
 김 선 옥 (고려대학교 교수) - 교육
 김 소 영 (성균관대학교 교수) - JSTS
 김 재 욱 (한국과학기술연구원 선임연구원) - 학술(하계)
 김 혁 (서울시립대학교 교수) - 교육
 문 상 미 (나사렛대학교 교수) - 여성/회원
 배 현 철 (한국전자통신연구원 책임연구원) - 학술(추계)
 변 영 재 (UNIST 교수) - ICCE Asia
 안 상 철 (한국과학기술연구원 책임연구원) - 신기술
 안 호 균 (한국전자통신연구원 실장) - 사업
 윤 상 민 (국민대학교 교수) - 신기술
 이 구 순 (파이낸셜뉴스 이사) - 홍보
 이 영 택 (ASML 전문) - 교육
 이 철 (동국대학교 교수) - 학회지
 임 동 구 (전북대학교 교수) - 정보화
 장 지 원 (연세대학교 교수) - 학회지
 정 해 준 (경희대학교 교수) - 국문논문
 진 성 훈 (인천대학교 교수) - 학술(하계)
 최 광 성 (한국전자통신연구원 실장) - 대외협력
 최 영 규 (인하대학교 교수) - 학술(하계)
 최 재 혁 (서울대학교 교수) - ICCE Asia
 하 정 우 (네이버 SI연구소장) - 신기술
 한 동 국 (국민대학교 교수) - 사업
 한 태 희 (성균관대학교 교수) - 국문논문
 허 재 두 (한국전자통신연구원 연구전문위원) - 사업
 황 태 호 (한국전자기술연구원 본부장) - 학술(하계)
 구 민 석 (서울시립대학교 교수) - 기획
 권 경 하 (KAIST 교수) - 사업
 김 건 우 (목포대학교 교수) - 회원
 김 기 현 (전북대학교 교수) - 회원
 김 대 영 (순천향대학교 교수) - 회원
 김 사 혁 (KISDI 책임연구원) - 대외협력
 김 정 석 (가천대학교/셀리코 교수/대표이사) - 대외협력
 김 형 진 (한양대학교 교수) - 회원/여성
 민 경 식 (국민대학교 교수) - JSTS
 배 준 성 (강원대학교 교수) - 정보화
 서 민 재 (서울시립대학교 교수) - 홍보/정보화
 서 중 열 (LG전자 그룹장) - 산학연
 송 준 영 (인천대학교 교수) - 홍보
 신 세 운 (POSTECH 교수) - 사업
 오 윤 호 (고려대학교 교수) - 학술(추계)
 유 동 훈 (디시일로 연구소장) - 산학연
 윤 명 국 (이화여자대학교 교수) - 신기술
 윤 희 인 (UNIST 교수) - 홍보
 이 재 규 (삼성전자 마스터) - 산학연
 이 정 원 (서울대학교 교수) - 회원/여성
 장 지 수 (삼성전자 PE) - 사업
 정 방 철 (충남대학교 교수) - 정보화
 정 완 영 (KAIST 교수) - 홍보
 차 철 웅 (한국전자기술연구원 센터장) - 표준화
 최 강 선 (한국기술교육대학교 교수) - SPC
 추 민 성 (한양대학교 교수) - 기획
 홍 상 완 (서강대학교 교수) - 대외협력

권 기 록 (부경대학교 교수) - 학술(하계)
 김 민 규 (LG이노텍 상무) - 산학연
 김 성 우 (서울대학교 교수) - 대외협력
 김 유 철 (LG SI연구원 부부장) - 신기술
 김 중 현 (고려대학교 교수) - 사업
 류 은 석 (성균관대학교 교수) - 사업
 박 성 정 (건국대학교 교수) - 국제협력
 백 종 덕 (연세대학교 교수) - 신기술
 안 광 호 (한국전자기술연구원 본부장) - 대외협력
 안 상 수 (명지전문대학 교수) - 학회지
 원 용 욱 (명지대학교 교수) - 신기술
 윤 종 윤 (파두 사장) - 교육
 이 상 만 (고려대학교 교수) - 산학연
 이 창 우 (가톨릭대학교 교수) - 국제협력
 이 태 동 (국제대학교 교수) - 홍보/정보화
 장 성 욱 (카카오모빌리티 부사장) - 신기술
 전 동 석 (서울대학교 교수) - 학술(하계)
 제 민 규 (KAIST 교수) - ICCE Asia
 채 관 엽 (삼성전자 마스터) - 대외협력
 최 병 호 (한국전자기술연구원 연구소장) - 산학연
 최 재 욱 (가천대학교 교수) - 국문논문
 최 정 욱 (한양대학교 교수) - 학회지
 하 태 준 (광운대학교 교수) - 학술(추계)
 한 진 호 (한국전자통신연구원 책임연구원) - 국제협력
 함 범 섭 (연세대학교 교수) - 학술(하계)
 황 진 영 (한국항공대학교 교수) - 홍보

권 건 우 (홍익대학교 교수) - 학술(하계)
 권 중 원 (한국산업기술시험원 책임연구원) - 학술(추계)
 김 근 영 (한국전자통신연구원 연구원) - 학술(하계)
 김 도 훈 (한국전자통신연구원 박사) - 표준화
 김 범 현 (한양대학교 교수) - 정보화
 김 승 환 (한국전자통신연구원 책임연구원) - 학술(추계)
 김 현 우 (고려대학교 교수) - 신기술
 류 성 주 (서강대학교 교수) - 학술(하계)
 박 관 서 (연세대학교 교수) - 학술(추계)
 배 준 호 (가천대학교 교수) - 표준화
 서 봉 상 (올포랜드 상무) - 홍보/정보화
 송 익 현 (한양대학교 교수) - 회원
 송 철 (DGIST 교수) - 신기술
 심 민 섭 (경상국립대학교 교수) - 회원
 유 경 창 (삼성전자 수석연구원) - 회원
 유 성 철 (쌍용정보통신 상무) - 홍보/정보화
 윤 상 훈 (한국전자기술연구원 책임연구원) - 사업
 이 권 형 (LG전자 팀장) - 학술(하계)
 이 정 석 (인하공업전문대학 교수) - 학회지
 임 매 순 (한국과학기술연구원 책임연구원) - 사업
 정 민 채 (세종대학교 교수) - 국문논문
 정 성 엽 (고려대학교 교수) - 학술(하계)
 정 재 욱 (인천대학교 교수) - 학회지
 채 주 형 (광운대학교 교수) - 총무
 최 욱 (인천대학교 교수) - 학회지
 한 상 욱 (삼성전자 연구원) - 학술(하계)

지부장 명단

강 원 지 부
 대구·경북지부
 부산·경남·울산지부
 제 주 지 부
 호 서 지 부
 미 국

강 문 식 (강릉원주대학교 교수)
 이 찬 수 (영남대학교 교수)
 고 진 환 (경상국립대학교 교수)
 고 석 준 (제주대학교 교수)
 강 윤 희 (백석대학교 교수)
 최 명 준 (텔레다인 박사)

광주·전남지부
 대전·충남지부
 전 북 지 부
 충 북 지 부
 일 본
 러 시 아 지 부

최 수 일 (전남대학교 교수)
 김 철 영 (충남대학교 교수)
 이 주 연 (전주비전대학교 교수)
 조 문 규 (한국교통대학교 교수)
 강 유 선 (Tokyo Polytechnic University 교수)
 Prof. Edis B. TEN (National University of Science and Technology)

위원회 명단

자문위원회

위원장	박 항 구 (소암시스템 회장)
부위원장	박 성 한 (한양대학교 명예교수)
위원	고 성 제 (고려대학교 명예교수)
	김 성 대 (KAIST 명예교수)
	김 재 희 (연세대학교 명예교수)
	박 규 태 (연세대학교 명예교수)
	서 승 우 (서울대학교 교수)
	이 문 기 (연세대학교 명예교수)
	이 진 구 (동국대학교 명예교수)
	임 혜 속 (이화여자대학교 교수)
	정 정 화 (한양대학교 명예교수)
	홍 승 홍 (인하대학교 명예교수)

공 준 진 (삼성전자공과대학교 교수)	구 용 서 (단국대학교 석좌교수)
김 수 중 (경북대학교 명예교수)	김 영 권 (건국대학교 명예교수)
나 정 웅 (KAIST 명예교수)	문 영 식 (한양대학교 명예교수)
박 진 옥 (육군사관학교 명예교수)	백 준 기 (중앙대학교 교수)
성 광 모 (서울대학교 명예교수)	윤 종 웅 (한국공학교육인증원 이사장)
이 상 설 (한양대학교 명예교수)	이 재 홍 (서울대학교 명예교수)
이 혁 재 (서울대학교 교수)	임 제 탁 (한양대학교 명예교수)
전 국 진 (서울대학교 명예교수)	전 홍 태 (중앙대학교 명예교수)
천 경 준 (씨젠 회장)	홍 대 식 (연세대학교 교수)

기획위원회

위원장	강 명 곤 (서울시립대학교 교수)
부위원장	김 현 (서울과학기술대학교 교수)
위원	구 민 석 (서울시립대학교 교수)
	배 종 호 (국민대학교 교수)
	조 문 규 (한국교통대학교 교수)

권 구 덕 (강원대학교 교수)	권 민 우 (강릉원주대학교 교수)
송 익 현 (한양대학교 교수)	심 원 보 (서울과학기술대학교 교수)
채 주 형 (광운대학교 교수)	추 민 성 (한양대학교 교수)

학술연구위원회 - 하계

위원장	김 훈 (인천대학교 교수)
위원	강 문 식 (강릉원주대학교 교수)
	김 영 민 (서울대학교 교수)
	이 권 형 (LG전자 팀장)
	정 승 원 (고려대학교 교수)
	한 상 옥 (삼성전자 연구원)

김 근 영 (한국전자통신연구원 연구원)	김 도 현 (제주대학교 교수)
김 익 군 (한국전자통신연구원 본부장)	김 혁 (서울시립대학교 교수)
인 치 호 (세명대학교 교수)	정 성 업 (고려대학교 교수)
조 성 현 (한양대학교 교수)	진 성 훈 (인천대학교 교수)
한 재 호 (고려대학교 교수)	

학술연구위원회 - 추계

위원장	이 강 윤 (성균관대학교 교수)
위원	고 형 호 (충남대학교 교수)
	박 관 서 (연세대학교 교수)
	하 태 준 (광운대학교 교수)

권 중 원 (한국산업기술시험원 책임연구원)	김 승 환 (한국전자통신연구원 책임)
배 현 철 (한국전자통신연구원 책임연구원)	오 윤 호 (고려대학교 교수)

논문편집위원회

위원장	조 성 현 (한양대학교 교수)
위원	강 성 복 (한국생산기술연구원 수석연구원)
	권 중 원 (한국산업기술시험원 책임연구원)
	김 상 범 (한국폴리텍대학 교수)
	김 영 선 (대림대학교 교수)
	여 영 정 (연세대학교 교수)
	이 덕 우 (계명대학교 교수)
	조 성 인 (동국대학교 교수)

강 제 원 (이화여자대학교 교수)	권 구 덕 (강원대학교 교수)
김 기 연 (한국산업기술시험원 선임연구원)	김 명 선 (한성대학교 교수)
김 선 용 (건국대학교 교수)	김 영 로 (명진전문대학 교수)
김 학 구 (중앙대학교 교수)	심 규 성 (한경국립대학교 교수)
유 재 준 (UNIST 교수)	윤 중 현 (조선이공대학교 교수)
이 중 호 (숭실대학교 교수)	임 민 중 (동국대학교 교수)
최 재 용 (가천대학교 교수)	추 민 성 (한밭대학교 교수)

국제협력위원회

위원장	노 원 우 (연세대학교 교수)
위원	김 태 욱 (연세대학교 교수)
	한 진 호 (한국전자통신연구원 책임연구원)

박 성 정 (건국대학교 교수)	이 창 우 (가톨릭대학교 교수)
------------------	-------------------

대외협력위원회

위원장	김 용 신 (고려대학교 교수)
위원	궁 재 하 (고려대학교 교수)
	김 정 석 (가천대학교/셀리코 교수/대표이사)
	최 광 성 (한국전자통신연구원 실장)

김 사 혁 (KISDI 책임연구원)	김 성 우 (서울대학교 교수)
안 광 호 (한국전자기술연구원 본부장)	채 관 업 (삼성전자 마스터)
홍 성 완 (서강대학교 교수)	

회원관리위원회

위원장	선우 경 (서울대학교 교수)		
위원	김 건 우 (목포대학교 교수)	김 기 현 (전북대학교 교수)	김 대 영 (순천향대학교 교수)
	김 민 주 (단국대학교 교수)	김 민 휘 (중앙대학교 교수)	김 형 진 (한양대학교 교수)
	문 상 미 (나사렛대학교 교수)	송 익 현 (한양대학교 교수)	심 민 섭 (경상국립대학교 교수)
	유 경 창 (삼성전자 수석연구원)	이 정 원 (서울대학교 교수)	

회지편집위원회

위원장	황 원 준 (아주대학교 교수)		
부위원장	김 영 로 (명지전문대학 교수)		
위원	김 상 범 (한국폴리텍대학 교수)	박 승 창 (유오씨 대표)	박 진 선 (부산대학교 교수)
	심 규 성 (한경국립대학교 교수)	안 성 수 (명지전문대학 교수)	윤 중 현 (조선이공대학교 교수)
	이 정 석 (인하공업전문대학 교수)	이 철 (동국대학교 교수)	정 승 원 (고려대학교 교수)
	최 욱 (인천대학교 교수)		

사업위원회

위원장	유 찬 세 (한국전자기술연구원 수석연구원)		
부위원장	권 혁 인 (중앙대학교 교수)		
위원	김 윤 (서울시립대학교 교수)	김 중 현 (고려대학교 교수)	류 은 석 (성균관대학교 교수)
	선우명훈 (아주대학교 교수)	손 교 민 (삼성전자 마스터)	연 규 봉 (한국자동차연구원 팀장)
	정 일 권 (한국전자통신연구원 본부장)		

교육연구위원회

위원장	변 대 석 (삼성전자 마스터)	동 성 수 (용인예술과학대학교 교수)	
위원	김 선 욱 (고려대학교 교수)	김 혁 (서울시립대학교 교수)	류 현 석 (서울대학교 교수)
	박 영 우 (TEL 부사장)	변 영 재 (UNIST 교수)	윤 종 윤 (파두 대표이사)
	이 영 택 (ASML 전문)	이 후 진 (한성대학교 교수)	정 성 업 (고려대학교 교수)
	최 병 수 (한국전자통신연구원 박사)		

홍보/정보화위원회

위원장	서 병 석 (상지대학교 교수)		
위원	공 규 열 (한성대학교 교수)	권 태 수 (서울과학기술대학교 교수)	김 범 현 (한양대학교 교수)
	배 준 성 (강원대학교 교수)	서 민 재 (서울시립대학교 교수)	서 봉 상 (울포랜드 상무)
	송 준 영 (인천대학교 교수)	유 성 철 (쌍용정보통신 상무)	윤 희 인 (UNIST 교수)
	이 구 순 (파이낸셜뉴스 이사)	이 태 동 (국제대학교 교수)	임 동 구 (전북대학교 교수)
	정 방 철 (충남대학교 교수)	정 완 영 (KAIST 교수)	황 진 영 (한국항공대학교 교수)

표준화위원회

위원장	김 원 종 (한국전자통신연구원 책임연구원)		
부위원장	연 규 봉 (한국자동차연구원 팀장)		
위원 (간사)	김 도 훈 (한국전자통신연구원 박사)		
위원	권 기 원 (성균관대학교 교수)	김 성 동 (서울과학기술대학교 교수)	박 재 영 (광운대학교 교수)
	배 준 호 (가천대학교 교수)	이 상 근 (성균관대학교 교수)	좌 성 훈 (서울과학기술대학교 교수)
	차 철 웅 (한국전자기술연구원 센터장)		

신기술위원회

위원장	한 재 호 (고려대학교 교수)		
위원	고 정 길 (연세대학교 교수)	김 혁 (서울시립대학교 교수)	김 현 우 (고려대학교 교수)
	송 철 (DGIST 교수)	원 용 욱 (명지대학교 교수)	윤 상 민 (국민대학교 교수)
	장 현 재 (KIST 선임연구원)		

지부담당위원회

위원장	이 승 호 (한밭대학교 교수)		
위원	강 문 식 (강릉원주대학교 교수)	강 윤 희 (백석대학교 교수)	고 석 준 (제주대학교 교수)
	고 진 환 (경상국립대학교 교수)	김 철 영 (충남대학교 교수)	이 주 연 (전주비전대학교 교수)
	이 찬 수 (영남대학교 교수)	최 수 일 (전남대학교 교수)	조 문 규 (한국교통대학교 교수)

선거관리위원회

위 원 장	전 홍 태 (중앙대학교 교수)		
위 원	강 명 곤 (서울시립대학교 교수)	강 석 주 (서강대학교 교수)	조 성 재 (이화여자대학교 교수)
	조 성 현 (한양대학교 교수)	한 재 호 (고려대학교 교수)	

포상위원회

위 원 장	임 해 숙 (이화여자대학교 교수)		
위 원	강 명 곤 (서울시립대학교 교수)	김 훈 (인천대학교 교수)	백 광 현 (중앙대학교 교수)
	조 성 재 (이화여자대학교 교수)	조 성 현 (한양대학교 교수)	
위원 및 간사겸임	강 석 주 (서강대학교 교수)		

인사위원회

위 원 장	이 총 용 (연세대학교 교수)		
위 원	강 석 주 (서강대학교 교수)	김 현 (서울과학기술대학교 교수)	백 광 현 (중앙대학교 교수)
	조 성 재 (이화여자대학교 교수)		

JSTS 편집위원회

위 원 장	김 재 준 (서울대학교 교수)		
위 원	강 명 곤 (서울시립대학교 교수)	강 인 만 (경북대학교 교수)	권 혁 인 (중앙대학교 교수)
	김 상 범 (서울대학교 교수)	김 상 완 (서강대학교 교수)	김 소 영 (성균관대학교 교수)
	김 윤 (서울시립대학교 교수)	김 지 훈 (이화여자대학교 교수)	남 일 구 (부산대학교 교수)
	민 경 식 (국민대학교 교수)	서 문 교 (성균관대학교 교수)	신 민 철 (KAIST 교수)
	이 가 원 (충남대학교 교수)	이 강 윤 (성균관대학교 교수)	전 동 석 (서울대학교 교수)
	정 재 경 (한양대학교 교수)	조 성 재 (이화여자대학교 교수)	조 일 환 (명지대학교 교수)
	차 호 영 (홍익대학교 교수)	최 우 석 (서울대학교 교수)	최 우 영 (서울대학교 교수)
	한 대 덕 (한양대학교 교수)		

SPC위원회

위 원 장	박 철 수 (광운대학교 교수)	송 병 철 (인하대학교 교수)	
자 문 위 원	김 선 옥 (고려대학교 교수)	김 창 수 (고려대학교 교수)	백 준 기 (중앙대학교 교수)
	심 동 규 (광운대학교 교수)	임 해 숙 (이화여자대학교 교수)	전 병 우 (성균관대학교 교수)
	조 남 익 (서울대학교 교수)	조 민 호 (고려대학교 교수)	
운영위원 (간사)	공 경 보 (부산대학교 교수)	조 성 인 (동국대학교 교수)	홍 승 혁 (수원대학교 교수)
운 영 위 원	김 종 옥 (고려대학교 교수)	유 양 모 (서강대학교 교수)	최 강 선 (한국기술교육대학교 교수)
운영위원 겸 편집위원	강 석 주 (서강대학교 교수)	김 영 민 (홍익대학교 교수)	김 원 준 (건국대학교 교수)
	김 재 곤 (한국항공대학교 교수)	서 영 호 (광운대학교 교수)	이 채 은 (한양대학교 교수)
	정 승 원 (고려대학교 교수)	황 원 준 (아주대학교 교수)	황 인 철 (강원대학교 교수)
편 집 위 원	권 건 우 (홍익대학교 교수)	권 준 석 (중앙대학교 교수)	김 병 서 (홍익대학교 교수)
	김 영 빈 (중앙대학교 교수)	김 용 태 (경북대학교 교수)	김 원 중 (한국전자통신연구원 책임연구원)
	김 진 술 (전남대학교 교수)	김 태 석 (광운대학교 교수)	김 태 환 (한국항공대학교 교수)
	김 현 (서울과학기술대학교 교수)	남 일 구 (부산대학교 교수)	민 경 식 (국민대학교 교수)
	박 승 창 (유오씨 대표이사)	신 영 주 (고려대학교 교수)	신 오 순 (송실대학교 교수)
	엄 찬 호 (중앙대학교 교수)	오 지 형 (중앙대학교 교수)	오 태 현 (POSTECH 교수)
	오 현 우 (쓰리웨이소프트 기업 연구소장)	우 성 민 (한국기술교육대학교 교수)	유 지 현 (광운대학교 교수)
	윤 은 준 (경일대학교 교수)	이 재 훈 (고려대학교 교수)	이 철 (동국대학교 교수)
	이 후 진 (한성대학교 교수)	이 훈 (부경대학교 교수)	임 성 훈 (DGIST 교수)
	장 승 진 (한밭대학교 교수)	장 주 용 (광운대학교 교수)	장 훈 석 (한국전자기술연구원 선임연구원)
	장 희 선 (평택대학교 교수)	정 성 윤 (영남대학교 교수)	조 성 현 (한양대학교 교수)
	진 훈 (경기대학교 교수)	최 상 호 (광운대학교 교수)	황 성 운 (가천대학교 교수)

바이오-메디컬연구회

위 원 장	김 동 현 (연세대학교 교수)
-------	------------------

Society 명단

통신소사이터티

회 장	신 오 순 (숭실대학교 교수)	김 재 현 (아주대학교 교수)	김 진 영 (광운대학교 교수)
부 회 장	김 선 용 (건국대학교 교수)	오 정 근 (ATNS 대표이사)	윤 석 현 (단국대학교 교수)
	김 훈 (인천대학교 교수)	이 재 진 (숭실대학교 교수)	이 정 우 (중앙대학교 교수)
	이 인 규 (고려대학교 교수)	허 준 (고려대학교 교수)	
감 사	최 천 원 (단국대학교 교수)	이 흥 노 (광주과학기술원 교수)	
협 동 부 회 장	유 명 식 (숭실대학교 교수)	김 연 은 (브로던 대표이사)	김 영 한 (숭실대학교 교수)
	김 병 남 (에이스테크놀로지 연구소장)	방 승 찬 (한국전자통신연구원 원장)	연 철 홀 (LGT 상무)
	김 용 석 (답스 대표이사)	장 병 수 (이노벨류네트웍스 부사장)	정 현 규 (한국전자통신연구원 부장)
	이 재 훈 (유정시스템 대표이사)		
	조 인 호 (에이스테크놀로지 박사)	김 성 훈 (한국전자통신연구원 박사)	김 정 호 (이화여자대학교 교수)
이 사	김 광 순 (연세대학교 교수)	김 중 현 (고려대학교 교수)	서 철 현 (숭실대학교 교수)
	김 준 수 (한국공항공대학교 교수)	손 인 수 (동국대학교 교수)	신 수 용 (금오공과대학교 교수)
	성 원 진 (서강대학교 교수)	윤 중 호 (한국항공대학교 교수)	윤 지 훈 (서울과학기술대학교 교수)
	신 요 안 (숭실대학교 교수)	이 재 훈 (동국대학교 교수)	이 중 호 (숭실대학교 교수)
	이 예 훈 (서울과학기술대학교 교수)	임 중 태 (홍익대학교 교수)	장 석 호 (건국대학교 교수)
	이 호 경 (홍익대학교 교수)	조 성 래 (중앙대학교 교수)	조 성 현 (한양대학교 교수)
	정 소 이 (아주대학교 교수)	최 진 식 (한양대학교 교수)	허 서 원 (홍익대학교 교수)
	최 지 응 (대구경북과학기술원 교수)	황 승 훈 (동국대학교 교수)	황 인 태 (전남대학교 교수)
간 사	홍 인 기 (경희대학교 교수)	정 소 이 (아주대학교 교수)	
연구회위원장	김 중 현 (고려대학교 교수)	윤 상 민 (국민대학교 교수) - 지능형네트워크	
	최 영 준 (아주대학교 교수) - 통신	이 철 기 (아주대학교 교수) - ITS	
	조 춘 식 (한국항공대학교 교수) - 마이크로파 및 전파전자	허 재 두 (한국전자통신연구원 연구전문위원) - 무선 PAN/BAN	
	김 강 욱 (경북대학교 교수) - 군사전자		

반도체소사이터티

회 장	장 성 진 (삼성전자 상근고문)	권 오 경 (한양대학교 교수, 반소전임회장)	김 영 환 (POSTECH 교수)
자 문 위 원	공 준 진 (삼성전자 마스터/ 교수)	김 진 상 (경희대학교 총장, 반소전임회장)	김 희 석 (청주대학교 교수)
	김 재 석 (연세대학교 교수)	선우명훈 (아주대학교 교수, 반소전임회장)	손 보 익
	박 홍 준 (POSTECH 교수)	신 현 철 (한양대학교 교수, 반소전임회장)	우 남 성 (반소전임회장)
	신 승 송 (반소전임회장)	임 신 일 (서경대학교 교수)	임 형 규 (반소전임회장)
	이 승 훈 (서강대학교 교수)	정 연 모 (경희대학교 교수)	정 항 근 (전북대학교 교수)
	전 영 현 (삼성전자 부회장, 반소전임회장)	조 경 순 (한국외국어대학교 교수)	조 상 복 (울산대학교 교수)
	소 경 신 (시놀시스 사장)	최 기 영 (서울대학교 교수)	최 승 중 (웰랑 대표이사)
	조 중 휘 (인천대학교 교수, 반소전임회장)	허 열 (실리콘마이터스 회장)	
감 사	이윤중 (KAIST 교수)	이 광 엽 (서경대학교 교수)	
부 회 장	이 강 윤 (성균관대학교 교수)	고 대 협 (LX세미콘 연구소장)	이 한 호 (인하대학교 교수)
	박 영 우 (EXICON 사장)	최 중 호 (서울시립대학교 교수)	김 동 규 (한양대학교 교수)
	이 희 덕 (충남대학교 교수)		
	안 기 현 (한국반도체산업협회 전무)	김 지 훈 (이화여자대학교 교수)	김 현 (서울과학기술대학교 교수)
총 무 이 사	고 형 호 (충남대학교 교수)	박 종 선 (고려대학교 교수)	윤 찬 호 (삼성전자 마스터)
	류 현 석 (서울대학교 교수)		
	황 상 준 (삼성전자 부사장)	유 창 식 (삼성전자 부사장)	조 성 재 (이화여자대학교 교수)
편 집 이 사	노 정 진 (한양대학교 교수)		
	한 태 희 (성균관대학교 교수)	김 철 우 (고려대학교 교수)	범 진 욱 (서강대학교 교수)
학 술 이 사	강 진 구 (인하대학교 교수)	송 민 구 (동국대학교 교수)	이 병 훈 (POSTECH 교수)
	변 영 재 (UNIST 교수)	이 혁 재 (서울대학교 교수)	이 희 덕 (충남대학교 교수)
	이 승 호 (한밭대학교 교수)	정 진 균 (전북대학교 교수)	차 호 영 (홍익대학교 교수)
	인 치 호 (세명대학교 교수)	최 창 환 (한양대학교 교수)	
	최 우 영 (연세대학교 교수)	공 배 선 (성균관대학교 교수)	공 정 택 (성균관대학교 교수)
사 업 이 사	강 운 병 (삼성전자 마스터)	김 소 영 (성균관대학교 교수)	김 시 호 (연세대학교 교수)
	김 동 순 (세종대학교 교수)	김 원 중 (한국전자통신연구원 실장)	김 종 선 (홍익대학교 교수)
	김 용 석 (가천대학교 석좌교수)	백 광 현 (중앙대학교 교수)	변 대 석 (삼성전자 마스터)
	나 영 선 (LX세미콘 상무)	송 용 호 (삼성전자 부사장)	엄 낙 웅 (한국전자통신연구원 책임연구원)
	손 고 민 (삼성전자 마스터)	이 상 훈 (웨이브피아 대표)	조 태 제 (DGIST 교수)
	오 정 우 (연세대학교 교수)	최 병 호 (한국전자기술연구원 연구소장)	최 윤 경 (고려대학교 교수)
	최 규 명 (서울대학교 교수)		
	최 준 립 (경북대학교 교수)	이 성 수 (숭실대학교 교수)	
재 무 이 사	권 기 원 (성균관대학교 교수)	김 동 현 (ICTK 사장)	김 보 은 (라온텍 사장)
산 학 이 사	김 경 수 (넥스트칩 대표이사)	나 준 호 (실리콘웍스 전무)	손 재 철
	김 준 석 (ADT 사장)	신 용 석 (케이던스코리아 사장)	오 문 욱 (삼성전자 부사장)
	송 태 훈 (휴인스 사장)	이 장 규 (텔레칩스 대표이사)	
	이 도 영 (옵토레인 사장)	문 용 (숭실대학교 교수)	
회 원 이 사	노 원 우 (연세대학교 교수)	김 상 인 (아주대학교 교수) - 광파워전자자공학	
연구회위원장	최 우 영 (서울대학교 교수) - 반도체소재및재료	황 인 철 (강원대학교 교수) - RF집적회로	
	이 강 윤 (성균관대학교 교수) - SoC설계	김 의 균 (한국전자통신연구원 본부장) - 정보보안시스템	
	정 원 영 (강원공업 본부장) - PCB&Package	김 한 구 (EESOP 대표) - ESD/EOS & Latchup	
	정 의 준 (경희대학교 교수) - 내방사전 반도체 설계 및 소자	송 민 규 (동국대학교 교수) - 이미지센서	
	노 원 우 (연세대학교 교수) - 인 메모리 컴퓨팅	권 구 덕 (강원대학교 교수)	
협 동 위 원	강 명 곤 (서울시립대학교 교수)	김 수 연 (동국대학교 교수)	김 영 민 (홍익대학교 교수)
	권 영 수 (한국전자통신연구원 책임연구원)	김 태 환 (한국항공대학교 교수)	류 성 주 (서강대학교 교수)
	김 재 욱 (KIST 그룹장)	송 준 영 (인천대학교 교수)	양 준 성 (연세대학교 교수)
	박 성 정 (건국대학교 교수)	윤 명 국 (이화여자대학교 교수)	이 영 주 (POSTECH 교수)
	오 윤 호 (성균관대학교 교수)	이 윤 명 (성균관대학교 교수)	이 형 민 (고려대학교 교수)
	이 우 주 (중앙대학교 교수)	전 성 훈 (삼성전자 상무)	정 무 경 (SK사피온 CTO)
	전 동 석 (서울대학교 교수)	제 민 규 (KAIST 교수)	채 형 일 (건국대학교 교수)
	정 윤 호 (항공대학교 교수)		

최재혁 (서울대학교 교수)
한정환 (충남대학교 교수)

최재혁 (성균관대학교 교수)
황태호 (한국전자기술연구원 박사)

추민성 (한양대학교 교수)

컴퓨터사이버티

회예회장

최수식 (신한대학교 교수)
김민우 (강릉원주대학교 교수)
박인호 (단국대학교 명예교수)
정영호 (동인대학교 교수)
정영호 (울지대학교 교수)
정영호 (가천대학교 교수)
정영호 (강원대학교 교수)
정영호 (국민대학교 교수)
정영호 (제주대학교 교수)
정영호 (고려대학교 교수)
정영호 (상명대학교 교수)
정영호 (강릉원주대학교 교수)
정영호 (인천대학교 교수)
정영호 (명지병원 책임연구원)
정영호 (울지대학교 교수)
정영호 (수명여자대학교 교수)
정영호 (한성대학교 교수)
정영호 (순천향대학교 교수)
정영호 (경희대학교 교수)
정영호 (서울과학기술대학교 교수)
정영호 (한국전자기술연구원 책임연구원)
정영호 (전북대학교 교수)
정영호 (김포대학교 교수)
정영호 (고려대학교 교수)
정영호 (전북대학교 교수)
정영호 (고려대학교 교수)
정영호 (조선대학교 교수)
정영호 (경기대학교 겸임교수)
정영호 (부경대학교 교수)
정영호 (군산대학교 교수)
정영호 (유오씨 대표이사)
정영호 (네이버시스템즈 대표이사)
정영호 (비온이노베이션 대표이사)
정영호 (연세세브란스병원 연구원)
정영호 (대보정보통신 부장)
정영호 (유비벨록스모바일 대표이사)
정영호 (세림TSG 전무)

김승천 (한성대학교 교수)
박명진 (한국교통대학교 교수)
박명진 (동명대학교 교수)
박명진 (스마트의료기기산업진흥재단 부이사장)
남상엽 (국제대학교 교수)
심정연 (UNIST 교수)
심정연 (강남대학교 교수)
김병서 (홍익대학교 교수)
우태진 (KAIST 교수)
이후진 (한성대학교 교수)
김장근 (공주대학교 교수)
노소현 (배재대학교 교수)
김민우 (월송출판 교수)
이민우 (김포대학교 교수)
이민우 (가천대학교 교수)
김민우 (금오공과대학교 교수)
김민우 (한성대학교 교수)
김민우 (다스퍼워 이사)
김민우 (연안공과대학교 교수)
김민우 (대림대학교 교수)
김민우 (신한대학교 교수)
김민우 (연세의료원 팀장)
차시호 (청운대학교 교수)
김대희 (한국정보통신 대표이사)
김대희 (투아시스템즈 이사)
김대희 (올프랜드 이사)
김대희 (주얼리 대표이사)
이학준 (이노지스스 코리아 연구소장)
조병준 (CNCInstrument 사장)

김형중 (고려대학교 교수)
신인철 (단국대학교 명예교수)
이규대 (공주대학교 교수)
홍유식 (상지대학교 교수)
정교일 (한국전자통신연구원 책임연구원)
윤은준 (경일대학교 교수)
김영학 (산업기술평가관리원 본부장)
유철성 (쌍용정보통신주식회사 본부장)
정민성 (홍익대학교 교수)
김선욱 (고려대학교 교수)
김선욱 (세종대학교 교수)
심민호 (한경국립대학교 교수)
이민호 (경북대학교 교수)
이찬수 (영남대학교 교수)
김종욱 (경동대학교 교수)
김종욱 (중앙대학교 교수)
이정선 (울지대학교 교수)
이재민 (명지병원 소장)
이재민 (순천향대학교 교수)
홍상중 (SK Telecom 박사)
황진영 (한국항공대학교 교수)
김성길 (우주텔레콤 이사)
김진수 (지엔소프트 대표이사)
송지훈 (웨이버스 이사)
이영욱 (진우ATS 팀장)
임준성 (대전정보통신 차장)
조병영 (태진인포텍 전무)

연구회위원장

한태주 (연세대의원 팀장) - 멀티미디어
최주환 (경일대학교 교수) - 융합컴퓨팅
김도현 (제주대학교 교수) - M2M/IOT
김성준 (가천대학교 교수) - 인공지능 및 보안
김영선 (한성대학교 교수) - 시 응용

진훈 (경기대학교 겸임교수) - 휴먼ICT
이민호 (경북대학교 교수) - 인공지능/신경망/퍼지
우영택 (KAIST 교수) - 증강현실
정민성 (홍익대학교 교수) - 블록체인

인공지능 신호처리사이버티

회예회장

김민우 (계명대학교 교수)
김민우 (이화여자대학교 교수)
김민우 (광주과학기술원 교수)
김민우 (광운대학교 교수)
김민우 (서울대학교 교수)
김민우 (서강대학교 교수)
김민우 (DGIST 교수)
김민우 (고려대학교 교수)
김민우 (충북대학교 교수)
김민우 (이주대학교 교수)
김민우 (한국전자통신연구원 연구원)
김민우 (성균관대학교 교수)
김민우 (경희대학교 교수)
김민우 (경일대학교 교수)
김민우 (세종대학교 교수)
김민우 (서울과학기술대학교 교수)
김민우 (성균관대학교 교수)
김민우 (광운대학교 교수)
김민우 (연세대학교 교수)
김민우 (성균관대학교 교수)
김민우 (연세대학교 교수)
김민우 (POSTECH 교수)
김민우 (KAIST 교수)
김민우 (중앙대학교 교수)
김민우 (연세대학교 교수)
김민우 (서울과학기술대학교 교수)
김민우 (연세대학교 교수)
김민우 (계명대학교 교수)
김민우 (한국기술교육대학교 교수)
김민우 (한국전자기술연구원 센터장)
김민우 (국기수리과학연구소 박사)
김민우 (고려대학교 교수)
김민우 (영남대학교 교수)
김민우 (네이버 연구원)
김민우 (연세대학교 교수)
김민우 (영남대학교 교수)
김민우 (전남대학교 교수)
김민우 (충북대학교 교수)
김민우 (서울대학교 교수)
김민우 (중앙대학교 교수)

김민우 (고려대학교 교수)
김민우 (한양대학교 교수)
김민우 (세종대학교 교수)
김민우 (송실대학교 교수)
김민우 (이주대학교 교수)
김민우 (동국대학교 교수)
곽준 (서울대학교 교수)
권구락 (조선대학교 교수)
김기백 (송실대학교 교수)
김기백 (연세대학교 교수)
김재준 (한국항공대학교 교수)
김재준 (한국공공연구원 연구원)
김재준 (한국전자통신연구원 연구원)
박성우 (명지대학교 교수)
박성우 (광운대학교 교수)
박성우 (한국전자통신연구원 교수)
송진호 (연세대학교 교수)
신학일 (캐논메디칼시스템즈코리아 박사)
신학일 (부산대학교 교수)
신학일 (인텔리비스 대표이사)
신학일 (한국외국어대학교 교수)
신학일 (연세대학교 교수)
신학일 (POSTECH 교수)
신학일 (한국항공대학교 교수)
신학일 (한국전자기술연구원 책임연구원)
신학일 (성균관대학교 교수)
신학일 (제주대학교 교수)
신학일 (제주한라대학교 교수)
신학일 (광주과학기술원 교수)
신학일 (한밭대학교 교수)
신학일 (HDXMILL 박사)
신학일 (제주대학교 교수)
신학일 (이화여자대학교 교수)
신학일 (세종대학교 교수)
신학일 (연세대학교 교수)
신학일 (LG전자 연구원)
신학일 (한국전자통신연구원 그룹장)
신학일 (KAIST 교수)

김창익 (KAIST 교수)
송철우 (인하대학교 교수)
전병우 (성균관대학교 교수)
김승룡 (고려대학교 교수)
곽진태 (고려대학교 교수)
권진석 (중앙대학교 교수)
김현현 (연세대학교 교수)
김현현 (한국전자기술연구원 수석연구원)
김진우 (계명대학교 교수)
김진우 (경성대학교 교수)
박진우 (서울과학기술대학교 교수)
박성우 (KAIST 교수)
배성우 (경희대학교 교수)
서진근 (연세대학교 교수)
신재섭 (픽스트리 대표이사)
양현중 (UNIST 교수)
오세훈 (한국외국어대학교 교수)
유양모 (서강대학교 교수)
이기 (건국대학교 교수)
이성현 (DGIST 교수)
이성현 (경북대학교 교수)
이성현 (서울대학교 교수)
이성현 (계명대학교 교수)
이성현 (카톨릭대학교 교수)
이성현 (한양대학교 교수)
이성현 (한국전자통신연구원 연구원)
정미라 (대구과학관 교수)
정호기 (한국교통대학교 교수)
조승룡 (KAIST 교수)
최승룡 (서울과학기술대학교 교수)
최주환 (한양대학교 교수)
최주환 (이주대학교 교수)
최주환 (KAIST 교수)
최주환 (부경대학교 교수)
최주환 (고려대학교 교수)
최주환 (LG전자 연구원)

[illegible]

제24대 평의원 명단

강명곤 (서울시립대학교 교수)
 강석판 (LG전자 상무)
 강제원 (이화여자대학교 교수)
 고성환 (고려대학교 명예교수)
 고진환 (경상대학교 교수)
 공배선 (성균관대학교 교수)
 구민석 (서울시립대학교 교수)
 궁재하 (고려대학교 교수)
 권구덕 (강원대학교 교수)
 권태수 (서울과학기술대학교 교수)
 김강태 (삼성전자 부사장)
 김기현 (전북대학교 교수)
 김도훈 (한국전자통신연구원 박사)
 김동현 (ICT 대표이사)
 김민휘 (중앙대학교 교수)
 김상태 (전남테크노파크 본부장)
 김성태 (한국과학기술원 명예교수)
 김수중 (경북대학교 명예교수)
 김시호 (연세대학교 교수)
 김영민 (홍익대학교 교수)
 김영재 (해동과학문화재단 이사장)
 김원종 (한국전자통신연구원 책임연구원)
 김은원 (대림대학교 교수)
 김재현 (아주대학교 교수)
 김종욱 (고려대학교 교수)
 김진상 (경희대학교 교수)
 김철우 (고려대학교 교수)
 김한구 (EESOP 대표)
 김현 (부천대학교 교수)
 김형진 (한양대학교 교수)
 나정웅 (한국과학기술원 명예교수)
 남일구 (부산대학교 교수)
 노정진 (한양대학교 교수)
 류성주 (서강대학교 교수)
 문상미 (나사렛대학교 교수)
 민경식 (국민대학교 교수)
 박규태 (연세대학교 명예교수)
 박성정 (건국대학교 교수)
 박영훈 (숙명여자대학교 교수)
 박종일 (한양대학교 교수)
 박철수 (광운대학교 교수)
 배순민 (케이티 Lab장)
 백광현 (중앙대학교 교수)
 범진욱 (서강대학교 교수)
 서민재 (서울시립대학교 교수)
 서승우 (서울대학교 교수)
 서지원 (연세대학교 교수)
 성광모 (서울대학교 명예교수)
 송민규 (동국대학교 교수)
 송익현 (한양대학교 교수)
 신세운 (포항공과대학교 교수)
 신창환 (고려대학교 교수)
 심용 (중앙대학교 교수)
 안상철 (한국과학기술연구원 책임연구원)
 안호균 (한국전자통신연구원 실장)
 연규봉 (한국자동차연구원 수석연구원)
 오태현 (포항공과대학교 교수)

강문식 (강릉원주대학교 교수)
 강성원 (한국전자통신연구원 부원장)
 고병철 (계명대학교 교수)
 고용남 (하나미크론 전무)
 고형호 (충남대학교 교수)
 공준진 (삼성전자공과대학교 주임교수)
 구본태 (한국전자통신연구원 본부장)
 권건우 (홍익대학교 교수)
 권기룡 (부경대학교 교수)
 권혁인 (중앙대학교 교수)
 김경기 (대구대학교 교수)
 김도현 (제주대학교 교수)
 김동규 (한양대학교 교수)
 김동현 (연세대학교 교수)
 김병서 (홍익대학교 교수)
 김선용 (건국대학교 교수)
 김성우 (서울대학교 교수)
 김승천 (한성대학교 교수)
 김영권 (후레대학교 명예총장)
 김영민 (서울대학교 교수)
 김영한 (UC San Diego/가우스랩스 대표이사)
 김원준 (건국대학교 교수)
 김익균 (한국전자통신연구원 본부장)
 김재희 (연세대학교 명예교수)
 김중현 (고려대학교 교수)
 김진영 (광운대학교 교수)
 김태욱 (연세대학교 교수)
 김혁 (서울시립대학교 교수)
 김현수 (삼성전자 상무)
 김형탁 (홍익대학교 교수)
 남광희 (포항공과대학교 교수)
 노승원 (LGI노텍 CTO)
 노태문 (한국전자통신연구원 연구전문위원)
 류수정 (서울대학교 교수)
 문영식 (한양대학교 명예교수)
 민동보 (이화여자대학교 교수)
 박성욱 (차세대지능형반도체사업단 이사장)
 박성한 (한양대학교 명예교수)
 박인규 (인하대학교 교수)
 박준석 (인하대학교 교수)
 박항구 (소암시스템 회장)
 배준성 (강원대학교 교수)
 백종덕 (연세대학교 교수)
 변대석 (삼성전자 마스터)
 서병석 (상지대학교 교수)
 서정목 (연세대학교 교수)
 서철현 (숭실대학교 교수)
 손교민 (삼성전자 마스터)
 송병철 (인하대학교 교수)
 송준영 (인천대학교 교수)
 신오순 (숭실대학교 교수)
 신현철 (광운대학교 교수)
 안광호 (한국전자기술연구원 본부장)
 안성수 (명지전문대학 교수)
 양준성 (연세대학교 교수)
 오윤제 (정보통신기획평가원 PM)
 우성민 (한국기술교육대학교 교수)

강석주 (서강대학교 교수)
 강윤희 (백석대학교 교수)
 고석준 (제주대학교 교수)
 고정환 (인하공업전문대학 교수)
 공규열 (한성대학교 교수)
 곽진태 (고려대학교 교수)
 구용서 (단국대학교 석좌교수)
 권경하 (한국과학기술원 교수)
 권종원 (한국산업기술시험원 책임연구원)
 권호열 (강원대학교 교수)
 김기연 (한국산업기술시험원 선임연구원)
 김도현 (국민대학교 명예교수)
 김동식 (인하공업전문대학 교수)
 김명선 (한성대학교 교수)
 김상범 (한국폴리텍대학 교수)
 김선욱 (고려대학교 교수)
 김소영 (성균관대학교 교수)
 김승환 (한국전자통신연구원 책임)
 김영로 (명지전문대학 교수)
 김영선 (대림대학교 교수)
 김용신 (고려대학교 교수)
 김윤 (서울시립대학교 교수)
 김재준 (서울대학교 교수)
 김중선 (홍익대학교 교수)
 김지훈 (이화여자대학교 교수)
 김철영 (충남대학교 교수)
 김학구 (중앙대학교 교수)
 김현 (서울과학기술대학교 교수)
 김형준 (한국과학기술연구원 소장)
 김훈 (인천대학교 교수)
 남기창 (동국대학교 교수)
 노원우 (연세대학교 교수)
 동성수 (용인예술과학대학교 교수)
 류은석 (성균관대학교 교수)
 문용 (숭실대학교 교수)
 박관서 (연세대학교 교수)
 박성욱 (강릉원주대학교 교수)
 박수현 (국민대학교 교수)
 박종선 (고려대학교 교수)
 박진욱 (육군사관학교 명예교수)
 방승찬 (한국전자통신연구원 원장)
 배현철 (한국전자통신연구원 책임연구원)
 백준기 (중앙대학교 교수)
 변영재 (울산과학기술원 교수)
 서봉상 (울포랜드 상무)
 서종모 (서울대학교 교수)
 선우경 (서울대학교 교수)
 송문섭 (심텍 회장)
 송상현 (중앙대학교 교수)
 송진호 (연세대학교 교수)
 신요안 (숭실대학교 교수)
 심동규 (광운대학교 교수)
 안병구 (홍익대학교 교수)
 안현식 (동명대학교 교수)
 엄낙웅 (한국전자통신연구원 책임연구원)
 오윤호 (고려대학교 교수)
 우운택 (한국과학기술원 교수)

우 정 호 (비전넥스트 대표이사)
 유 성 철 (쌍용정보통신 상무)
 유 찬 세 (한국전자기술연구원 수석연구원)
 윤 명 국 (이화여자대학교 교수)
 윤 영 권 (삼성전자 마스터)
 이 강 윤 (성균관대학교 교수)
 이 규 복 (한국전자기술연구원 부원장)
 이 동 규 (카카오모빌리티 부사장)
 이 상 만 (고려대학교 교수)
 이 서 규 (한국팹리스산업협회 회장)
 이수 인 (텔레칩스 상무)
 이 우 주 (중앙대학교 교수)
 이 장 규 (텔레칩스 대표이사)
 이재 진 (송실대학교 교수)
 이정 석 (인하공업전문대학 교수)
 이 중 호 (서울대학교 교수)
 이 찬 수 (영남대학교 교수)
 이 천 희 (전주대학교 교수)
 이 태 동 (국제대학교 교수)
 이 형 민 (고려대학교 교수)
 이 희 덕 (충남대학교 교수)
 임 성 훈 (대구경북과학기술원 교수)
 장 길 진 (경북대학교 교수)
 전 국 진 (서울대학교 명예교수)
 전 선 익 (파이낸셜뉴스 부회장)
 전 재 욱 (성균관대학교 교수)
 정 방 철 (충남대학교 교수)
 정 승 원 (고려대학교 교수)
 정 원 영 (강원공업 본부장)
 정 이 품 (연세대학교 교수)
 정 준 (솔리드 대표이사)
 조 남 익 (서울대학교 교수)
 조 성 재 (이화여자대학교 교수)
 주 영 복 (한국기술교육대학교 교수)
 채 영 철 (연세대학교 교수)
 최 강 선 (한국기술교육대학교 교수)
 최 성 민 (해치텍 대표이사)
 최 영 돈 (삼성전자 마스터)
 최 우 영 (연세대학교 교수)
 최 재 혁 (서울대학교 교수)
 최 창 식 (DB하이텍 부회장)
 하 태 준 (광운대학교 교수)
 한 은 혜 (에스에스앤씨 대표이사)
 한 태 화 (연세대의료원 팀장)
 함 철 철 (삼성전자 마스터)
 홍 국 태 (X세미콘 연구위원)
 홍 승 홍 (인하대학교 명예교수)
 황 성 윤 (가천대학교 교수)
 황 인 태 (전남대학교 교수)

원 제 형 (도쿄엘렉트론코리아 대표이사)
 유 윤 섭 (한경대학교 교수)
 유 창 동 (한국과학기술원 교수)
 윤 상 훈 (한국전자기술연구원 책임연구원)
 윤 종 용 (삼성전자 비상임고문)
 이 광 엽 (서경대학교 교수)
 이 덕 우 (계명대학교 교수)
 이 명 재 (한국과학기술연구원 책임연구원)
 이 상 설 (한양대학교 명예교수)
 이성 준 (한양대학교 교수)
 이승 호 (한밭대학교 교수)
 이 윤 구 (광운대학교 교수)
 이재 관 (한국자동차연구원 소장)
 이재 훈 (서울대학교 명예교수)
 이정 원 (서울대학교 교수)
 이 주 연 (전주비전대학교 교수)
 이 창 우 (가톨릭대학교 교수)
 이 철 (동국대학교 교수)
 이 한 호 (인하대학교 교수)
 이 후 진 (한성대학교 교수)
 인 치 호 (세명대학교 교수)
 임 제 탁 (한양대학교 명예교수)
 장 성 진 (삼성전자 상근고문)
 전 동 석 (서울대학교 교수)
 전 세 영 (서울대학교 교수)
 전 홍 태 (중앙대학교 명예교수)
 정 범 진 (서울과학기술대학교 교수)
 정 영 모 (한성대학교 교수)
 정 은 성 (홍익대학교 교수)
 정 일 권 (한국전자통신연구원 본부장)
 정 해 준 (경희대학교 교수)
 조 도 현 (인하공업전문대학 교수)
 조 성 현 (한양대학교 교수)
 진 훈 (인양대학교 교수)
 채 주 형 (광운대학교 교수)
 최 광 성 (한국전자통신연구원 실장)
 최 수 일 (전남대학교 교수)
 최 용 수 (신한대학교 교수)
 최 욱 (인천대학교 교수)
 최 정 욱 (한양대학교 교수)
 최 천 원 (단국대학교 교수)
 한 동 국 (국민대학교 교수)
 한 재 호 (고려대학교 교수)
 한 태 회 (성균관대학교 교수)
 허 재 두 (한국전자통신연구원 연구전문위원)
 홍 대 식 (연세대학교 교수)
 홍 인 기 (경희대학교 교수)
 황 원 준 (아주대학교 교수)
 황 진 영 (한국항공대학교 교수)

유 명 식 (송실대학교 교수)
 유 재 준 (울산과학기술원 교수)
 유 창 식 (삼성전자 부사장)
 윤 석 현 (단국대학교 교수)
 윤 종 윤 (파두 사장)
 이 규 대 (공주대학교 교수)
 이 덕 진 (전북대학교 교수)
 이 문 기 (연세대학교 명예교수)
 이상 훈 (웨이브피아 대표이사)
 이수 민 (한국센서연구소 대표이사)
 이영택 (ASML 전무)
 이 인 규 (고려대학교 교수)
 이재 규 (삼성전자 마스터)
 이재 훈 (유정시스템 대표이사)
 이 중 호 (서울대학교 교수)
 이 진 구 (동국대학교 석좌교수)
 이재 은 (한양대학교 교수)
 이충용 (연세대학교 교수)
 이 혁 재 (서울대학교 교수)
 이 흥 노 (광주과학기술원 교수)
 임 대 영 (한국산업기술시험원 박사)
 임 혜 숙 (이화여자대학교 교수)
 장 익 준 (경희대학교 교수)
 전 병 우 (성균관대학교 교수)
 전 영 현 (삼성전자 부회장)
 정 민 채 (세종대학교 교수)
 정 성 엽 (고려대학교 교수)
 정 완 영 (KAIST 교수)
 정 은 승 (삼성전자 고문)
 정 정 화 (한양대학교 석좌교수)
 제 민 규 (한국과학기술원 교수)
 조 문 규 (한국교통대학교 교수)
 조 영 민 (SkyMirr CEO)
 채 관 엽 (삼성전자 마스터)
 천 경 준 (씨젠 회장)
 최 병 호 (한국전자기술연구원 연구소장)
 최 영 규 (인하대학교 교수)
 최 우 영 (서울대학교 교수)
 최 재 용 (가천대학교 교수)
 최 중 호 (서울시립대학교 교수)
 추 민 성 (한양대학교 교수)
 한 동 석 (경북대학교 교수)
 한 진 호 (한국전자통신연구원 책임연구원)
 함 범 섭 (연세대학교 교수)
 허 준 (고려대학교 교수)
 홍 성 완 (서강대학교 교수)
 홍 철 호 (중앙대학교 교수)
 황 인 철 (강원대학교 교수)
 황 태 호 (한국전자기술연구원 본부장)

사무국 직원 명단

송기원 국장 - 산학연관 협력, 신규 사업, 저문/서울IT포럼, 지부, 인사, 규정, 회장단 관련, 총회 등 사무국 총괄
 이안소 부장 - 학회지, 추계학술대회, 이사회/평의원회, 총무업무(선거, 공문, 임원관련, 송년회, 포상 및 Wiset 등), 산업전자소사이터티
 배지영 부장 - 국제학술대회 총괄(ITC-CSCC, ICBC, ICCE-Asia), 하계종합학술대회, 신기술(총괄), 시스템 및 제어소사이터티
 배기동 부장 - 사업행사(기술워크숍 등), 국문논문, 표준화, 용역업무, 인공지능신호처리소사이터티
 변은정 부장 - 본회/소사이터티/연구회 재무, 회원관리(개인회원 및 특별회원), 기획, 통신소사이터티
 이소진 서기 - 국제학술대회 담당(ITC-CSCC, ICBC, ICCE-Asia), 외국 기관과 국제협력(Joint Award 등), JSTS 및 SPC 발간, 컴퓨터소사이터티
 김예빈 서기 - 정보화(홈페이지 관리 및 디지털 업무지원, 장비관리), 교육, 홍보, 신기술(담당), 기타 지원업무, 반도체소사이터티

학회소식

2024년도 정기총회 및 추계학술대회

11월 22일(금)~23일(토) 양일간에 걸쳐 소사이더티 주관으로 하이원리조트 컨벤션타워(강원도 정선)에서 학계, 산업계 및 연구계에서 연구한 총 370여편의 학술논문을 발표하는 추계학술대회가 개최되었다. 그리고 한 해를 마무리하는 정기총회와 아울러 금년도 학회를 위해 수고하거나 탁월한 업적을 이룬 회원에게 학회상을 수여하였다. 한편 이번 총회에서는 2024년도 결산, 2025년도 사업계획 및 예산(안) 승인, 2025년도 임원 선출 보고 등의 안건들이 원안대로 승인되었다.



정기총회 - 이충용 학회장 인사말



추계학술대회 기조강연 - 박명재 부사장(SK하이닉스)



단체사진

ESD/EOS & Latchup 워크숍 2024

반도체소사이어티(운영위원장 : 김한구 대표(EESOP))에서는 “ESD/EOS & Latchup 워크숍 2024”을 11월 14일(목)부터 15일(금)까지 한양대학교 HIT(한양종합기술연구원) 2층 대강의실에서 개최 하였다. IC가 등장하며 등장한 ESD/EOS Issue는 IC 제조 단계부터 조립 및 테스트를 거쳐 Module과 Product 제조 단계에 이르기까지 IC와 관련된 전분야에서 발생할 수 있는 신뢰성 문제 중의 하나라고 볼 수 있다.

Foundry 업체의 Leading Company인 TSMC의 경우, 2015년부터는 공식적으로 고객들에게 ESD Cell을 제공하지 않는 것을 원칙으로 하고 있으며, Fabless 업체에서 IP 및 IC를 설계함에 있어서 ESD/EOS 문제를 해결함에 있어서 많은 부담이 되고 있는 상황이다.

IDM 업체들은 그들이 직접 공정을 개발하는 한편, IP와 Product도 개발하고 있기 때문에 공정개발 단계부터 Silicon-Level에서 ESD/EOS에 대한 많은 평가를 거쳐 IC-Level에서의 ESD/EOS Solution을 도출하고 있고, IP와 Product 및 고객사(Module/Product)에서 발생하는 System-Level ESD/EOS에 대해서도 많은 경험을 가지고 있는 ESD Engineer들이 있다. 그러나, 대부분의 Fabless 업체들은 1~2명 정도의 ESD Engineer를 보유하고 있기 때문에 경험과 지식이 부족한 상황이며, 제품 양산 단계에서 ESD/EOS & Latchup 등의 Issue가 발생하는 경우에는 원인분석과 개선안 도출에도 어려움이 따르고 있고, 양산 지연의 문제가 발생하는 경우가 많을 수밖에 없다.

따라서 본 워크숍은 ESD/EOS & Latchup에 대한 교육을 통해 IC와 관련된 다양한 업무에 종사하고 있는 Engineer들에 대한 지식을 업그레이드하는 것을 목적으로 하였다.



반도체소사이어티 장성진 회장 축하



워크숍 전경

제6회 Memory-Centric 컴퓨팅의 미래와 도전 워크숍

반도체소사이어티(운영위원장 : 노원우 교수(연세대), 장익준 교수(경희대))에서는 11월 15일(금)

경희대학교 국제캠퍼스 피스홀에서 “제6회 Memory-Centric 컴퓨팅의 미래와 도전 워크숍”을 개최하였다. 기조 강연으로 Duke University의 Hai Li 교수가 엡지 컴퓨팅 환경에서의 AI 모델 경량화와 하드웨어 최적화 방안을 제시하고 효율적인 연산을 위한 하드웨어 인식 기반의 최적화 기법을 탐색하였다.

오전 첫번째 경량화 세션에서는 네이버의 권세중 리더를 모시고 하이퍼스케일 AI 시대에 총소유비용(TCO) 절감과 소프트웨어 에코시스템의 변화를 다루며, AI 컴퓨팅의 경제성과 확장성에 대한 심층적인 분석을 제시 하였고, 김형준 대표(스퀴즈비츠)는 대형 언어 모델(LLM)의 효율적인 서비스 방안을 주제로 vLLM과 TensorRT-LLM을 비교하며, 비용 효율성을 극대화할 수 있는 구현 기술에 대한 심도 있는 발표를 하였다.

오후 첫 세션으로는 “DRAM 세션”이란 제목으로 데이터센터부터 엡지 디바이스까지 활용될 수 있는 메모리 기반 AI 컴퓨팅 솔루션을 중심으로 발표가 이어졌다.

SK하이닉스의 임의철 Fellow는 PIM(Processing in Memory) 기술을 통한 LLM의 효율적인 컴퓨팅 방법을 제시하며, 데이터 중심 연산 구조가 메모리와 프로세서 간의 병목을 어떻게 해결할 수 있는지를 다루었다. 이어서 삼성전자의 손교민 마스터는 AI 시대에 최적화된 메모리 솔루션을 소개하며, 대규모 데이터 연산에 특화된 차세대 메모리 기술을 집중 조명하였고, 마지막으로 메티스엑스의 김진영 대표는 CXL 기술을 기반으로 AI 데이터센터의 미래를 전망하며, 대규모 컴퓨팅에서의 데이터 이동성과 처리 속도 향상에 대해 설명하였다.

마지막 선도 연구자 세션에서는 KAIST의 제민규 교수가 뉴럴 레코

딩을 위한 CIM 응용 사례를 통해 최신 연구 동향을 공유하였고, 이어
서 연세대학교의 노원우 교수는 지난 7년간의 본 워크숍 구성 및 참
여 경험을 바탕으로 메모리 중심 컴퓨팅의 미래와 도전에 대해 심도
깊은 회고와 통찰을 나누었다.

이번 세미나는 다양한 산업 및 학계의 연구 성과와 향후 과제를 포
괄적으로 다루며, AI와 메모리 기술의 융합을 통해 만들어갈 미래 비
전을 모색하는 기회가 제공 되었고, 약 120명이 참석하였다.



워크숍 전경



참석자 단체 사진



학회일지

www.theieie.org

THE INSTITUTE OF
ELECTRONICS AND INFORMATION
ENGINEERS

(2024년 11월 13일 ~ 12월 13일)

1. 행사 개최

구분	행사명	기간	장소
반도체소사이어티	ESD/EOS & Latchup 워크숍 2024	11.14-15	한양대학교 HIT
반도체소사이어티	제 6회 Memory-Centric 컴퓨팅의 미래와 도전 워크숍 (온라인병행)	11.15	경희대학교 국제캠퍼스
학술위원회	2024년도 추계학술대회	11.22~23	하이원리조트 컨벤션타워

AI 반도체를 위한 하드웨어-소프트웨어 공동 경량화 기술



류 성 주 편집위원
(서강대학교)

최근 스마트폰, 자동차, 웨어러블 디바이스를 비롯한 엣지 디바이스는 물론이고, 대규모 데이터를 처리하는 데이터센터에 이르기까지 인공지능의 활용 범위가 급격히 확대되고 있다. 이와 더불어 사용자들은 더욱 빠르고 정확하며 실시간으로 반응하는 인공지능 서비스를 요구하

고 있으며, 이러한 요구를 충족시키기 위해서는 고성능 인공지능 반도체의 개발이 필수적이다. 특히, 제한된 자원을 효율적으로 활용하면서도 높은 성능을 발휘할 수 있는 경량화 기술은 인공지능 반도체 분야의 핵심적인 과제로 떠오르고 있다.

본 특집호에서는 이러한 시대적 요구에 발맞춰 인공지능 반도체를 위한 소프트웨어 및 하드웨어 공동 경량화 기술을 심층적으로 다룬다. 다섯 기관 전문가들이 각자의 연구 결과와 통찰력을 바탕으로 인공지능반도체 기술의 현주소와 미래 전망을 제시한다.

첫째, “AI 반도체를 위한 SW/HW 연구동향 (박종선 외)”에서는 다양한 인공지능 애플리케이션에 널리 활용되는 Transformer 및 Convolutional neural network 모델들을 효율적으로 하드웨어에서 가속연산하기 위한 Quantization 및 Pruning 등의 경량화 기법을 소개하며 이를 위한 하드

웨어 아키텍처 설계 방법에 대해 소개한다. 둘째, “영상처리를 위한 경량화 뉴럴프로세서 설계 (류성주)”에서는 영상 및 동영상 처리를 저전력으로 진행하기 위한 소프트웨어 경량화 방법 및 이를 위한 하드웨어 설계 방법을 소개한다. 셋째, “AI 반도체를 위한 TinyML 기술 동향 (이성수 외)”에서는 엣지디바이스의 저전력 디버닝 연산을 위한 다양한 TinyML 응용 방법을 소개하고 이를 위한 최신 기술 동향을 설명한다. 넷째, “플래시 메모리 기반의 인공지능 반도체 (이성태)”에서는 플래시 메모리에서 학습가능한 인공지능 반도체 설계 방법에 대해 설명한다. 다섯째, “반도체 설계를 위한 AI/ML 기술 (최구명)”에서는 인공지능 반도체 외에도 반도체 설계를 위한 인공지능 기술을 소개하고 이를 통해 EDA 툴의 성능을 높이는 연구 결과들을 소개하였다.

이렇게 연말의 바쁜 일정 중에 본 특집호를 위해 소중한 칼럼을 보내주신 집필진에 감사의 인사를 전하며, 본 특집호가 AI 반도체 소프트웨어-하드웨어 공동 최적화를 위한 여러 분야 전문가들의 교류와 협력을 위한 새로운 계기가 되어 우리나라 IT융합산업의 발전과 경쟁력 강화에 기여할 수 있기를 기원한다.

AI 반도체를 위한 SW/HW 연구 동향

I. 서론

최근 AI 기술의 비약적인 발전은 우리의 일상에 큰 변화를 가져오고 있다. Chat-GPT와 같은 대화형 인공지능이 자연스럽게 사람과 상호 작용하며, 자율 주행 자동차가 스스로 주행 환경을 인식하고 의사 결정을 내리는 것이 대표적인 예다. AI 기술의 성장은 2010년대부터 본격적으로 가속화되었는데, 그 배경에는 컴퓨팅 하드웨어의 발전이 있다. 특히, 그래픽 처리를 위해 개발된 GPU(Graphic Processing Unit)가 핵심적인 역할을 했다. 병렬 연산에 최적화된 구조인 GPU를 AI 모델 연산에 활용함으로써 반복적인 대규모 데이터의 연산을 효율적으로 처리할 수 있었다. 이러한 장점을 바탕으로 GPU는 AI 분야에 광범위하게 활용되며 기술 발전을 가속화시켰다.

2020년대에 들어서며 대규모 언어 모델(LLM, Large-Language Model)과 같은 복잡한 AI 모델들이 등장했다. 최신 AI 모델의 파라미터 수는 2010년대에 비해 수십 배에서 수천 배로 증가했으며, 이에 따라 연산량과 메모리 접근 요구량도 기하급수적으로 늘어났다. 이러한 복잡성의 증가는 데이터 처리 시간 지연과 막대한 에너지 소모를 초래했다. 2023년에 Open AI에서 발표한 대표적인 대규모 언어 모델인 GPT-4는 훈련하는데 평균 100일이 걸리며, 62,319MWh의 에너지가 소모된다. 2020년에 발표된 GPT-3가 훈련하는데 평균 34일이 걸리며, 1,287MWh의 에너지가 소모되는 것에 비교하면 훈련하는데 걸리는 시간은 약 3배, 소모되는 에너지는 약 50배 증가하는 것이다. 이러한 문제를 해결하기 위해서는 AI 모델에 최적화된 AI 반도체의 개발이 필수적이다.

AI 반도체 연구는 소프트웨어/하드웨어 최적화 방법이 상호협력하며 활발히 진행되고 있다. 소프트웨어 최적화 방법은 하드웨어 구조에



박재범
고려대학교



안민영
고려대학교



박종선
고려대학교

맞게 AI 모델을 최적화하는 방법으로, 연산 방법 또는 모델 구조를 변경하여 AI 모델의 복잡도를 줄인다. 하드웨어 최적화 방법은 AI 모델의 특성을 반영하여 데이터 처리 방식이나 물리적 구조를 변경하는 것에 주안을 두고 있다. 이러한 최적화 방법들이 상호적으로 결합되어 더욱 효율적인 AI 반도체를 설계할 수 있다. 본 원고에서는 최근 대표적으로 사용되는 AI 모델을 간략히 소개하고, AI 반도체를 위한 소프트웨어 및 하드웨어 연구 동향을 살펴보고자 한다.

II. Transformer & CNN 배경 설명

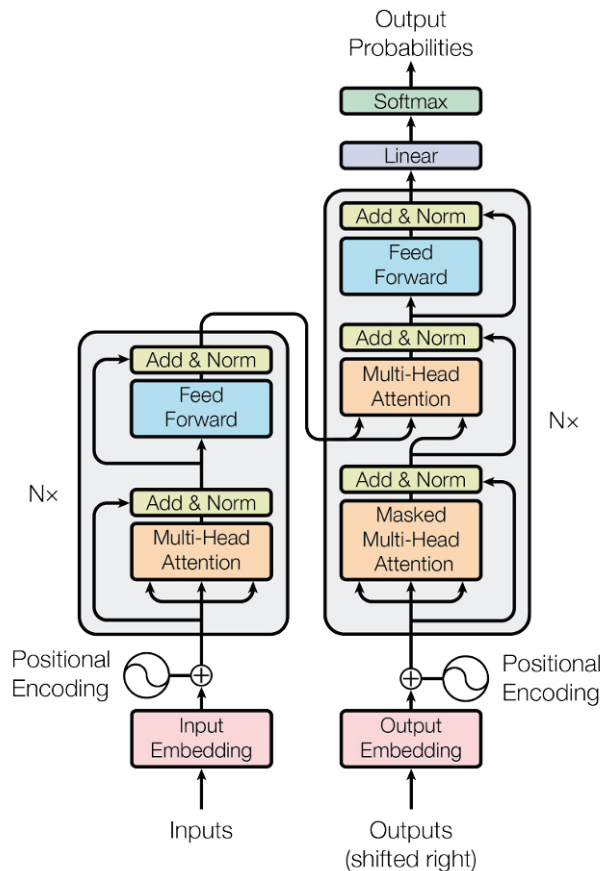
1절 서론에서 언급했듯이, 대표적인 AI 응용 분야에는 대화형 인공지능, 자율주행 기술 등이 있다. Chat-GPT 등의 대화형 인공지능에 사용되는 AI 모델은 인간의 언어를 이해하고 생성하는 Transformer 기반 Large-Language Model (LLM)이며, 자율주행 기술에서 활발하게 사용되는 AI 모델은 시각적 데이터 처리에 적합한 Convolutional Neural Network (CNN)이다. 이 절에서는 최근 대표적인 AI 모델인 Transformer와 CNN의 기본 구조와 주요 계층에 대한 연산 방법을 소개한다.

1. Transformer 모델

LLM의 성능 향상에는 Transformer 구조가 중요한 역할을 하였으며, 이는 [1]에서 처음 제안되었다. Transformer의 주요 혁신 중 하나는 Self-Attention 메커니즘이다. Self-Attention은 입력 데이터의 각 요소가 다른 요소들과의 상관관계를 분석하여 중요도가 높은 요소에 더 큰 가중치를 부여하는 방식이다. 이를 통해 문장 내 각 단어가 다른 모든 단어와 동시에 상호작용할 수 있으며, 이 처리 방식을 통해 모델의 정보 처리 능력이 크게 향상되었다.

1) Transformer의 기본 구조

〈그림 1〉에서 보듯이, Transformer는 크게 세 가지 주요 구성 요소로 나눌 수 있다. 먼저, 임베딩(Embedding) 단계에서는 입력된 토큰(단어)에 의미 정



〈그림 1〉 Transformer 모델의 기본 구조 ^[1]

보를 포함한 고차원 벡터를 할당한다. 다음으로, 멀티헤드 어텐션(Multi-Head Attention)은 앞서 설명한 Self-Attention 메커니즘을 여러 개 병렬로 처리하여 모델의 표현력을 크게 확장시킨다. 마지막으로, 피드포워드 네트워크(Feed-Forward Network)는 비선형 변환을 통해 모델이 복잡한 관계를 학습할 수 있게 한다. 이러한 구성 요소들이 결합되어 Transformer는 높은 수준의 문맥 이해와 복잡한 의미 관계를 효과적으로 학습할 수 있다.

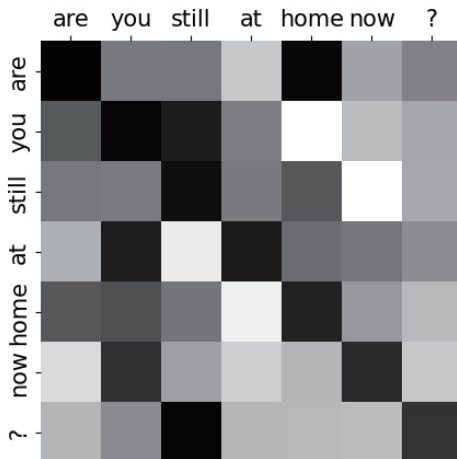
2) Multi-Head Attention의 연산 방법

Transformer의 Self-Attention 연산 과정은 다음과 같은 단계로 진행된다. 먼저, 토큰의 임베딩 과정을 거친 후 가중치와 곱하여 Query(Q), Key(K), Value(V)라는 세 가지 행렬을 생성하게 되는데, 이 행렬들은 각 토큰에 대한 고유한 특성을 담고 있다.

〈그림 2〉은 Q, K, V를 사용해 Self-Attention의 연산

$$\text{Softmax}\left(\frac{\mathbf{Q} \times \mathbf{K}^T}{\sqrt{d_k}}\right) \mathbf{V} = \mathbf{Z}$$

〈그림 2〉 Self-Attention의 연산 과정



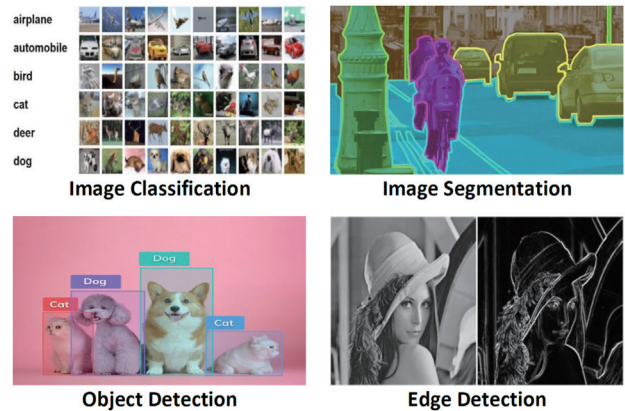
〈그림 3〉 Self-Attention의 히트맵

과정을 보여준다. 먼저, 앞서 구한 Q와 K 간의 유사도를 계산하기 위해 두 벡터의 점곱(Dot Product)을 수행한 후, 값이 지나치게 커지지 않도록 적절한 값으로 나누어 정규화한다. 이후 Softmax 함수를 적용해 유사도를 가중치로 변환한다. 마지막으로 이 가중치를 Value 행렬에 곱해 Self-Attention의 최종 출력을 얻는다. 이 출력은 입력된 문장의 단어들이 서로 어떠한 관계를 형성하는지를 반영한 새로운 표현이다. 이 과정을 여러 개의 헤드로 병렬 처리한 것이 Multi-Head Attention으로, 각 헤드는 서로 다른 관계를 학습하여 모델의 표현력을 더욱 풍부하게 만든다.

〈그림 3〉은 Self-Attention 연산 결과를 히트맵(Heatmap)으로 시각화한 것으로, 각 단어가 다른 단어들과 얼마나 연관되어 있는지를 유사성 값으로 시각적으로 확인할 수 있다. 이를 통해 Transformer는 Self-Attention을 활용하여 문맥 속에서 단어들 간의 관계를 명확하게 파악할 수 있다.

2. CNN 모델

CNN은 인접한 데이터가 먼 데이터보다 연관성이 높을 것이라고 가정하고 인접한 거리의 데이터들 사이에서 특



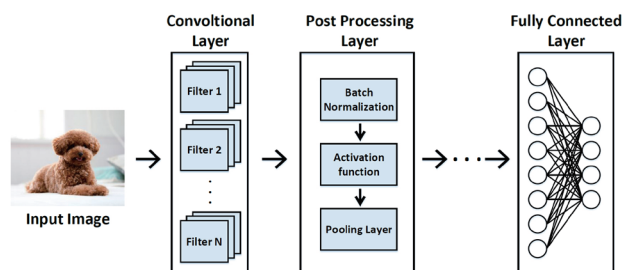
〈그림 4〉 CNN의 다양한 활용 분야

징을 추출하여 연관성을 찾아내는 컨볼루션 연산을 활용한다. 실제로 이미지 데이터는 가까운 데이터들 사이에 연관성을 가지기 때문에, CNN의 연산 특징은 이미지 데이터에 적합한 AI 모델이다. 따라서 CNN은 〈그림 4〉와 같이 이미지 분류, 객체 검출, 이미지 분할 등의 컴퓨터 비전 분야에서 뛰어난 성능을 보이고 있다.

1) CNN 모델의 기본 구조

CNN은 이미지 분류, 객체 검출, 이미지 분할 등의 활용 분야에 따라 모델의 깊이, 계층 구조, 특징 추출 방식 등이 달라지게 된다. 〈그림 5〉는 CNN의 대표적인 활용 분야인 이미지 분류의 대표적인 모델의 구조를 나타내었다.

〈그림 5〉는 CNN의 기본 구조이며, 크게 컨볼루션 계층(Convolution Layer), 후처리 계층(Post-Processing Layer), 완전 연결 계층(Fully-Connected Layer)으로 이루어져 있다. 컨볼루션 계층은 CNN 모델의 핵심 계층으로, 이미지를 작은 조각으로 나누어 특징을 뽑아내

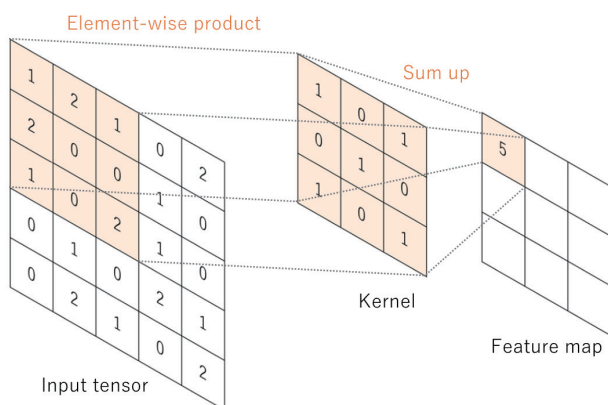


〈그림 5〉 CNN 모델의 기본 구조

는 역할을 한다. 후처리 계층은 배치 정규화 계층(Batch Normalization Layer), 활성화 함수 계층(Activation Function Layer), 풀링 계층(Pooling Layer) 등으로 구성되어 있다. 배치 정규화 계층은 입력 데이터의 분포를 정규화해서 학습이 더 안정적이고 빠르게 진행될 수 있도록 돕는 역할을 하며, 활성화 함수 계층은 모델에 비선형성을 추가하여 복잡한 패턴을 학습할 수 있게 해준다. 풀링 계층은 컨볼루션 계층 뒤에 위치하며, 이전 계층의 출력에서 중요한 데이터를 추출하여 특징의 공간적 크기를 축소한다. CNN 모델의 종류에 따라 컨볼루션 계층과 후처리 계층은 다양한 구조로 반복될 수 있다. CNN 모델의 마지막에 위치한 완전 연결 계층(Fully-Connected Layer)은 컨볼루션 계층에서 추출한 특징을 바탕으로 분류나 예측 등의 최종적인 결과를 만들어 낸다.

2) 컨볼루션 계층의 연산 방법

CNN 모델에서 가장 큰 연산량을 차지하는 계층은 컨볼루션 계층이다. 따라서 컨볼루션 계층에서 사용되는 주요 연산인 컨볼루션 연산의 방법과 특징을 소개한다. 컨볼루션 연산은 필터와 입력 데이터의 연산으로 이루어진다. 필터는 매트릭스 형태로 정의되며, 보통 3×3 , 5×5 , 7×7 과 같은 크기를 가진다. 필터는 이미지를 한 번에 일정 부분씩 훑으며, 해당 영역의 정보를 분석하게 된다. 필터는 이미지의 한 부분과 겹쳐지며, 이 영역의 각 픽셀 값과 필터의 값을 곱한 후 그 결과를 모두 더하는 연산을 수행한다. <그림 6>은 가중치(Weight) 크기가 3×3 일 때



<그림 6> 2D 컨볼루션 연산 예시^[2]

2D 입력 데이터와의 컨볼루션 연산을 나타낸다. 필터가 이미지를 한 번 훑은 결과는 하나의 숫자로 표현되며, 필터가 이미지 전체를 이동하면서 같은 연산을 반복해 출력 데이터를 생성한다. 이처럼, 컨볼루션 계층에서는 단순한 곱셈 및 덧셈 연산이 여러 번 반복된다.

III. AI 반도체를 위한 SW 연구 동향

AI 모델의 규모가 커짐에 따라 연산량과 메모리 사용량이 급격히 증가하면서, 성능을 유지하면서도 효율성을 높이기 위한 경량화 기법들이 필수적으로 요구되고 있다. 이를 해결하기 위해 다양한 기술이 활발히 연구되고 있으며, 대표적인 기법으로 가지치기(Pruning), 양자화(Quantization), 지식 증류(Knowledge Distillation)가 있다.

가지치기는 중요하지 않은 파라미터를 제거하여 모델의 크기와 연산량을 줄이고, 양자화는 가중치(Weight)와 활성화(Activation)를 적은 비트로 변환하여 메모리와 에너지 효율을 극대화하는 방법이다. 지식 증류는 대규모 모델의 성능을 작은 모델에 전달해 경량화하면서도 성능 저하를 최소화한다.

본론에서는 이러한 경량화 기법 중 대표적인 양자화와 가지치기의 원리와 종류, 그리고 각 기법의 장단점에 대해 구체적으로 살펴보고자 한다.

1. AI 모델 경량화를 위한 양자화 기법

양자화는 신경망의 가중치와 활성화를 더 적은 비트 수로 변환하여 모델의 성능과 효율성을 높이는 기술이다. 이를 통해 모델 크기를 줄이고, 계산 속도를 향상시키며, 메모리 사용량을 줄여 모델을 더욱 효율적으로 활용할 수 있다.

<그림 7>에서 보듯이, 가속기에서는 연산보다 Off-Chip 메모리(DRAM)에서 데이터를 불러오는 데 더 많은 에너지가 소모된다. 따라서, 양자화를 통해 데이터를 적은 비트로 표현하면 모델 크기를 줄일 뿐만 아니라, Off-Chip 메모리 접근에 따른 에너지 소모도 크게 절감할 수 있다. 양자화 과정에서 일부 정보 손실이 발생할 수 있지

Operation	Energy [pJ]	Relative Cost
32 bit int ADD	0.1	1
32 bit float ADD	0.9	9
32 bit int MULT	3.1	31
32 bit float MULT	3.7	37
32 bit 32KB SRAM	5	50
32 bit DRAM	640	6400

〈그림 7〉 연산 별 에너지 [3]

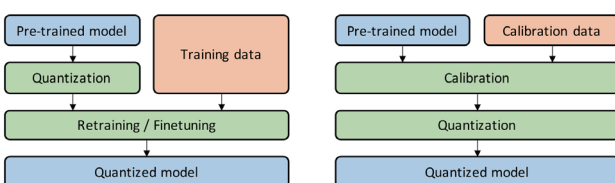
만, 모델이 커짐에 따라 양자화는 필수적인 기술로 자리 잡고 있다.

1) 양자화 방법에 따른 분류

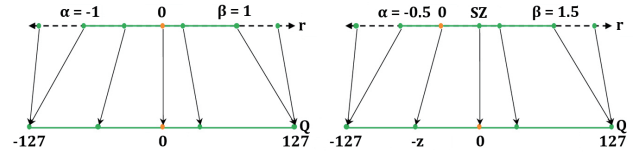
양자화는 〈그림 8〉에서 나타난 것처럼 방법에 따라 양자화 인식 훈련(Quantization-Aware Training)과 사후 훈련 양자화(Post-Training Quantization)으로 구분된다.

양자화 인식 훈련은 모델 학습 단계에서부터 양자화를 반영하여 진행되는 방식이다. 이 방식은 양자화 후에도 모델의 성능 저하를 최소화할 수 있다는 장점이 있지만, 추가적인 학습 과정이 필요하다는 단점이 존재한다. 양자화 인식 훈련을 이용한 소프트웨어적 기법의 종류는 대표적으로 [5,6] 등이 있다.

사후 훈련 양자화는 이미 학습이 완료된 모델을 양자화하는 방식이다. 이때 활성화의 양자화 방법에 따라 정적 양자화(Static Quantization)와 동적 양자화(Dynamic Quantization)로 나뉜다. 정적 양자화는 가중치와 활성화를 모두 미리 양자화하는 방식으로, 메모리 효율성을 극대화할 수 있다. 반면, 양자화 가중치는 가중치는 미리 양자화하되 활성화는 추론 중에 동적으로 양자화함으로써 성능을 높일 수 있다. 사후 훈련 양자화를 이용한 소프트웨어적 기법의 종류는 대표적으로 [7,8] 등이 있다.



〈그림 8〉 (a) QAT (b) PTQ [4]



〈그림 9〉 (a) 대칭 양자화 (b) 비대칭 양자화 [4]

2) Formula Definition에 따른 분류

양자화는 〈그림 9〉에서 나타난 것처럼 Formula Definition에 따라 대칭 양자화(Symmetric Quantization)와 비대칭 양자화(Asymmetric Quantization)로 구분된다.

대칭 양자화는 입력 데이터의 범위를 양의 최대값과 음의 최대값으로 동일하게 설정한 후, 그 범위 내에서 데이터를 양자화하는 방식이다. 이 방식은 구현이 간단하고 효율적이지만, 데이터의 분포가 균일하지 않을 경우 모델의 성능이 감소할 수 있는 단점이 있다. 대칭 양자화를 이용한 소프트웨어적 기법의 종류는 대표적으로 [9,10] 등이 있다.

비대칭 양자화는 입력 데이터의 최소값과 최대값을 각각 다른 값으로 지정하여 그 범위에 맞춰 데이터를 양자화하는 방식이다. 이를 통해 데이터의 비대칭적인 분포를 더욱 정교하게 반영할 수 있으며, 모델의 성능 저하를 최소화하는 데 유리하다. 이 때문에 비대칭 양자화는 특히 높은 성능이 요구되는 응용 분야에서 주로 사용된다. 비대칭 양자화를 이용한 소프트웨어적 기법의 종류는 대표적으로 [11,12] 등이 있다.

2. AI 모델 경량화를 위한 가지치기 기법

가지치기란 중요하지 않은 가중치 또는 입력 데이터를 0으로 치환하여 AI 모델의 전체 파라미터 수를 줄이는 기법이다. AI 모델은 수많은 가중치와 같은 파라미터로 구성되어 있으며 일부 파라미터는 모델 성능에 큰 영향을 미치지 않기 때문에 불필요한 파라미터를 제거하여 AI 모델의 크기와 연산량을 줄일 수 있다.

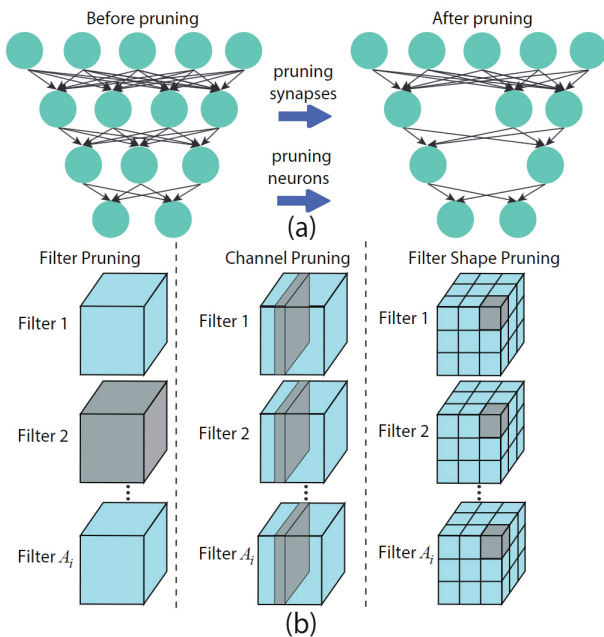
1) 가지치기 대상에 따른 분류

〈그림 10〉과 같이, 가지치기는 가지치기를 하는 대상에 따라 구조적 가지치기(Structured Pruning)와 비구조

적 가지치기(Unstructured Pruning)로 구분할 수 있다.

비구조적 가지치기는 네트워크의 임의의 개별 가중치를 제거하는 방식으로, 모델 내에서 중요도가 낮은 가중치를 선택적으로 제거하여 가중치의 수를 줄이는 데 중점을 둔다. 이 방식의 가장 큰 장점은 제거할 가중치를 자유롭게 선택할 수 있어 높은 비율의 가중치 제거가 가능하다는 것이다. 하지만, 비구조적 가지치기는 가중치가 불규칙하게 제거되기 때문에, 남은 가중치의 위치 정보를 저장해야 하는 문제가 있다. 이는 추가적인 인덱스 저장 공간을 요구하게 되어, 오히려 전체 저장 공간이 커지거나 하드웨어 구현이 어려워지는 단점이 있다. 비구조적 가지치기를 이용한 소프트웨어적 기법의 종류는 대표적으로 [13,14] 등이 있다.

구조적 가지치기는 네트워크 구조를 고려하여 일정한 패턴에 따라 가중치나 뉴런을 제거하는 방식이다. 대표적으로 필터 가지치기(Filter Pruning)와 채널 가지치기(Channel Pruning), 그리고 필터 형태 가지치기(Filter Shape Pruning)가 있다. <그림 10> (b) 참고) 필터 가지치기는 AI 모델의 필터 전체를 제거하는 방식으로, 특정 계층에서 전체 필터를 제거함으로써 모델의 복잡성을 줄인다. 채널 가지치기는 여러 연속된 채널을 제거하는 방식



<그림 10> (a) 비구조적 가지치기 (b) 구조적 가지치기의 3가지 종류 [17]

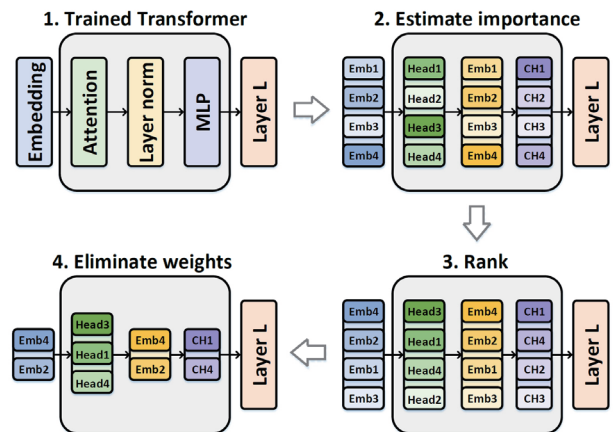
이며, 필터 형태 가지치기는 모든 필터에서 동일한 위치의 가중치를 제거하는 방식이다. 구조적 가지치기는 규칙적인 매트릭스 형태를 유지하기 때문에, 가중치의 위치 정보(인덱스)를 저장할 필요가 없다. 따라서, 구조적 가지치기는 하드웨어 구현에 더 적합하고 병렬 처리가 용이한 장점을 가지고 있다. 구조적 가지치기를 이용한 소프트웨어적 기법의 종류는 대표적으로 [15,16] 등이 있다.

2) 가중치 제거 기준에 따른 분류

가지치기는 가중치를 제거하는 기준에 따라 크기 기반 가지치기(Magnitude-based Pruning)와 중요도 기반 가지치기(Importance-based Pruning)로 분류된다.

크기 기반 가지치기는 가중치의 크기(절댓값)을 기준으로 작은 가중치를 제거하는 방식이다. 즉, 가중치가 작을수록 성능에 덜 중요한 것으로 간주한다. 계산량이 적으며 간단한 조건으로 가중치를 제거하기 때문에 하드웨어 구현이 쉽다. 크기 기반 가지치기를 이용한 소프트웨어적 기법의 종류는 대표적으로 [18],[19] 등이 있다.

중요도 기반 가지치기는 가중치가 모델 성능에 미치는 영향을 평가하여 성능 저하가 적은 가중치를 제거하는 방식이다. <그림 11>은 Transformer에 중요도 기반 가지치기를 적용했을 때의 과정을 나타낸 그림이다. 크기 기반 가지치기와는 달리 성능 분석이 필요하기 때문에 하드웨어에서 구현 시 연산량과 메모리 사용량이 커질 수 있지만, 성능 저하를 최소화할 수 있다는 장점이 있다. 중요



<그림 11> Transformer 중요도 기반 가지치기 [22]

도 기반 가지치기를 이용한 소프트웨어적 기법의 종류는 대표적으로 [20,21] 등이 있다.

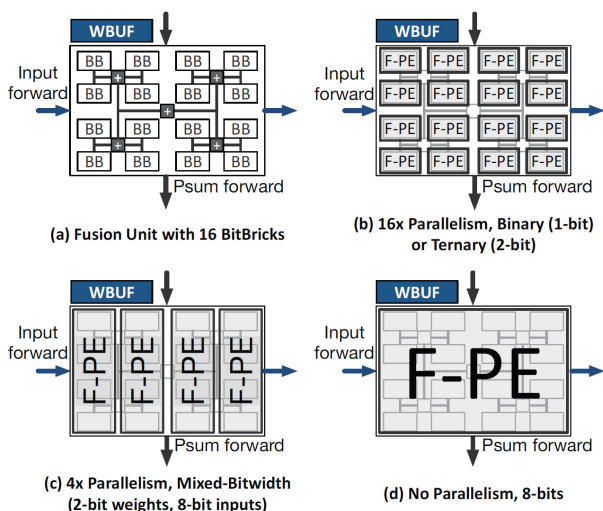
IV. AI 반도체를 위한 HW 연구 동향

양자화 및 가지치기와 같은 AI 반도체를 위한 소프트웨어적 기법은 불규칙적인 연산 및 데이터 접근 등으로 인해 기존의 범용 CPU/GPU와 같은 하드웨어에는 효율적으로 적용할 수 없기 때문에 소프트웨어적 기법을 고려한 하드웨어 설계 연구가 필요하다. 따라서 이 절에서는 가지치기 및 양자화 기법에 의해 경량화된 모델을 하드웨어에서 효율적으로 동작시키는 가속기 연구 동향을 소개한다.

1. 혼합-정밀도 양자화 (Mixed-Precision Quantization)를 지원하는 가속기 아키텍처

기존의 CPU/GPU와 같은 하드웨어에서는 고정된 비트폭을 사용한다. 그러나 AI 모델의 각 계층마다 결과에 미치는 중요도가 다를 수 있다. 따라서 중요한 계층에 더 높은 정밀도를 적용하는 혼합-정밀도 양자화를 지원하는 가속기(Mixed-Precision Quantization Accelerator)가 필요하며, 이를 통해 연산 자원을 효율적으로 사용하면서도 모델 성능 저하를 최소화할 수 있다.

〈그림 12〉는 혼합-정밀도 양자화를 지원하는 대표적인 디지털 가속기 아키텍처로, 비트 레벨에서 동적으로



〈그림 12〉 Fused PE 내 Bitbrick들의 동적인 결합 [23]

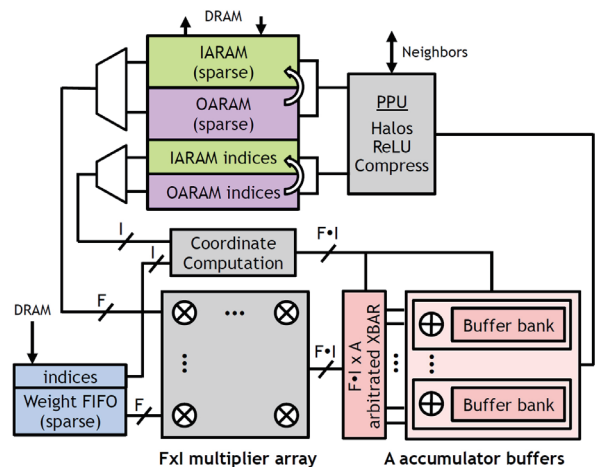
조합 가능한 구조를 가진다. [23]는 AI 모델에서 계층별로 필요한 비트폭에 맞추어 가속기 구조를 조정할 수 있다. 아키텍처의 기본 구성 요소인 'Bitbrick'은 2비트 단위로 곱셈 및 덧셈 연산을 수행한다. 이 Bitbrick이 동적으로 결합되어 Fused PE(Processing Engine)를 형성하며, 2비트, 4비트, 8비트, 16비트 등 2의 제곱수의 비트폭 연산을 지원한다. AI 모델에 더 최적화된 양자화가 가능하기 때문에 기존의 하드웨어보다 에너지 소비는 낮으면서도 연산 속도는 빠른 하드웨어를 구현할 수 있다.

2. 가지치기 기법이 가능한 희소(Sparse) 데이터 연산 가속기

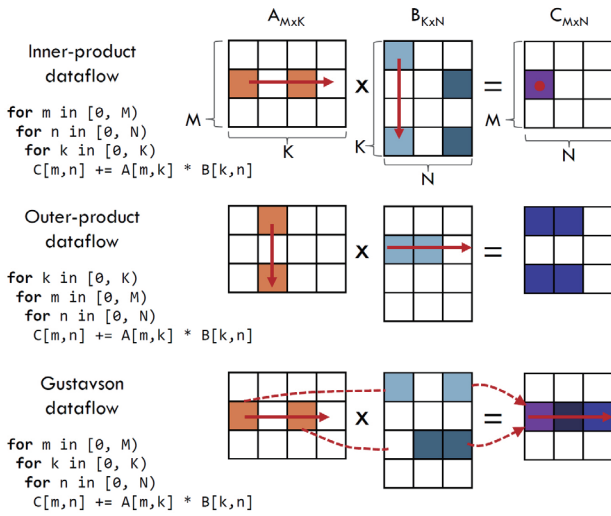
가지치기 기법을 적용한 후에는 AI 모델의 데이터가 희소해진다. 희소해진 데이터를 효율적으로 연산하는 방법에 대한 가속기 연구를 소개한다.

1) CNN 기반 가지치기 가속기 데이터플로우(Dataflow)

[24]는 CNN 모델에 대하여 희소 데이터들의 연산을 최적화한 가속기이며, 〈그림 13〉은 SCNN의 Dataflow를 나타낸 그림이다. [24]는 압축 인코딩과 데이터의 인덱싱 값을 이용하여 불필요한 연산을 선별하여 데이터 전송을 없앴 후, 카르테지안 곱(Cartesian-Product)를 활용하여 0이 아닌 값만 선택하여 곱셈하고, 이 곱셈 결과 값을 재사용한다. 또한 평면 타일링(Plannar-Tiled)을 통해 입력 데이터를 조각으로 나누어 병렬화하며, 입력 데이터



〈그림 13〉 CNN 기반 가지치기 가속기 SCNN의 데이터플로우 [24]



〈그림 14〉 3가지 행렬 연산 방식 데이터플로우 비교 [26]

고정(Input - Stationary)을 통해 데이터 재사용을 극대화하여 희소 데이터들 간의 연산을 가속한다.

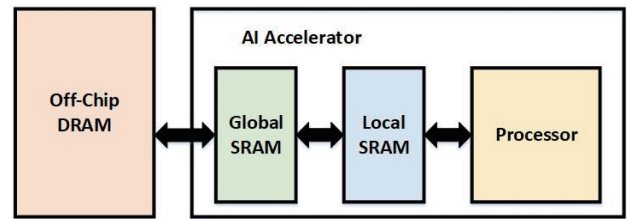
2) 희소 행렬 연산을 최적화한 가속기 데이터플로우

〈그림 14〉는 기존의 행렬 연산 방식인 내적(Inner Product), 외적(Outer Product) 데이터플로우와 희소 행렬 처리를 최적화한 구스타프슨 데이터플로우의 동작을 설명한 그림이다.

내적 및 외적 방식은 전통적인 행렬 처리 방식으로, 0 값에 대해서도 연산을 진행하기 때문에 불필요한 연산이 발생하며, 메모리 접근에 비효율적이다. 반면, 구스타프슨 알고리즘은 희소 행렬 간 곱셈을 가속화하기 위한 알고리즘으로, 메모리 접근 패턴을 최적화한다. 행렬의 0이 아닌 값 요소를 행 단위로 처리하면서 열 인덱스 값을 정렬하여 효율적으로 계산한다. 또한, 하드웨어 가속기에서 이를 구현할 때, 여러 코어가 동시에 독립적으로 계산할 수 있도록 하드웨어 가속기를 설계하여 연산 처리 효율이 극대화된다.

3. HW 경량화 기법을 구현한 Processing-In-Memory (PIM)

〈그림 15〉는 기존 AI 가속기에서의 데이터 흐름을 설명하는 구조이다. 이는 폰 노이만 아키텍처의 전형적인 구조로, 현대의 컴퓨터 및 스마트폰에서 주로 활용된다.



〈그림 15〉 컴퓨팅 구조

그러나 이 구조는 메모리 병목 현상이라는 근본적인 문제를 가지고 있다. 메모리에서 프로세서로 데이터를 전송하는 과정에서 전송 지연(Latency)이 발생해 연산 속도가 저하되고, 특히 대규모 데이터를 처리하는 AI 연산에서는 이 문제가 더 심각해진다. 또한, 데이터 전송 과정에서 많은 에너지가 소모되기 때문에 전력 효율에도 한계가 있다. 이러한 문제는 AI 반도체 설계에서 해결해야 할 중요한 과제로 인식되고 있다.

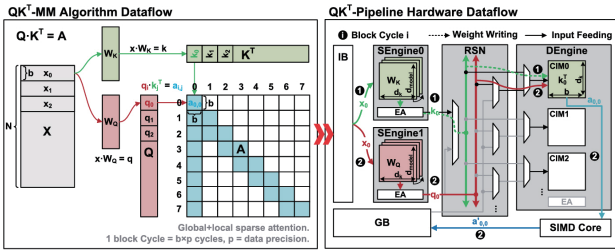
이 문제를 해결하기 위해 Processing-In-Memory(PIM) 기술이 활발히 연구되고 있다. PIM이란 기존 폰 노이만 아키텍처의 한계를 해결하기 위해 메모리 내부에 연산 작업을 수행할 수 있는 프로세서를 통합시킨 차세대 AI 반도체이다. 이를 통해 메모리 간 데이터 이동을 최소화하여 전력 소비를 크게 줄일 수 있다. 또한, AI 연산의 대부분을 차지하는 MAC 연산을 병렬로 처리함으로써 에너지 효율적으로 작업을 수행할 수 있어, 폰 노이만 구조의 한계를 극복할 유망한 대안으로 주목받고 있다.

1) 가지치기 기법이 가능한 희소 데이터 연산 PIM

희소성은 데이터나 모델에서 대부분의 값이 0이거나 비어 있는 상태를 의미한다. 즉, 희소성이 높다는 것은 0이 차지하는 비율이 높다는 뜻이다. 이를 통해 가중치와 활성화의 희소성을 인식하고, 불필요한 연산을 건너뛰어 성능을 최적화할 수 있다. 주로 PIM 아키텍처에서는 활성화나 가중치를 제어하는 컨트롤러를 통해 희소 데이터를 효율적으로 처리한다. 아래는 컨트롤러를 사용해 CNN과 Transformer의 희소 데이터 연산을 수행하는 PIM 아키텍처의 예시이다.

a. Transformer 기반 Pruning PIM 아키텍처

블록 희소 어텐션(Block Sparse Attention)은 토큰들



〈그림 16〉 블록 희소 어텐션 데이터플로우 [26]

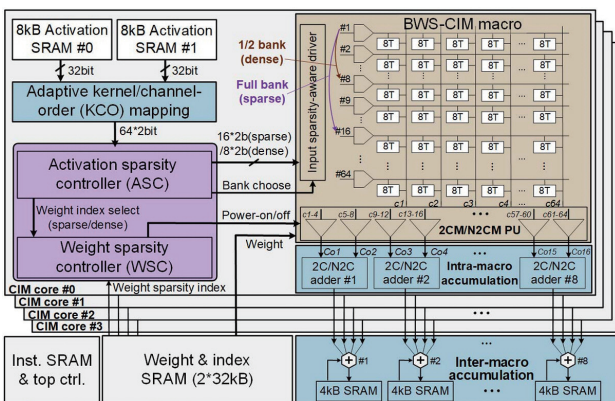
을 여러 블록으로 나눈 후, 선택된 블록에 대해서만 연산을 수행하여 전체 연산량을 줄이는 방법이다. 이를 통해 연산 효율을 크게 향상시킬 수 있지만, 모델 성능은 선택된 어텐션 패턴에 따라 달라질 수 있다. 〈그림 3〉의 Self-Attention 히트맵에서 볼 수 있듯이, 블록마다 중요도가 다르다. 따라서 상대적으로 중요한 블록들을 선택해 계산량을 줄이면서도 성능 저하를 최소화할 수 있다.

[26]은 블록 희소 어텐션을 적용한 PIM 아키텍처를 제안하였는데, 〈그림 16〉의 블록 희소 어텐션 패턴을 선택하여 연산을 수행함으로써, 전체 연산량을 줄이고 성능 저하를 최소화하였다.

b. CNN 기반 Pruning PIM 아키텍처

〈그림 17〉은 [27]에서 제안한 PIM 기반 가속기의 전체 아키텍처를 나타낸다.

이 아키텍처는 동적 희소성 스케일링(Dynamic Sparsity Scaling)을 통해 활성화와 가중치의 희소성을 감지하고 연산을 최적화한다. CNN 모델에서는 각 계층의 중요도가 다르므로, 중요도가 낮은 계층에서는 필



〈그림 17〉 CNN-based PIM 아키텍처 [27]

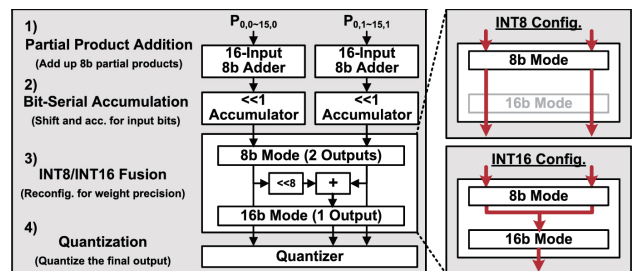
터 단위로 희소성을 적용해 에너지 효율을 높일 수 있다. [27]은 가중치에 필터 또는 블록 단위로 희소성을 부여하고, WSC(Weight Sparsity Controller)를 사용하여 에너지 효율을 향상시켰다. WSC는 0으로 이루어진 행을 감지해 해당 연산 유닛의 전원을 차단함으로써 불필요한 연산을 방지한다. 이를 통해 필터 단위에서 0으로 구성된 블록의 연산을 건너뛰어 연산 효율성을 극대화할 수 있었다. 또한, ASC(Activation Sparsity Controller)는 활성화 데이터가 모두 0이거나 중요하지 않은 데이터를 실시간으로 감지해 불필요한 연산을 생략한다. 이를 통해 활성화 값이 모두 0일 때 연산을 건너뛰고, 일부 희소성이 있는 경우에는 최소한의 연산만 수행함으로써 연산 효율을 높였다.

2) 혼합-정밀도 양자화 (Mixed-Precision Quantization)를 지원하는 PIM

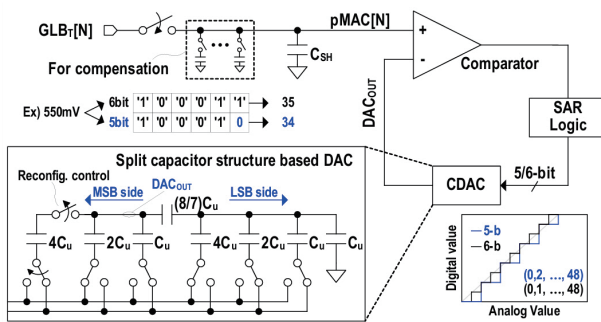
혼합-정밀도 양자화는 중요도가 낮은 연산에는 낮은 정밀도를, 성능에 중요한 핵심 연산에는 높은 정밀도를 적용하여 성능 저하를 최소화하는 기술이다. 주로 PIM 아키텍처에서 Digital 방식은 부분 합을 이용해 혼합-정밀도 양자화를 지원하고, Analog 방식은 ADC(Analog-to-Digital Converter)를 사용해 혼합-정밀도 양자화를 지원한다. 아래는 Digital과 Analog 방식의 혼합-정밀도 양자화를 지원하는 PIM 아키텍처의 예시이다.

a. 혼합-정밀도 양자화를 지원하는 Digital PIM

Transformer 모델에서 높은 성능을 유지하기 위해 어텐션 계층은 높은 정밀도가 요구된다. 반면, 완전 연결 계층은 상대적으로 낮은 정밀도를 사용해도 성능을 유지할 수 있다. [26]은 혼합-정밀도 양자화 방식을 지원하도록



〈그림 18〉 혼합-정밀도 지원 방식 [26]



(그림 19) 5/6비트 재구성 가능한 SAR-ADC [28]

설계하여 어텐션 계층은 16비트, 완전 연결 계층은 8비트 정밀도를 지원한다. <그림 18>에서 볼 수 있듯이, 어텐션 계층은 두 개의 8비트 덧셈 결과를 계산한 후 상위 8비트(MSB)에 8번의 시프트 연산을 수행하고, 이를 하위 8비트(LSB)와 더해 최종적으로 16비트로 연산한다. 반면, 완전 연결 계층은 시프트 연산 없이 바로 연산을 수행한다. 이를 통해 성능 저하를 최소화하면서 불필요한 연산을 줄여 에너지 효율을 향상시켰다.

b. 혼합-정밀도 양자화를 지원하는 Analog PIM

Analog PIM은 Digital PIM과 달리, 아날로그 출력을 디지털로 변환하기 위해 ADC(Analog-to-Digital Converter)를 사용해야 한다. ADC의 해상도가 높을수록 모델 성능은 증가하지만, 에너지 소모도 함께 증가한다. 따라서 중요한 연산에만 높은 해상도를 적용한다면 성능 저하를 최소화하면서 에너지 효율을 높이는 것이 가능하다.

[28]에서는 <그림 19>처럼 5/6비트 재구성 가능한 ADC를 사용하여 계층의 중요도에 따라 해상도를 조정함으로써 에너지 효율을 향상시켰다.

V. 결론

AI 분야의 급속한 발전으로 인해 모델 파라미터 수에 따른 연산량이 크게 증가하면서 에너지 소비와 성능 효율성이 중요한 문제로 대두되고 있다. 이에 본 원고에서는 효율적인 AI 반도체의 주요 기술 동향을 살펴보았다. 앞으로 이러한 효율적인 가속기는 지속적으로 발전하여 의료 진단, 스마트 시티, 환경 모니터링 등 다양한 분야에서

중요한 역할을 할 것으로 기대된다.

참고 문헌

- [1] A. Vaswani, et al., "Attention is all you need," NIPS, 2017
- [2] R. Yamashita, et al., "Convolutional neural networks: an overview and application in radiology," Insights Imaging, Aug. 2018
- [3] S. Han, et al., "EIE: Efficient Inference Engine on Compressed Deep Neural Network," ISCA, Jun, 2016
- [4] A. Gholami, et al., "A Survey of Quantization Methods for Efficient Neural Network Inference," CRC Press, 2022
- [5] R. Novkin et al., "Approximation- and Quantization-Aware Training for Graph Neural Networks," TOC, Feb, 2024
- [6] C. Sakr, et al., "Optimal Clipping and Magnitude-aware Differentiation for Improved Quantization-aware Training," PMLR, Jun, 2022
- [7] Y.-S. Tai et al., "TSPTQ-ViT: Two-Scaled Post-Training Quantization for Vision Transformer," ICASSP, Jun, 2023
- [8] Y. Wang et al., "AQA: An Adaptive Post-Training Quantization Method for Activations of CNNs," TOC, Aug, 2024
- [9] S.-T. Lin et al., "LG-LSQ: Learned Gradient Linear Symmetric Quantization for Low-Precision Integer Hardware," AICAS, Jun, 2023
- [10] X. Zhao et al., "Linear Symmetric Quantization of Neural Networks for Low-precision Integer Hardware," ICLR, 2020
- [11] T. Zhao et al., "Lightweight Network for Modulation Recognition Based on Stochastic Pruning-Asymmetric Quantization," APCC, Jan, 2023
- [12] L. Ma et al., "Deep Progressive Asymmetric Quantization Based on Causal Intervention for Fine-Grained Image Retrieval," TMM, 2024
- [13] Z. Liao, et al., "Can Unstructured Pruning Reduce the Depth in Deep Neural Networks?," ICCV, 2023
- [14] M. Sun, et al., "A simple and effective pruning approach for large language models," ICLR, 2023
- [15] M. Xia, et al., "Structured pruning learns compact and accurate models," ACL, 2022

- [16] X. Zhang, et al., "Filter Pruning via Learned Representation Median in the Frequency Domain," TCYB, May 2023
- [17] X. Ma et al., "Non-Structured DNN Weight Pruning—Is It Beneficial in Any Platform?," TNNLS, 2022
- [18] Li, Guiying, et al., "Stage-wise magnitude-based pruning for recurrent neural networks," TNNLS, 2022
- [19] J. Back, et al., "Magnitude Attention-based Dynamic Pruning," CVPR, Jun. 2023
- [20] S. Ren, et al., "Pruning Pre-trained Language Models with Principled Importance and Self-regularization," ACL, May 21, 2023
- [21] H. Wang, et al., "Neural pruning via growing regularization," ICLR, 2022.
- [22] S. Muralidharan, et al., "Compact Language Models via Pruning and Knowledge Distillation," arXiv, Jul. 2024
- [23] H. Sharma et al., "Bit Fusion: Bit-Level Dynamically Composable Architecture for Accelerating Deep Neural Network," ISCA, Jun. 2018
- [24] A. Parashar, et al., "SCNN: An Accelerator for Compressed-sparse Convolutional Neural Networks," SIGARCH News, Jun. 2017
- [25] G. Zhang, et al., "Gamma: leveraging Gustavson's algorithm to accelerate sparse matrix multiplication," ASPLOS, Apr. 2021
- [26] F. Tu, et al., "TranCIM: Full-Digital Bitline-Transpose CIM-based Sparse Transformer Accelerator With Pipeline/Parallel Reconfigurable Modes," JSSC, Jun. 2023
- [27] J. Yue, et al., "14.3 A 65nm Computing-in-Memory-Based CNN Processor with 2.9-to-35.8TOPS/W System Energy Efficiency Using Dynamic-Sparsity Performance-Scaling Architecture and Energy-Efficient Inter/Intra-Macro Data Reuse," ISSCC, Feb. 2020
- [28] K. Lee, et al., "A Charge-Sharing based 8T SRAM In-Memory Computing for Edge DNN Acceleration," DAC, Feb. 2021

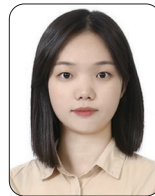


박재범

- 2024년 2월 광운대학교 전자융합공학과 학사
- 2024년 9월 ~ 현재 고려대학교 전기전자공학부 석사과정

〈관심 분야〉

SRAM, Processing-In-Memory 설계



안민영

- 2021년 2월 경희대학교 전기전자공학부 학사
- 2024년 9월 ~ 현재 고려대학교 전기전자공학부 석사과정

〈관심 분야〉

저전력 뉴럴네트워크 가속기 설계



박종선

- 1998년 2월 고려대학교 전자공학과 학사
- 2000년 8월 Purdue Univ., School of ECE 석사
- 2005년 12월 Purdue Univ., School of ECE 박사
- 2005년 5월 ~ 2008년 2월 미국 Marvell
Semiconductors, Staff
Design Engineer
- 2008년 ~ 현재 고려대학교 전기전자공학부 교수

〈관심 분야〉

신뢰성 강한 저전력 통신 및 디지털 신호처리 프로세서
설계, 저전력 임베디드 메모리 설계, 저전력 뉴럴네트워크
가속기 설계

영상처리를 위한 경량화 뉴럴프로세서 설계

I. 서론

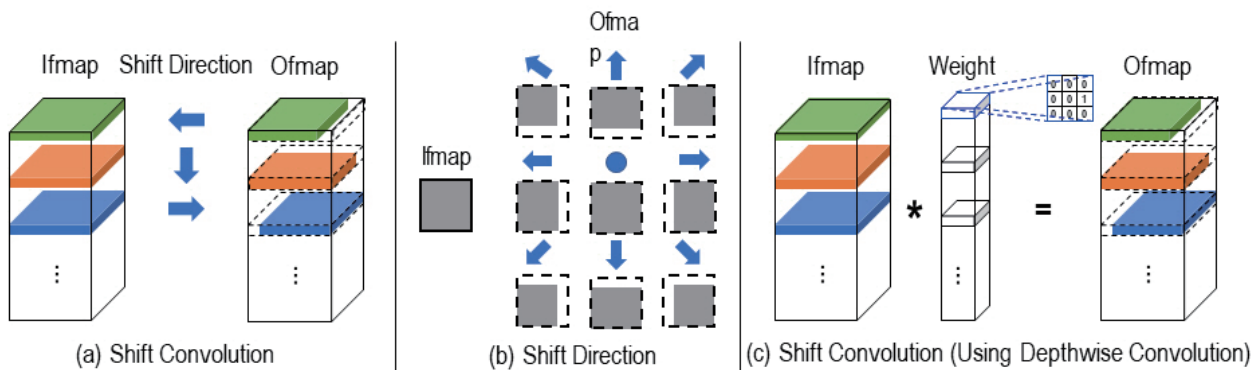
인공지능 기술의 발전과 함께 뉴럴네트워크는 다양한 분야에서 활용되고 있으며, 그 중에서도 Convolutional neural network (CNN)은 객체 감지나 이미지 필터링과 같이 일상 생활에서 자주 사용되는 영상 처리 작업들에 많이 활용되고 있다. 최근의 CNN 모델들은 높은 정확도를 얻기 위해 이전 모델들보다 더 많은 레이어를 가지고 있으며, 이는 계산 복잡성 증가로 이어진다. 그러나, 엣지 디바이스를 이용하여 이러한 모델을 계산하는 것은 실시간 처리를 어렵게 만들기 때문에 정확도 하락을 최소화하는 동시에 네트워크 모델을 압축해야하는 어려운 문제가 있다.

한편, 스마트 홈, 산업 자동화와 같은 보안 애플리케이션에서는 여러 영상을 연속적으로 처리하는 동영상 처리를 위한 뉴럴네트워크 모델을 흔히 활용하고 있다. 비디오 입력에서 객체 감지를 수행할 때 비디오의 각 프레임은 많은 수의 Multiply-accumulate (MAC) 계산을 필요로 하며 안정적인 서비스와 모니터링을 위해 항상 켜져있어야 한다. 결과적으로 많은 전력과 에너지를 소모하게 된다. 이러한 문제를 해결하기 위해 본 칼럼에서는 경량화 영상^[1] 및 동영상 인공지능 모델의 처리^[2]를 위한 소프트웨어-하드웨어 동시 최적화 연구 결과를 소개한다.

경량화 영상처리를 위해 본 연구팀은 일반적인 Convolution 연산을 Shift convolution으로 대체하여 모델 복잡성을 줄였다^[1]. Shift convolution에서 Spatial dimension convolution은 Shift 연산으로 대체되어 연산복잡도가 Zero-flop이 되도록 구현하였다^[3]. 이러한 Shift convolution 연산은 실제 MAC 연산에 포함되지 않아 총 연산 수를 크게 감소시킬 수 있지만, 기존 하드웨어들은 이러한 Shift convolution 연산을 지원하지 않아 하드웨어에서 실제 연산시에는 결



류 성 주
서강대학교



〈그림 1〉 Shift convolution 및 기존 하드웨어의 맵핑 예시^[1,3].

국 Standard convolution에 비해 훨씬 많은 연산 시간을 초래한다. 따라서 본 연구팀은 Teleport 아키텍처를 소개하였으며^[1], 이를 통해 Shift 연산을 지원하는 모듈을 추가하여 일반적인 Standard convolution에 비해 연산량 및 연산시간을 크게 감소시켰다.

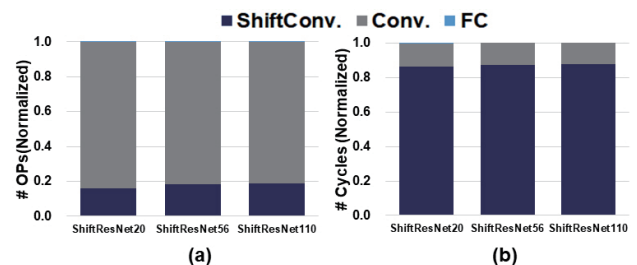
다음으로 경량화 동영상처리를 위해 본 연구팀은 유사한 동영상 프레임을 선택적으로 Skip하여 모델 복잡도를 줄였다^[2]. 비디오 객체 감지 작업에서 현재 프레임이 이전 프레임과 유사하여 추론 결과가 이전 프레임과 거의 동일할 것으로 예상되는 경우 현재 프레임의 Object detection을 위한 Backend 뉴럴네트워크를 건너뛰는 알고리즘을 제안하였고 이를 가속화하기 위한 Statues 하드웨어 구조를 제안 및 설계하였다.

따라서 본 칼럼에서는 위에서 언급한 두 가지의^[1-2] 영상 및 동영상의 경량화 인공지능 처리를 위한 소프트웨어-하드웨어 공동 최적화 방법을 소개한다. II장에서는 영상 모델의 경량화 처리를 위한 Teleport 소프트웨어-하드웨어 공동 최적화 설계 방법을 소개하고, III장에서는 동영상 모델의 경량화 처리를 위한 Statues 소프트웨어-하드웨어 공동 최적화 설계 방법을 소개하며 IV장에서 본 칼럼을 마무리한다.

II. 영상처리 경량화 가속기 설계

1. 배경지식: Shift Convolution

Shift convolution의 핵심 개념은 Input feature를 Channel 단위로 하나의 방향씩 Shift 처리하는 것이다



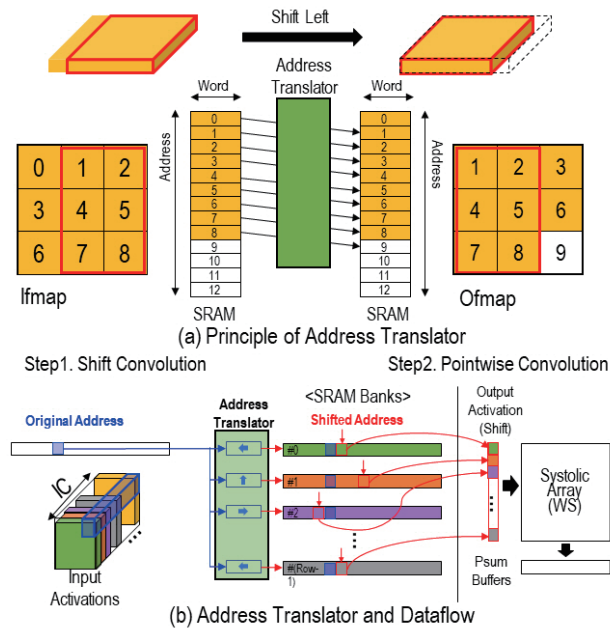
〈그림 2〉 Shift convolution 연산량 및 연산시간^[1].

〈그림 1a-b〉. 즉, 각 Channel의 Input feature들은 해당 Channel로 미리 학습된 이동 방향에 따라 독립적으로 이동한다는 것이다. 이러한 채널별 Shift 연산은 실제로 Depthwise convolution과 유사하다. MobileNet이 기존의 Standard convolution을 Depthwise convolution과 Pointwise convolution으로 분해하는 것처럼, ShiftNet은 Standard convolution을 Shift convolution과 Pointwise convolution으로 분해한다. 그러나, 기존 인공지능 하드웨어 가속기들은 이러한 Shift convolution을 지원하지 않는다. ShiftNet을 기존 하드웨어 가속기에서 연산할 때, Shift convolution은 3×3 Filter를 갖는 Depthwise convolution으로 맵핑되어 〈그림 1c〉 하드웨어에 입력된다. 결국 〈그림 2〉와 같이 Shift convolution의 연산량은 전체 연산의 16-19%만 차지하는 반면 Systolic array에서 연산할 때 전체 연산시간의 86-88%의 매우 큰 부분을 차지하는 문제가 있다. 따라서, 이러한 Shift convolution 경량화 방법은 기존 하드웨어에서의 Shift convolution 맵핑시에 성능향상을 보여주지 못한다.

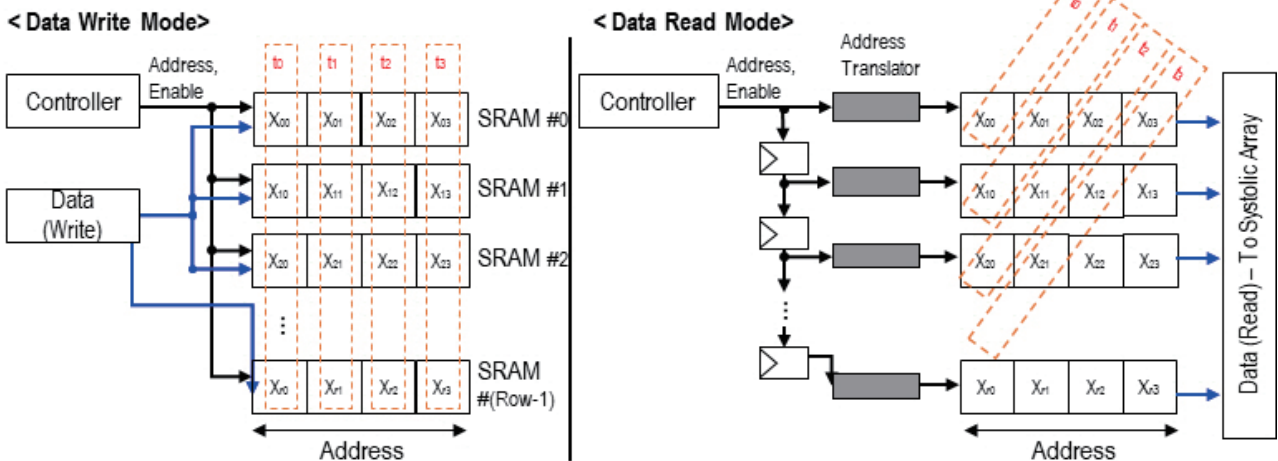
2. 영상처리 인공지능 모델 경량화를 위한 Teleport 하드웨어 가속기

본 연구에서는 Zero-flop 기반의 Shift convolution을 계산하기 위한 Custom logic 설계방법을 제안한다. Shift convolution은 Activation memory에 저장된 Feature들을 다른 메모리 주소로 옮기는 것이 필요하며, 이러한 방식은 메모리에서 읽은 후에 이를 다른 곳에 쓰는 일이 필요하므로 반복적인 메모리 읽기 및 쓰기 작업에 대한 부담이 필요하다. 이러한 문제를 해결하기 위

해 Address translator를 설계하였으며^[1], 이를 통해 메모리 읽기 및 쓰기 작업을 줄이고, Shift convolution을 Depthwise convolution이나 소프트웨어 추가 구현으로 대체하지 않아도 된다. <그림 3a>는 Shift convolution을 위한 Address translator를 보여준다. 이는 왼쪽으로 Shift하는 예시만 설명하며, 다른 방향의 Shift 연산 또한 동일한 방식으로 쉽게 구현할 수 있다. 이를 통해 Shift convolution 연산에서 Feature를 새로운 메모리 주소로 이동하는 부담 대신, 메모리 읽기 동작 시에 원래의 Feature 주소를 가상의 Shift된 Feature 주소로 수정한다. Shift convolution의 정의에 따라 Input feature의 각 Channel 데이터는 각 Channel에 미리 학습된 Shift 방향으로 독립적으로 이동한다. 따라서, 이를 위해 Input feature의 각 Channel마다 하나의 Address translator가 할당된다. <그림 3b>는 이러한 Address translator를 적용한 하드웨어에서의 Dataflow를 보여준다. 먼저, 각 Input SRAM bank는 Input channel마다 하나씩 할당된다. 즉, Spatial 영역에서 연속적인 Input activation들은 하나의 SRAM bank에 저장되고, 서로 다른 Input channel의 Activation들은 서로 다른 SRAM bank에 저장된다. 둘째, Address translation 회로에서 Input activation의 메모리 주소가 이동된 주소로 변경 후에 출력된다. 이 단계에서 Shift 연산은 간단한 메모리 주소 변환 연산으로 대체된다. 각 SRAM bank에 할당된 Address translator는 미리 학습 후 정의된 Shift



<그림 3> Address translator^[1].



<그림 4> Systolic data loader를 이용한 쓰기 및 읽기 동작 모드^[1].

립적인 내용이므로 이에 대한 설명은 생략한다.

III. 동영상처리 경량화 가속기 설계

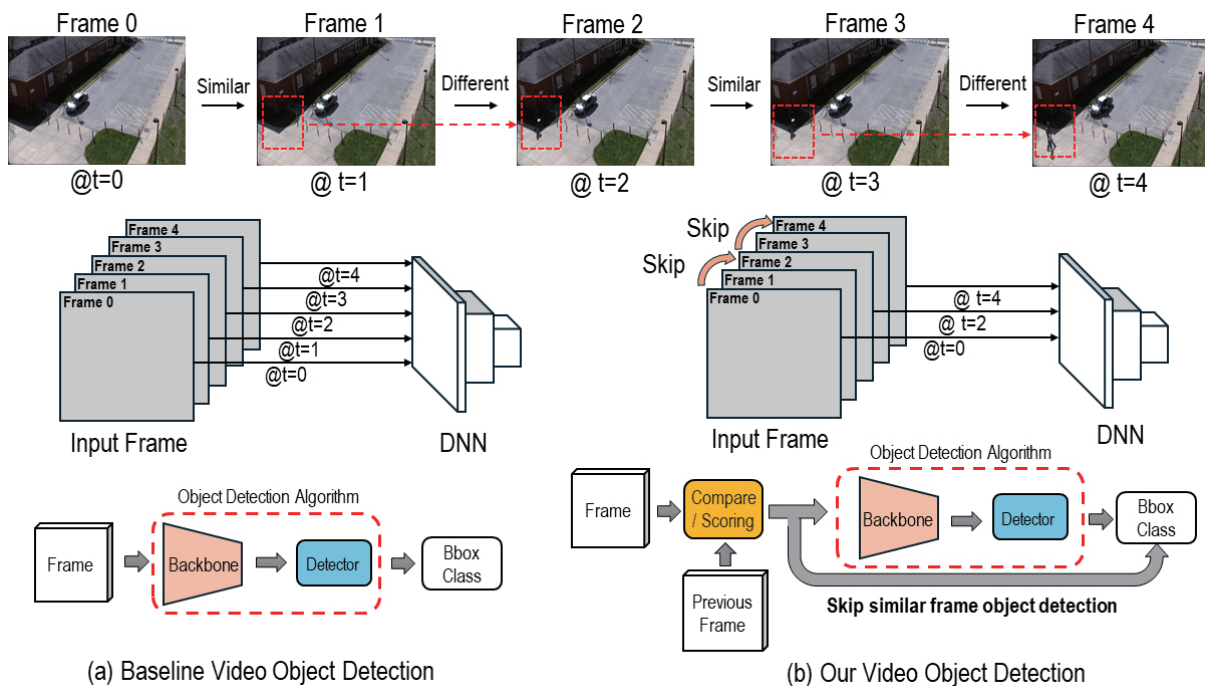
1. 배경지식: Video object detection

동영상에서의 Object detection 작업은 이미지 프레임에서 Object detection 작업을 연속으로 진행하는 것이며, 이 때 각 프레임 이미지는 MobileNet, ResNet과 같은 Feature extractor에 입력된다. 대부분의 Backbone 뉴럴네트워크는 일반적으로 많은 수의 Convolutional layer, Fully connected layer 및 기타 Special function 들로 구성된다. 이러한 Backbone depth와 Width는 일반적으로 Accuracy에 영향을 미치므로 높은 Accuracy를 위해 MAC 연산 수를 크게 증가시키는 경향이 있고 이는 높은 추론 연산 능력을 필요로 한다. 위에서 언급한대로 Video object detection은 Image object detection이 확장이므로, Video object detection에서는 Sequential한 Video frame이 입력으로 제공되고 Backbone 뉴럴네트워크는 주어진 Input frame에 대한 Inference 작업을 수행하여 해당 Frame에 대한 최종 Bounding box (bbox)

와 Class를 얻는다 <그림 7a>. 결과적으로 비디오 클립의 길이가 증가함에 따라 계산 복잡성도 선형으로 증가한다. 보안 및 모니터링 시스템에서는 여러 카메라에서 오는 방대한 양의 비디오 프레임을 동시에 처리해야 하고, 안정적인 모니터링 서비스를 제공하기 위해 장치가 항상 켜져 있어야 하기 때문에 많은 수의 엣지디바이스들의 전력 소모 문제가 발생한다.

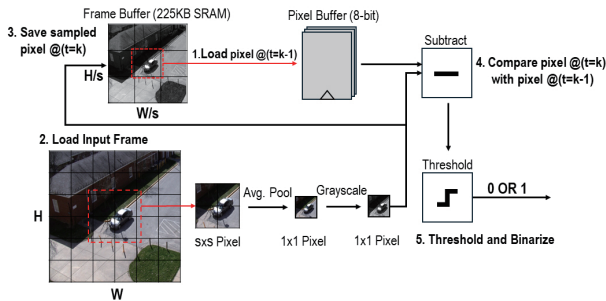
2. 동영상처리 인공지능 모델 경량화를 위한 Statues 하드웨어 가속기

본 연구에서는 기존 Video object detection 알고리즘의 한계를 극복하기 위해 엣지 보안 시스템에 알맞은 다음 내용을 제안하였다. <그림 7b>는 우리의 Video object detection 방법을 보여준다. 기존 방법 <그림 7a>와 비교하여, 우리는 연속적인 프레임 유사도에 따라 Backend 뉴럴네트워크를 선택적으로 연산하지 않는다. 조금 더 정확히 말하면, 현재 입력되는 프레임과 지난 프레임이 Frame downscaling module (FDM) 및 Component-based scoring module (CBSM) 모듈에 입력되어 두 프레임을 비교하고 유사도를 측정한다. 본 예시에서는 프레



<그림 7> 동영상 처리를 위한 Statues 인공지능 모델에 대한 개요 [2].

Frame Downscaling Module (FDM)



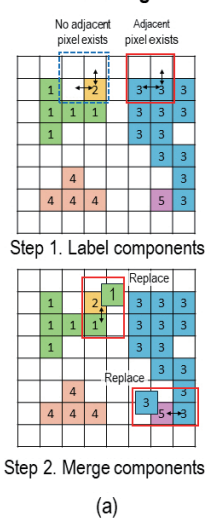
〈그림 8〉 Frame downscaling module (FDM)에 대한 설명 [2].

임 1/3이 직전 프레임 0/2와 유사하다. 따라서, 프레임 1/3은 Backend object detection 연산을 계속해서 수행하지 않으며, 그렇더라도 직전 프레임과 유사하기 때문에 직접 프레임의 뉴럴네트워크 연산 결과를 재사용한다. 반면, 프레임 2/4는 이전 프레임과 다소 차이가 있기 때문에 계속해서 Backend 뉴럴네트워크 연산을 진행한다.

〈그림 8〉은 제안한 FDM을 설명한다. 먼저 225KB SRAM으로 구성된 Frame buffer는 저용량으로 샘플된 이전 입력 프레임 ($t = K - 1$)을 갖는다. 이후 $t = K$ 에서의 현재 프레임을 Fetch할 때 이전 프레임과 마찬가지로 저용량으로 샘플링한다. 본 예시에서는 샘플링 Parameter로 $s = 2$ 를 사용하였으며 상황에 따라 다른 샘플링 수를 사용할 수 있다. 인접한 두 샘플된 프레임을

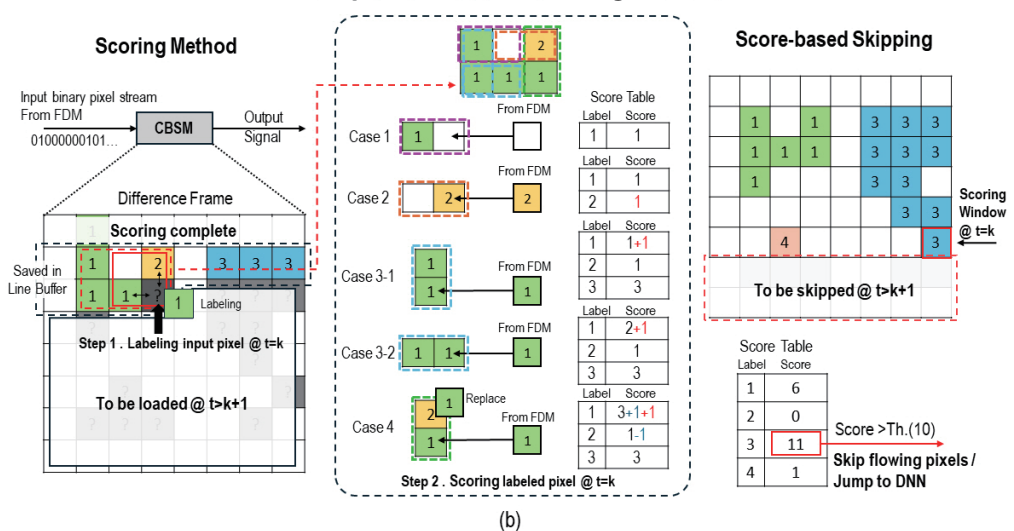
Pixel 단위로 뺄셈비교한 후에 미리 정의된 Threshold 수를 통해 Binarized 프레임을 만들고 이를 이후의 CBSM으로 전달한다. 〈그림 9〉는 제안한 CBSM을 설명한다. CBSM은 Connected-component labeling^[4]을 기반으로 하였으며, Connected-component labeling은 두 스텝으로 나뉜다. Pixel labeling stage에서는 Labeling을 왼쪽 위에 위치한 Pixel부터 오른쪽 아래에 위치한 Pixel까지 순차적으로 진행한다. 만약 왼쪽 혹은 위의 방향으로 인접한 Pixel이 유효한 경우, 그 Pixel과 같은 수로 Labeling을 진행한다 〈그림 9a〉. 반면, 만약 인접 Pixel이 비어있다면, Labeling을 위해 새로운 수를 사용한다. 그러나, 이러한 방법은 오른쪽과 아래에 위치하게 될 Pixel의 Label을 미리 예측할 수 없기 때문에 정확한 Labeling이 어려운데 이 경우에는 인접한 Pixel이 서로 다른 수로 Labeling 되어있다면, 같은 수를 갖도록 Component를 합치는 작업이 추가로 필요하다. 이렇게 위에서 언급한 Connected-component labeling 방법을 기반으로 CBSM은 Difference frame에 대한 Labeling을 다룬다 〈그림 9b〉. 만약 하나의 Component가 미리 정의된 Threshold보다 큰 경우에 이는 해당 Component가 연속된 두 프레임이 매우 유사하다는 점을 알려준다. 이렇게 Scoring을 진행하다가 미리 정해둔 유사도를 기준

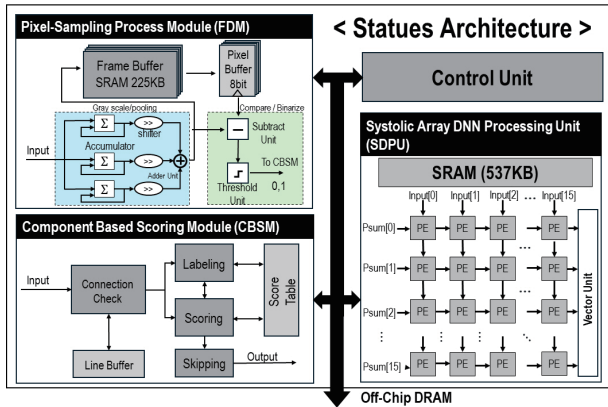
Connected-Component labeling



(a)

Component-based Scoring Module





〈그림 10〉 Statues 하드웨어 가속기의 Top-level 아키텍처 도식화 [2].

으로 두 프레임의 유사성이 판단되면 남은 Scoring 작업 및 Backend 뉴럴네트워크를 연산하지 않고 해당 Frame의 남은 연산을 모두 건너뛰고 이전 Frame의 Inference 결과를 그대로 재사용하여 현재 Frame의 연산 결과를 근사화한다. 〈그림 10〉은 Statues 하드웨어 가속기의 Top-level 아키텍처를 보여준다. Frame 유사도 판단을 위한 FDM 및 CBSM을 포함하고, Backend 뉴럴네트워크를 연산하기 위한 Systolic array DMM processing unit (SDPU)을 담고 있으나 Systolic array 대신 다른 연산기 마이크로아키텍처를 사용하여도 무방하다. 전체 진행 절차에서 FDM/CBSM과 SDPU는 Pipeline되어 성능의 추가 향상을 구현하였다.

IV. 전망과 결론

본 칼럼에서는 우리 연구팀의 영상 및 동영상 처리 경량화를 위한 소프트웨어-하드웨어 공동 최적화 연구 결과를 소개하였다^[1-2]. Teleport 및 Statues 하드웨어 아키텍처를 통해 엣지디바이스에서 저전력으로 경량화 모델을 충분히 구현 가능하며, 빠르게 발전하는 여러 뉴럴 네트워크에서 이러한 기술을 적용할 수 있길 기대한다. 또한, 본 연구팀은 새롭게 등장하는 여러 인공지능 모델에 대한 경량화 소프트웨어-하드웨어 공동 최적화 연구를 계속해서 진행하고 있으며 이를 통해 인공지능 모델을 저전력으로 구현 가능하는데 끊임없이 기여할 계획이다.

참고 문헌

- [1] Hyunmin Kim, Sungju Ryu, "Teleport: A High-Performance ShiftNet Hardware Accelerator with Fused Layer Computation," ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), ACM/IEEE, 2023.
- [2] Yeonggeon Kim, Hyunmin Kim, Sungju Ryu, "Statues: Energy-Efficient Video Object Detection on Edge Security Devices with Computational Skipping," ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED), ACM/IEEE, 2024.
- [3] Bichen Wu, Alvin Wan, Xiangyu Yue, Peter Jin, Sicheng Zhao, "Shift: A Zero FLOP, Zero Parameter Alternative to Spatial Convolutions," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE/CVF, 2018.
- [4] Jia Wei Tang, Nasir Shaikh-Husin, Usman Ullah Sheikh, and M. N. Marsono, "linked list run-length-based single-pass connected component analysis for real-time embedded hardware," Journal of Real-Time Image Processing, 2018.



류성주

- 2015년 2월 부산대학교 전자전기공학부 학사
- 2021년 2월 POSTECH 창의IT융합공학과 박사
- 2021년 1월 ~ 2021년 8월 삼성종합기술원 연구원
- 2021년 9월 ~ 2023년 2월 숭실대학교 조교수
- 2023년 3월 ~ 현재 서강대학교 조교수

〈관심 분야〉

디지털회로설계, 인공지능프로세서, 인메모리컴퓨팅

AI 반도체를 위한 TinyML 기술 동향

I. 서론

딥러닝(Deep Learning)은 다양한 산업에 혁신을 가져온 기술로, 이미지 인식, 자연어 처리, 자율 주행, 의료 진단 등 여러 분야에서 그 잠재력을 입증하고 있다. 딥러닝 모델은 대량의 데이터를 처리하고 복잡한 패턴을 학습하는 강력한 기능을 통해, 인간의 인지 능력에 근접한 수준의 정확도를 달성하고 있다. 하지만 이러한 강력한 성능에는 상당한 비용이 따른다. 예를 들어, 수백만 개 이상의 매개변수를 가진 복잡한 신경망 모델은 학습과 추론 과정에서 막대한 연산을 수행해야 하며, 이로 인해 GPU와 같은 고성능 하드웨어가 필수적이다. 이는 곧 딥러닝 모델을 실행하기 위해서 강력한 컴퓨팅 리소스가 필요함을 의미한다. 또한, 이러한 모델은 복잡하고 반복적인 작업을 처리하기 위해 대용량 메모리를 요구하며, 일반적으로 클라우드 기반에서 실행된다. 하지만 클라우드로 데이터를 전송할 경우 높은 대기 시간, 왕복 지연, 보안 및 개인 정보 보호 문제 등의 이슈가 발생할 수 있다.

따라서 막대한 연산 비용과 클라우드 전송 비용을 줄이기 위해 엣지 디바이스(Edge Device)에서 딥러닝 모델을 수행하는 연구들이 활발히 진행되고 있다. 클라우드 기반 처리에서는 수집된 데이터를 클라우드로 전송해야 하는 반면, 엣지 디바이스에서는 데이터를 수집한 현장에서 직접 처리할 수 있다. 이는 특히 실시간 처리와 저지연 응답이 중요한 애플리케이션에서 큰 이점을 제공한다. 엣지 디바이스에서의 데이터 처리는 네트워크 대역폭 사용을 최소화할 수 있고, 클라우드로의 의존도를 줄여 안정성과 보안을 향상시킬 수 있다.

예를 들어, 자율 주행 차량, 스마트 제조, 헬스케어 등과 같은 분야에서는 데이터가 빠르게 처리되어야 하고, 인터넷 연결이 불안정하거나 없는 환경에서도 원활한 작동이 필요하다. 이러한 이유로, 엣지 컴퓨팅



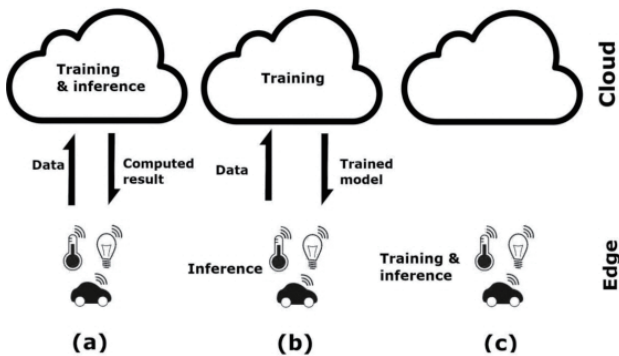
임형철
숭실대학교



이주석
숭실대학교



이성수
숭실대학교



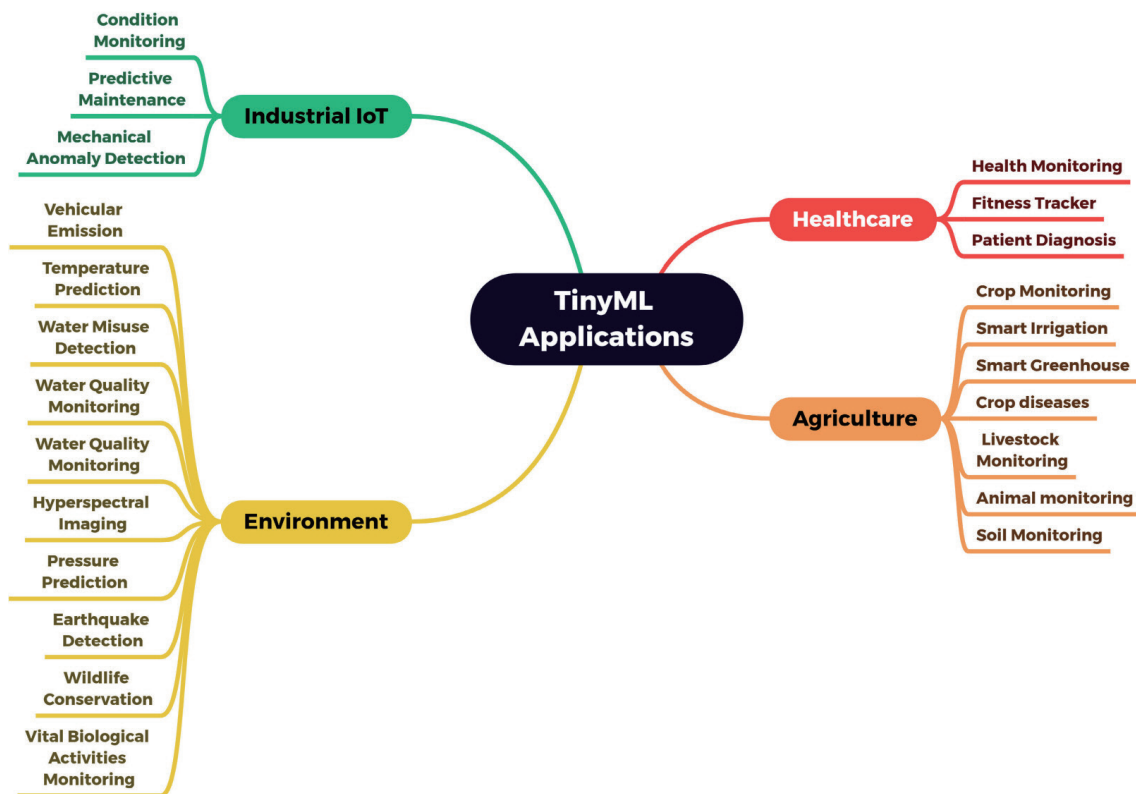
〈그림 1〉 훈련 및 추론을 위한 세 가지 접근 방식

이 딥러닝을 활용한 실시간 응용 프로그램에서 점점 더 중요한 역할을 차지하고 있다. 더 나아가, 엣지 디바이스에서 딥러닝 모델을 실행하면 사용자의 개인 정보 보호도 강화될 수 있다. 데이터를 중앙 서버로 전송하지 않고, 로컬에서 처리함으로써 데이터 유출의 위험을 줄일 수 있기 때문이다. 따라서 엣지 디바이스에서의 딥러닝 모델 수행은 비용 절감, 성능 향상, 그리고 보안 강화를 동시에 달성할 수 있는 중요한 기술로 자리 잡고 있다.

딥러닝 모델의 훈련과 추론은 〈그림 1〉에서 볼 수 있듯이 크게 세 가지 방식으로 나눌 수 있다. 첫째, 1(a)와 같이 클라우드에서 훈련을 수행하고, 클라우드에서 추론한 결과를 전달하는 방식이 있다. 둘째, 〈그림 1(b)〉처럼 클라우드에서 훈련을 수행한 후, 엣지 디바이스에서 추론을 수행하는 방식이 있다. 마지막으로, 〈그림 1(c)〉에서 보이는 것처럼 엣지 디바이스에서 훈련과 추론을 모두 수행하는 방식이 있다. 본 고에서는 엣지 디바이스에서의 추론에 초점을 맞추고자 한다. 따라서 다음 장에서는 엣지 추론을 위한 TinyML의 정의와 과제에 대해 살펴본다.

II. TinyML

TinyML은 매우 낮은 전력으로 엣지 디바이스에서 센서 데이터를 분석할 수 있는 알고리즘, 하드웨어, 소프트웨어를 포함한 빠르게 성장하는 머신 러닝(Machine Learning) 기술 및 응용 분야이다. TinyML 시스템은 점차 여러 상업적 응용 분야에 채택되고 있으며, 새로운 시



〈그림 2〉 TinyML 응용 분야의 분류

시스템들도 등장하고 있다. 동시에, 알고리즘, 네트워크, 모델 분야에서도 상당한 진전이 이루어지고 있다. 이는 TinyML 기술이 딥러닝 모델을 수십억 개의 IoT 장치와 마이크로컨트롤러(MCU)에 압축하여 AI 응용의 범위를 확장하고, 어디서나 지능적인 서비스를 제공하는 유비쿼터스 인텔리전스(Ubiquitous Intelligence)를 가능하게 하기 때문이다. 결과적으로, TinyML은 오늘날 저전력의 소형 기기들을 위한 기술과 응용에 중점을 둔 유망한 AI 대안으로 간주되고 있다.

TinyML은 다양한 실용적 응용 분야를 가지고 있으며, 이를 <그림 2>에 나타내었다. 본 고에서는 이러한 대표적인 응용 분야들에 대해 간략히 설명하고자 한다.

- 맞춤형 의료 서비스 : TinyML 기술은 스마트워치와 같은 웨어러블 기기에 적용되어 사용자의 활동과 산소 포화 상태를 지속적으로 추적하며, 이를 통해 건강에 대한 리포트를 제공한다. 또한, 신체 자세 추정으로 건강 관리에 활용될 수 있다.
- 스마트 차량 및 운송 : TinyML은 클라우드 연결 없이 물체 감지, 차선 감지 및 의사 결정을 수행할 수

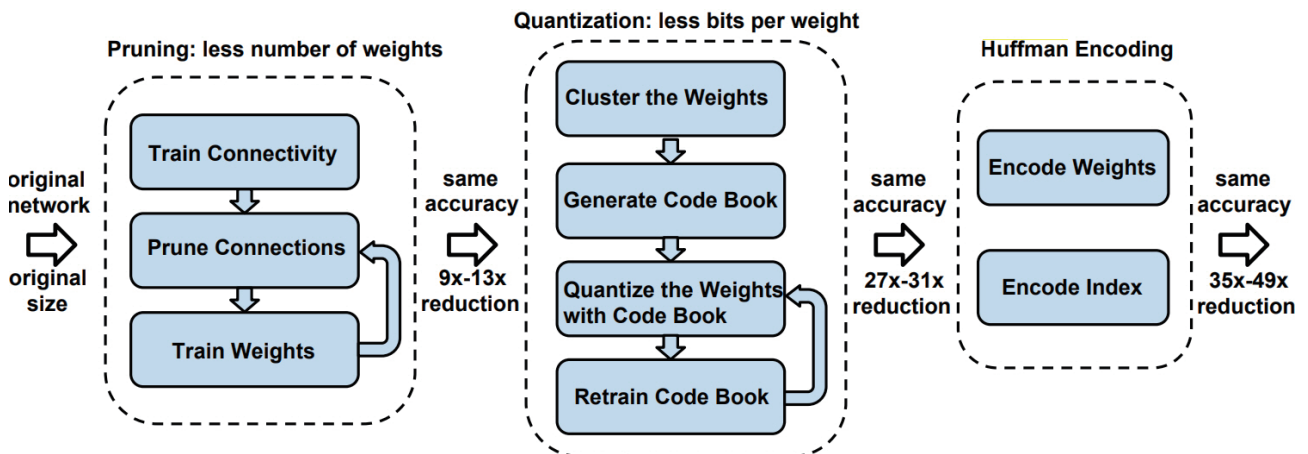
있어 자율 주행 시나리오에서 낮은 지연 시간으로 높은 정확도의 결과를 제공한다.

- 이상 감지 : 강력한 컴퓨팅 장치에 연결하지 않고도 로봇과 센서에 이상 감지 기능을 탑재할 수 있어, 비정상적인 패턴과 행동을 실시간으로 모니터링하고 식별할 수 있다.
- 생태학과 농업 : TinyML은 멸종 위기에 처한 종의 보존이나 기상 활동을 예측하기 위해 생태, 농업, 환경 애플리케이션을 지원할 수 있다.

하지만 일반적으로 딥러닝 모델은 높은 연산 비용을 수반하며, 이는 마이크로컨트롤러와 같은 디바이스의 엄격한 자원 제약 때문에 TinyML 응용 분야에서의 실현이 매우 어렵다. DRAM이 없고 운영 체제(OS)도 없으며, SRAM은 256 kB 이하이고, FLASH는 읽기 전용인 등 엄격한 메모리 제약이 존재하기 때문이다. 이러한 디바이스에서 사용할 수 있는 자원은 모바일 플랫폼보다 훨씬 적다. 실제로 온칩 메모리는 모바일 장치보다 1,000배, 클라우드 GPU보다 10만에서 100만 배 작아서 딥러닝 모델의 배포가 극도로 어렵다는 점을 알 수 있다. <표 1>에

<표 1> GPU, 모바일, 마이크로컨트롤러 하드웨어 제약 사항 및 널리 사용되는 딥러닝 모델의 메모리 사용량

	Cloud AI (NVIDIA V100)	Mobile AI (iPhone 11)	Tiny AI (STM32F746)	ResNet-50	MobileNetV2	MobileNetV2 (int8)
Memory	16GB	4GB	320kB	7.2MB	6.8MB	1.7MB
Storage	TB~PB	64GB	1MB	102MB	13.6MB	3.4MB



<그림 3> Deep Compression : 가지치기, 양자화 및 허프만 코딩

서 볼 수 있듯이, ARM Cortex-M7 MCU (STM32F746)는 320 kB SRAM과 1 MB 플래시 메모리를 갖추고 있어 기존 딥러닝 모델을 실행하는 것은 불가능하다. 예를 들어, ResNet-50은 해당 메모리 제한을 100배 초과하고, MobileNetV2는 22배를 초과한다. 심지어 int8 양자화 버전의 MobileNetV2조차도 여전히 메모리 제한을 5.3배 초과하여, 필요한 하드웨어 리소스와 실제 사용 가능한 메모리 사 사이에 큰 격차가 있음을 보여준다. 다음 장에서는 이와 같은 기존 딥러닝 모델을 리소스가 제한된 디바이스와 호환되도록 만드는 압축 기술들에 대해 살펴본다.

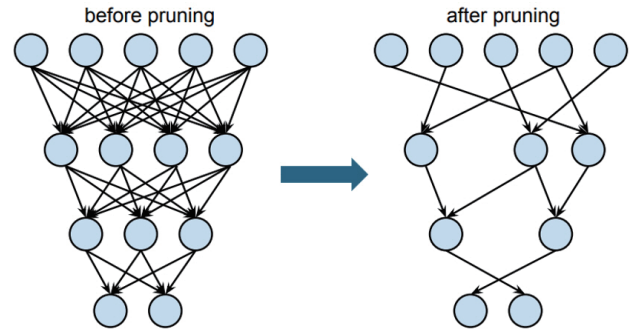
III. 딥러닝 경량화 기술

딥러닝 모델을 경량화하는 대표적인 방법으로 Deep Compression 기법이 있다. 2016년 ICLR에서 발표된 것으로, <그림 3>에 그 과정을 나타내었다^[5]. Deep Compression은 가지치기(pruning), 양자화(Quantization), 허프만 코딩(Huffman Coding)의 3단계로 이루어져 있다. 이번 장에서는 이 세 가지 기법들에 대해 간략히 소개하고자 한다.

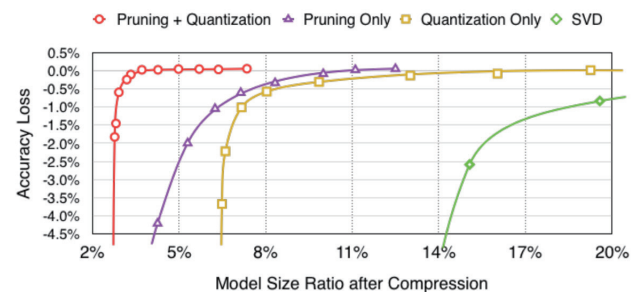
1. 가지치기(Pruning)

Pruning은 가지치기라는 뜻으로 딥러닝 모델에서 덜 중요한 가중치나 필터를 제거하는 대표적인 압축 기술 중 하나로 모델의 효율성을 높이는 데 큰 역할을 한다. 가지치기의 주요 장점은 매개변수의 크기를 줄여 메모리 사용을 최소화하고, 추론 시간을 단축시킬 수 있다는 점이다. 실제로 이러한 가지치기 과정은 우리의 뇌에서도 일어난다. 유아기를 거쳐 생성된 시냅스 중, 반복적인 자극을 받는 신경망의 시냅스는 더욱 강화되지만, 자극이 적은 시냅스는 가지치기 과정을 통해 사라진다. 이를 통해 인간의 뇌는 불확실한 환경 변화에 빠르게 적응할 수 있으며, 유전적 부담을 덜어주면서 효율적인 학습 및 적응 능력을 극대화할 수 있다.

<그림 4>는 가지치기 전후의 모델을 비교한 것이다. 가지치기는 일반적으로 구조적 가지치기와 비구조적 가지



<그림 4> Pruning 전과 후의 뉴런 구조



<그림 5> 경량화 방법에 따른 압축률 및 정확도 손실

치기로 구분되며, 이는 규칙성과 관련이 있다. 가지치기 방식에 따라 하드웨어 아키텍처의 효율성 및 예측 정확도에 차이를 미치게 된다. 가지치기를 통해 상대적으로 큰 가중치를 유지함으로써 정확도를 보존할 수 있으며, 규칙적인 가지치기를 적용하면 하드웨어 가속이 더 용이해진다. 예를 들어, 특정 블록, 행, 또는 필터 단위로 가중치를 제거하는 coarse-grained sparsity 기법^[x]으로 정확도의 손실 없이 더 높은 압축률을 달성할 수 있음을 실험을 통해 확인할 수 있다.

참고로, 드롭아웃(Dropout)은 과적합을 방지하기 위해 모델 훈련 중 무작위로 뉴런을 비활성화하여, 각 뉴런이 독립적으로 학습하도록 만드는 기법이다. 드롭아웃의 경우, 각 매개변수는 훈련 중에 확률적으로 비활성화되지만, 추론 시에는 모든 뉴런이 다시 활성화되어 영향을 받지 않는다. 반면, 가지치기는 한 번 제거된 가중치나 뉴런이 영구적으로 삭제되어, 훈련뿐만 아니라 추론 단계에서도 더 이상 사용되지 않는다는 차이점이 있다.

2. 양자화(Quantization)

대부분의 딥러닝 모델에서 가중치와 활성화는 32비트

또는 64비트 부동소수점 형식으로 표현된다. 양자화는 이러한 표현을 8비트 이하의 고정소수점 수로 줄이는 방법을 말한다. 고정소수점 연산은 부동소수점 연산에 비해 더 빠르고 에너지 효율적이기 때문에 모델의 메모리 사용량과 연산에 필요한 처리량을 크게 줄일 수 있다. 또한, 양자화로 비트 수가 줄어들면 하드웨어에서 더 효율적인 연산이 가능하다. 양자화는 모델 학습 중에 적용할 수 있으며, 학습이 완료된 후에도 적용할 수 있다. 하지만 양자화 시 정밀도가 지나치게 낮아지면 모델의 정확도가 심각하게 떨어질 수 있기 때문에 모델의 정확도와 가중치 및 활성화의 정밀도 사이에 적절한 타협이 필요하다.

양자화 비트 길이는 고정되거나 가변적일 수 있다. 고정 양자화에서는 네트워크의 모든 레이어, 필터, 활성화 등에 동일한 비트 길이가 적용된다. 이는 구현이 간단하고 일관성을 유지할 수 있지만, 특정 레이어나 가중치가 더 높은 정밀도를 필요로 할 때는 효율적이지 않을 수 있다. 반면, 가변 양자화는 네트워크의 레이어별, 필터별, 또는 가중치와 활성화에 대해 서로 다른 비트 길이를 적용하는 방식이다. 이 방법은 성능 최적화를 위해 정밀도가 필요한 부분에는 더 높은 비트 길이를 할당하고, 그렇지 않은 부분은 낮은 비트 길이를 적용하여 메모리 사용량과 연산량을 더욱 줄일 수 있다. 그러나 설계 및 구현이 고정 양자화에 비해 더욱 복잡하기 때문에 하드웨어 친화적이지 않다는 단점이 있다. 따라서 양자화 전략은 모델의 특성과 대상 하드웨어의 제약을 고려하여 신중히 선택해야 한다. 양자화는 가지치기와 함께 사용할 경우, 정확도를 유지하면서 모델의 크기를 대폭 감소시킬 수 있다.

〈그림 5〉는 가지치기와 양자화를 개별적으로 또는 함께 수행할 때 다양한 압축률에 따른 정확도를 보여준다. 먼저 보라색 선은 가지치기만을 적용한 모델을 나타내며, 원래 크기의 8% 이하로 압축하면 정확도가 크게 저하되기 시작한다. 노란색 선은 양자화만을 적용한 모델로, 마찬가지로 원래 크기의 8% 이하로 압축할 경우 정확도가 급격하게 떨어짐을 알 수 있다. 그러나 빨간색 선에서 볼 수 있듯이, 가지치기와 양자화를 모두 적용한 모델은 정확도의 손실 없이 모델 크기를 대폭 감소시킬 수 있음을 확인할 수 있다.

심볼 (Symbol)	고정 길이 부호 (Fixed-Length Code)	허프만 부호 (Huffman Code)
A	00	0
B	01	10
C	10	110
D	11	111

3. 허프만 코딩(Huffman Coding)

허프만 코딩은 무손실 압축 기법으로, 데이터를 압축한 후 정보 손실 없이 원래 상태로 복원할 수 있다. 이 기법은 데이터 세트의 각 심볼에 이진 코드를 할당하는 방식으로 작동하며, 자주 나타나는 심볼에는 더 짧은 코드를, 드물게 나타나는 심볼에는 더 긴 코드를 할당한다. 이 방법은 자주 등장하는 심볼을 짧게 표현함으로써 전체적인 메모리를 절약할 수 있다는 원리에 기반하고 있다.

예를 들어, 저장하려는 데이터가 'AAAABBBCCD'일 때, 고정 길이 부호화를 사용하면 표 1의 왼쪽 열과 같이 나타낼 수 있다. 따라서, 이 데이터는 '0000000001010110101100000000010101101011'으로 변환되며, 총 40비트로 저장된다. 이를 허프만 코딩을 적용하면 이보다 효율적으로 압축할 수 있다. 먼저, 각 심볼을 출현 빈도와 함께 하나의 노드로 구성한 후, 빈도에 따라 이진 트리를 생성한다. 빈도가 높은 심볼일수록 트리의 상위에 배치되고, 빈도가 낮은 심볼은 더 깊은 위치에 배치된다. 이후 트리의 경로를 따라 각 심볼에 고유한 이진 코드를 할당하는데, 왼쪽 노드로 이동하면 0, 오른쪽 노드로 이동하면 1을 추가하는 방식이다. 트리의 끝까지 도달하면 해당 심볼에 대한 최종 이진 코드가 결정된다. 이는 표 1의 오른쪽 열과 나타낼 수 있으며, 동일한 데이터는 '0000101010110110111000010101011011011'으로 총 38비트로 저장된다.

이와 같이 허프만 코딩을 적용하면, 빈도에 기반한 이진 코드 할당을 통해 자주 나타나는 심볼이 많은 데이터에서 매우 높은 압축 효율을 보일 수 있다. 앞서 설명한 가지치기와 양자화 기법을 결합하여 AlexNet 모델을 27배 압축할 수 있는데, 허프만 코딩까지 적용하면 무려 35배까지 압축이 가능하다. 이는 AlexNet 모델을 240MB

에서 6.9MB로 줄일 수 있다는 의미로, 이렇게 압축된 모델은 SRAM에 적재할 수 있을 만큼 작아진다. 결과적으로 에너지 소모가 큰 DRAM에 모델을 저장할 필요가 없어 전력 효율성을 향상시킬 수 있다.

IV. TinyML 추론을 위한 지표

딥러닝 모델 추론을 위한 하드웨어 플랫폼 간의 비교는 여러 지표를 바탕으로 이루어질 수 있다. 이번 장에서는 딥러닝 모델의 추론 성능을 평가하는 데 사용되는 몇 가지 표준 지표를 소개하고자 한다.

1. 모델 크기

딥러닝 모델의 크기는 추론 성능을 평가하는데 중요한 지표 중 하나이다. 엣지 디바이스 환경에서는 컴팩트한 네트워크 아키텍처가 특히 선호된다. 레이어의 수, 레이어당 노드 수, 활성화 함수, 노드 간 연결 수와 같은 하이퍼파라미터는 모델의 추론 성능에 직접적인 영향을 미친다. 따라서 모델을 설계할 때는 이러한 네트워크 아키텍처와 하이퍼파라미터를 신중히 고려해야 한다. 또한 가중치 및 편향과 같은 매개변수의 수는 모델의 메모리 요구사항을 반영하므로, 이는 성능을 평가하는 또 다른 중요한 지표로 작용할 수 있다.

2. 칩 면적과 비용

칩 면적은 추론 연산을 수행하기 위해 필요한 모든 계산을 구현하는 데 소요되는 실리콘 면적을 의미하며, 이는 필요한 게이트 수와 직접적으로 연관된다. 게이트 수가 많을수록 칩에 더 많은 공간이 필요하게 되어 칩의 전체 면적이 증가한다. 따라서 칩 설계 시 게이트 수를 줄이는 것은 크기와 비용을 절감하는 데 중요한 요소이다.

딥러닝 모델을 위한 칩 설계에서는 게이트 수가 모델의 연산 복잡도와 밀접한 관련이 있다. 모델의 매개변수와 연산 요구사항이 증가할수록 더 많은 게이트가 필요하며, 특히 모델의 연산 복잡도를 나타내는 FLOPs(Floating-Point Operations)는 게이트 수에 큰 영향을 미친다. 반면, 파라미터 수의 경우 SRAM 또는 DRAM에 저장할 수

있기 때문에 상대적으로 적은 영향을 미친다. 이 때문에 네트워크 경량화와 효율적인 회로 설계가 매우 중요한 과제로 떠오른다.

3. 전력 소비와 에너지 효율

밀리와트(mW) 또는 마이크로와트(μ W)로 표현되는 전력 소비는 딥러닝 모델 추론을 실행하는 데 필요한 총 전력을 나타낸다. 일반적으로 엣지 장치는 배터리로 작동하므로 초저전력 소비와 높은 에너지 효율이 필수적이다. 특히, 에너지 효율은 딥러닝 모델 추론 성능을 평가하는 핵심 기준 중 하나이다.

에너지 효율은 모델의 연산량과 메모리 접근 횟수에 큰 영향을 받으며, 특히 메모리 접근이 에너지 소비를 주도한다. 예를 들어, 45nm CMOS 기술에서 32비트 정수 덧셈 연산은 0.1pJ를 소모하는 반면, 32비트 SRAM 메모리 접근은 5pJ, 32비트 DRAM 메모리 접근은 640pJ를 소모한다. 이로 인해 대형 딥러닝 모델은 온칩 메모리 접근에 적합하지 않기 때문에 더 많은 에너지가 소모되는 DRAM 접근이 필요하게 된다. 만약, 100억 개의 연결 신경망을 20fps에서 실행하려면 (20Hz)(10G)(640pJ)=128W의 전력이 필요하며, 이는 일반적인 모바일 장치가 감당할 수 있는 전력 범위를 훨씬 초과하는 수치이다.

4. 추론 정확도와 견고성

추론 모델의 정확도는 엣지 디바이스가 할당된 작업을 얼마나 잘 수행하는지 보여준다. 특히, 자율 주행이나 얼굴 인식을 통한 장치 잠금 해제와 같은 보안 관련 작업은 매우 높은 정확도를 요구한다. 추론 정확도는 입력 데이터를 처리하는 속도에 따라 달라질 수 있다. 예를 들어, 비디오 프레임을 처리할 때 장치의 제약과 빠른 입력 속도로 인해 일부 프레임이 누락되어 정확도가 떨어질 수 있다.

또한, 모델의 견고성은 추론 성능 평가에서 매우 중요한 요소이다. 이상적인 데이터 세트와 시뮬레이션 환경에서 매우 정확한 모델일지라도, 실제 상황에서는 기대 이하의 성능을 보일 수 있다. 엣지 장치에서 추론 모델이 실

질적인 딥러닝 솔루션을 제공하기 위해서는 이러한 견고성이 필수적이다. 따라서 모델의 견고성은 다양한 데이터 세트를 통해 성능을 평가함으로써 정확하게 측정되어야 한다.

5. 지연 시간과 처리량

지연 시간은 모델이 완전한 추론을 얼마나 빠르게 수행할 수 있는지를 정의하는 개념이다. 이는 입력 데이터가 도착한 후 추론 결과가 생성되기까지 걸리는 시간으로, 주로 밀리초(ms) 또는 마이크로초(us) 단위로 측정된다. 낮은 지연 시간은 특히 엣지 추론에 중요한 요소이다. 예를 들어, 실시간 증강 현실(AR) 또는 가상 현실(VR) 애플리케이션과 지능형 로봇 비전과 같은 응용 분야에서는 지연 시간이 몇 밀리초 이내로 엄격하게 제한될 수 있다. 지연 시간은 엣지 장치의 종류, 가용 리소스, 추론 작업이 처리되는 방식 등 여러 요인에 따라 달라진다.

처리량은 추론 모델이 시간 단위당 처리할 수 있는 최대 입력 수를 의미하며, 종종 추론 가속기가 초당 수행할 수 있는 연산 수로 표현된다. 처리량이 높을수록 성능이 우수하다고 평가되며, 이를 정량화하기 위해 초당 기가 연산(GOPS) 또는 초당 테라 연산(TOPS)이 사용된다. 딥러닝 모델의 주요 연산량은 곱셈-누산(MAC)이므로, 처리량은 GMAC/s 또는 TMAC/s로 나타낼 수도 있다. 내비게이션, 의료 진단, 보안 및 자동화 같은 다양한 실시간 응용 분야에서는 높은 처리량이 필수적이다. 따라서 낮은 지연 시간과 높은 처리량은 실시간 성능을 위한 빠른 추론에 중요한 요소이다. 또한, 처리량은 칩의 코어 수와 밀접한 관련이 있으며, 코어 수는 성능 최적화를 위한 중요한 매개 변수로 고려해야 한다.

V. 결론

딥러닝 기술은 이미지 인식, 자율 주행, 의료 진단 등 다양한 분야에서 혁신을 이끌고 있으며, 탁월한 성과를 보여주고 있다. 하지만 딥러닝 모델의 높은 연산 요구와 자원 소모는 이를 효율적으로 배포하고 사용하는 데 있어 큰 도전 과제다. 이를 해결하기 위해 엣지 디바이스에서

딥러닝 모델을 실행하는 연구가 활발히 진행되고 있으며, 특히 TinyML과 같은 경량화 기술은 저전력 장치에서도 높은 성능을 유지하는 데 기여하고 있다.

딥러닝 모델 경량화의 핵심인 가지치기, 양자화, 호프만 코딩 등과 같은 알고리즘 기법 외에도, 하드웨어 측면에서도 다양한 경량화 기술이 활용되고 있다. 예를 들어, 저전력 프로세서 설계, 메모리 대역폭 최적화, 특수 목적의 AI 가속기와 같은 기술들은 엣지 디바이스의 효율성을 극대화하고, 에너지 소비를 줄이면서도 높은 성능을 유지하는 데 기여하고 있다. 이러한 하드웨어 최적화는 모델 압축 기법과 결합되어, 제한된 리소스에서도 고성능 딥러닝을 가능하게 한다.

결과적으로 엣지 디바이스에서 딥러닝을 수행하는 것은 비용 절감, 성능 향상, 보안 강화 측면에서 큰 이점을 제공한다. 앞으로 엣지 컴퓨팅과 TinyML 기술이 더욱 발전하면서, 딥러닝의 적용 범위는 한층 넓어질 것이며, 이는 더욱 스마트하고 지능화된 미래를 이끄는 중요한 기술적 토대로 자리잡을 것이다. 나아가, 이러한 연구 배경을 토대로 다양한 산업에서 큰 발전을 이룰 수 있기를 기대한다. 특히, 엣지 컴퓨팅과 AI 기술의 발전이 국가의 경쟁력을 높이는 데 기여할 것이며, 이에 대한 연구와 지원이 지속적으로 확대되어 산업 전반에서 혁신을 주도하는 기반이 되길 희망한다.

참고 문헌

- [1] M. M. H. Shuvo, et al, "Efficient Acceleration of Deep Learning Inference on Resource-Constrained Edge Devices: A Review", Proc. IEEE, vol. 111, no. 1, pp. 42-91, Jan. 2023.
- [2] J. Lin, et al, "Tiny Machine Learning: Progress and Futures [feature]", IEEE Circuits Syst. Mag., vol. 23, no. 3, pp. 8-34, Oct. 2023.
- [3] Y. Abadade, et al, "A Comprehensive Survey on TinyML", IEEE Access, vol. 11, pp. 96892-96922, 2023.
- [4] S. Han, et al, "Learning both weights and connections for efficient neural network", in Proc. Adv. Neural Inf. Process. Syst., 2015, pp. 1135-1143.
- [5] S. Han, et al, "Deep compression: Compressing deep neural

networks with pruning, trained quantization and Huffman coding", in Proc. Int. Conf. Learn. Represent. (ICLR), San Juan, Puerto Rico, May 2016, pp. 1–14.

- [6] J. Chen, et al. "Deep learning with edge computing: A review", Proc. IEEE, vol. 107, no. 8, pp. 1655–1674, Aug. 2019.
- [7] M. Shafique, et al. "TinyML: Current Progress, Research Challenges, and Future Roadmap", in Proc. 58th ACM/IEEE Design Autom. Conf. (DAC), Dec. 2021, pp. 1303–1306
- [8] J. Lin, et al. "MCUNet: Tiny deep learning on IoT devices", in Proc. Adv. Neural Inf. Process. Syst., 2020, pp. 11711–11722.
- [9] J. Garland, et al. "Low complexity multiply accumulate unit for weight-sharing convolutional neural networks", IEEE Comput. Archit. Lett., vol. 16, no. 2, pp. 132–135, Jul. 2017.
- [10] Y.-H. Chen, et al. "Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks", IEEE J. Solid-State Circuits, vol. 52, no. 1, pp. 127–138, Jan. 2017.



임형철

- 2021년 2월 숭실대학교 기계공학부
- 2021년 3월 ~ 현재 숭실대학교 지능형반도체학과

〈관심 분야〉

Vehicular Security, Artificial Intelligence, Automotive SoC, AI SoC



이주석

- 1983년 2월 서강대학교 전자공학과 학사
- 1985년 9월 고려대학교 전자공학과 석사
- 1999년 2월 고려대학교 전자공학과 박사
- 1984년 12월 ~ 1995년 3월 LG전자 선임연구원
- 1997년 3월 ~ 2000년 9월 용인송담대학 전자과 교수
- 2004년 6월 ~ 2006년 5월 (주)임택비전 연구소장
- 2007년 5월 ~ 2011년 9월 충북테크노파크 SoC 센터 센터장
- 2014년 2월 ~ 2018년 9월 (주)동일비전 연구소장
- 2022년 9월 ~ 현재 숭실대학교 IT대학 교수

〈관심 분야〉

AI SoC, Power Management SoC, Battery Management SoC



이성수

- 1991년 2월 서울대학교 전자공학과 학사
- 1993년 2월 서울대학교 전자공학과 석사
- 1998년 8월 서울대학교 전기공학부 박사
- 1998년 ~ 2000년 Research Associate, University of Tokyo
- 2000년 4월 ~ 2002년 8월 이화여자대학교 정보통신학과 연구교수
- 2002년 9월 ~ 현재 숭실대학교 전자정보공학부 교수
- 2016년 ~ 현재 한국전기전자학회 편집위원장, 수석부회장, 회장
- 2023년 ~ 현재 국가전략기술 반도체분야 조정위원장

〈관심 분야〉

Automotive SoC, AI SoC, Security SoC, Processor SoC, Battery Management SoC

플래시 메모리 기반의 인공지능 반도체

I. 서론

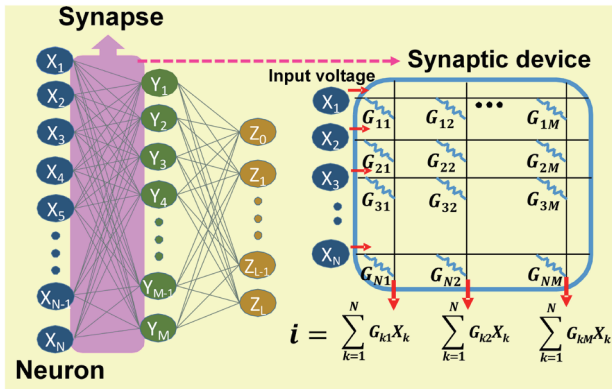
최근 인공지능(AI)은 자연어 처리, 음성 인식, 이미지 분류 등 다양한 인지 작업에서 상당한 발전을 보이고 있다. 그러나 고성능의 Deep neural networks (DNN)를 구현하기 위해서는 신경망 크기와 매개변수 수의 증가가 필요하며, 이로 인해 고성능 그래픽 처리 장치(GPU), 방대한 메모리 저장소, 그리고 높은 컴퓨팅 능력이 요구된다. 또한, 폰 노이만 병목 현상으로 인해 메모리와 프로세서 간의 대규모 데이터 전송이 발생하면서 Vector-matrix multiplication (VMM) 연산 시 많은 에너지와 시간이 소모된다.

뉴로모픽 시스템에서는 시냅스 소자 어레이를 활용하여 메모리 내 연산을 통해 폰 노이만 병목 현상을 해결하는 방안이 연구되고 있다. 입력 전압이 시냅스 소자 어레이에 인가되면, 전류는 시냅스 소자의 컨덕턴스와 입력 전압의 곱으로 결정된다(그림 1). 키르히호프의 전류 법칙(KCL)에 따라, 단일 소스 라인(SL)에 연결된 여러 소자에서 발생하는 전류가 합산된다. 어레이 내 각 SL에서 전류의 합산이 동시에 이루어지므로, 시냅스 소자 어레이는 하나의 펄스를 사용하여 병렬로 대규모 VMM을 수행할 수 있어 전통적인 폰 노이만 아키텍처에 비해 VMM 연산을 훨씬 효율적으로 처리할 수 있다.

최근까지 주로 저항성 랜덤 액세스 메모리(RRAM)와 상변화 메모리(PCM)를 뉴로모픽 시스템의 시냅스 소자로 사용해왔다. 하지만 RRAM 소자는 대규모 어레이 구현을 위한 selector, device variation, stochastic programming, reliability 등의 다양한 측면에서 개선될 필요가 있다. 또한, 금속 와이어에서 발생하는 IR 전압 강하는 RRAM 소자 어레이에서, VMM 연산의 정확도를 낮출 가능성이 있다. RRAM 소자의 제한된 on/off 전류 비율 역시 여러 시냅스 소자의 전류를 누적



이 성 태
홍익대학교



〈그림 1〉 뉴로모픽 시스템에 활용되는 시냅스 소자 어레이

합산하는 과정에서 오차를 일으킬 수 있는 잠재적 원인이 된다. 또한, PCM은 저항 드리프트 문제를 해결하기 위해 추가적인 연구가 필요하다. PCM의 비정질 칼코게나이드 물질에서 발생하는 시간 의존적인 저항 드리프트는 주요 신뢰성 문제 중 하나이며, 특히 다중 레벨 셀(MLC) 동작에 영향을 미친다.

이와 같은 문제를 해결하기 위해 NOR 플래시 메모리 및 SRAM과 같은 기술적으로 성숙하고 상업적으로 경쟁력 있는 실리콘 기반 소자를 시냅스 소자로 사용하는 방안이 있다. 그러나 NOR 플래시와 SRAM 모두 집적도 측면에서 제약이 있다. NOR 플래시 메모리의 경우, 단일 셀이 비교적 큰 면적을 차지하고, SRAM 셀은 여러 트랜지스터로 구성되기 때문이다. 반면, NAND 플래시 메모리는 그라운드 와이어와 비트 라인(BL) 컨택을 줄여 셀 집적도를 높인다.

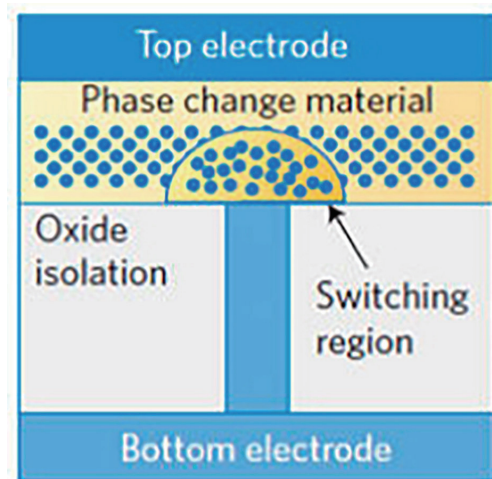
최신 DNN 알고리즘은 뛰어난 성능을 보이지만, 많은 수의 매개변수와 큰 네트워크 크기를 필요로 한다. 현재 NAND 플래시 메모리 기술은 대규모 데이터 저장 요구를 처리하는 데 가장 경쟁력 있는 솔루션 중 하나이다. 따라서 NAND 플래시 메모리 셀은 방대한 매개변수를 효과적으로 처리할 수 있는 능력을 가지고 있으며, 셀 집적도와 칩당 저장 용량 측면에서 큰 이점을 제공한다. 또한 NAND 플래시 메모리는 기술적으로 성숙하고 상업적으로 경쟁력 있는 기술이다. 최근 연구에서는 셀 스트링 구조에도 불구하고 NAND 플래시 메모리를 시냅스 소자로 활용하는 뉴로모픽 아키텍처를 제안했다. 본 문헌에서는

앞에서 간략히 소개한 시냅스 소자들을 소개하고 (II장), 낸드 플래시 메모리 기반의 시냅스 소자들에 대해 III장에서 소개한 후에, IV장에서 마무리한다.

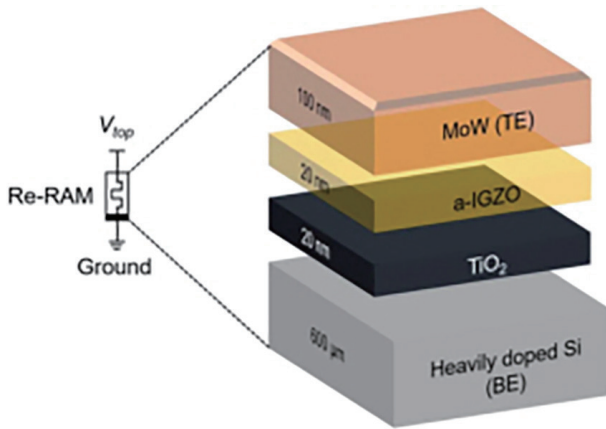
II. 시냅스로 사용되는 소자들

PCM은 물질의 상 변화에 따른 전기 저항의 큰 차이를 이용한다 〈그림 2〉. PCM의 상 변화는 비정질 상태(HRS)와 결정질 상태(LRS), 두 가지 상태로 나눌 수 있다. LRS 상태에서 HRS 상태로 '리셋'하기 위해서는 용융 급랭이 필요하며, 상변화 물질(GST)에 큰 전류를 가해야 한다. 뉴로모픽 시스템의 시냅스 소자로 사용할 경우 전력 소모 측면에서 리셋 전류를 줄여야 한다. 이러한 단점을 극복하기 위해 다양한 구조를 가진 PCM에 대한 여러 연구가 진행되었다. 저항이 높은 GST로 채워진 작은 부피의 공극이 전기 경로에 포함되면서, 이 제한된 구조는 기존 평면 셀 구조보다 더 큰 저항을 가진다. 또한 추가적인 하부 GST 층을 포함한 이중 GST 구조도 제작되었다^[1]. 제한된 GST 영역 아래에 추가 하부 GST 층을 삽입함으로써 열 손실을 효과적으로 방지할 수 있으며, 활성 영역의 온도 분포가 더 균일해진다.

멤리스터는 저항 스위칭을 기반으로 한 2단자 비휘발성 메모리 장치이다. 그 중에서도 RRAM은 금속-절연체-금속 구조를 갖는 소자로, 절연막 내 결함 체인을 기반으로 한 필라멘트 RRAM과 전기화학적 금속 필라멘트



〈그림 2〉 Phase change memory (PCM) 소자의 구조

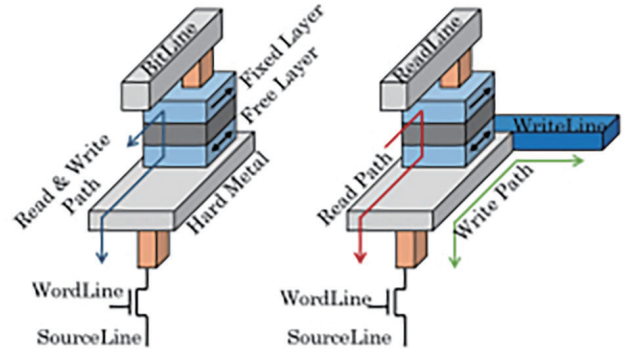


〈그림 3〉 Resistive random access memory (RRAM) 소자의 구조

형성에 기반한 CBRAM으로 나눌 수 있다 〈그림 3〉. 시냅스 장치로 사용되기 위해, 두 메모리 모두 간단한 구조, 저전력, CMOS 호환성, 고밀도 확장성 측면에서 매력적이다. 특히 크로스바 구조로 구성할 경우, RRAM 기반 컴퓨팅 시스템은 신경망의 벡터-행렬 곱셈(VMM)을 간단히 구현할 수 있다. 그러나 필라멘트 형성 과정이 본질적으로 급격하고 제어하기 어려워 신뢰성 문제와 큰 소자 산포 같은 문제가 있다. 또한 크로스바 구조는 셀렉터가 없을 경우 인접 셀을 통해 흐르는 불필요한 간섭 전류인 “sneak 전류”에 취약하다.

어레이 내 큰 소자 산포는 네트워크 성능 저하에 크게 기여하며, 표준 데이터셋인 MNIST와 CIFAR-10의 패턴 인식에서 낮은 인식 정확도를 초래한다. 특히 계산된 시냅스 가중치가 어레이에 매핑된 네트워크에서는 그 영향이 더욱 크다. 효과적인 아날로그 연산을 위해 필요한 다중 레벨 상태를 보장하려면 시냅스 어레이의 산포를 제어하는 것도 필수적이다.

스핀 기반 메모리인 스핀 전달 토크(STT)와 스핀 궤도 토크(SOT) 자기 저항 메모리(MRAM)는 고정 자기층, 터널층, 자유 자기층으로 구성된 자기 터널 접합(MTJ)을 기반으로 한다 〈그림 4〉. 스핀 기반 메모리의 저항 상태는 자유층의 스핀 자화 방향이 고정 자기층에 대해 어떤 방향을 가지는지에 따라 결정된다. RRAM과 PCM에 비해 읽기 및 쓰기 동작에서 높은 속도($\sim$ ns)와 낮은 에너지 소비($\sim$ pJ)의 특징을 가지기 때문에, 스핀 기반 메모리는 뉴로모픽 시스템의 시냅스 장치로 적합한 것으로 보고



〈그림 4〉 Spin Transfer Torque Magnetic Random Access Memory (STT-MRAM)과 Spin-orbit Torque Magnetic Random Access Memory (SOT-MRAM) 소자의 구조

되었다. 그러나 다중 레벨 상태를 위한 낮은 온/오프 저항 비율과 스핀 기반 장치의 확률적 스위칭 특성 때문에 한계가 있다. 이러한 단점을 극복하기 위해 다양한 메모리 셀 구조와 소재가 연구되고 있다. 여러 개의 수직으로 적층된 MTJ로 구성된 복합 스핀트로닉 소자가 다중 저항 상태를 구현하기 위한 시냅스 소자로 제안되었다^[2]. 이러한 MTJ는 CoFeB/MgO/CoFeB 박막으로 구성되었다. 제안된 복합 스핀트로닉 소자는 구성 요소의 소재 공학을 통해 설계 가능하고 안정적인 다중 저항 상태를 달성할 수 있다.

III. 낸드 플래시 메모리 기반의 시냅스 소자

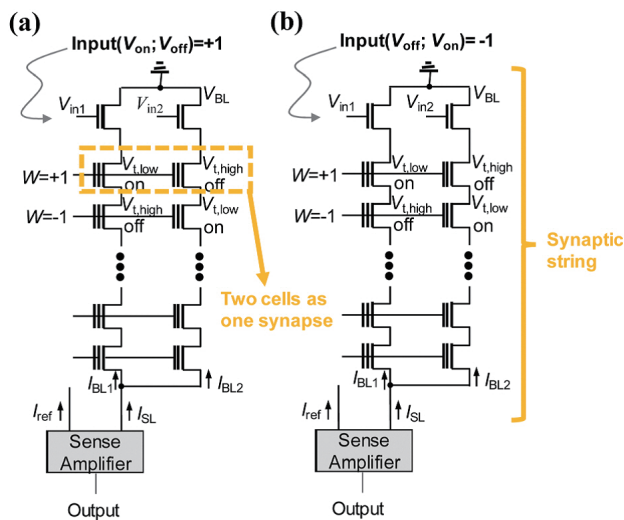
1. BNN의 시냅스로 사용

이진 신경망(BNN)은 가중치와 입력값을 1 또는 -1로 이진화하여 계산 오버헤드와 메모리 사용량을 크게 줄일 수 있다. 고정밀 부동소수점 곱셈 및 덧셈 (MAC) 연산 대신, BNN은 XNOR 연산과 비트 카운팅을 활용한다. 따라서 BNN은 하드웨어 신경망 구현에 유용한 해결책이 될 수 있다. 최근 연구에 따르면, BNN은 MNIST, CIFAR-10, ImageNet과 같은 잘 알려진 이미지 데이터셋에서 만족스러운 분류 정확도를 달성할 수 있음을 보여주었다.

[3]에서는 BNN에서 XNOR 연산을 수행하기 위해 2T2S(2개의 트랜지스터와 2개의 NAND 셀 스트링) 시

냅스 스트링 구조를 제안했다. <그림 5>는 XNOR 연산을 위한 2T2S(2개의 트랜지스터와 2개의 NAND 셀 스트링) 구조를 보여준다. 이 구조는 단일 시냅스 스트링을 위해 직렬로 연결된 NAND 셀들로 구성된 두 개의 NAND 셀 스트링을 사용하며, 두 입력 전압이 게이트에 적용되는 두 개의 입력 트랜지스터로 구성된다. 두 셀 스트링의 인접한 두 NAND 셀로 구성된 각 시냅스는 왼쪽 NAND 셀이 낮은 문턱 전압($V_{th,low}$)을, 오른쪽 NAND 셀이 높은 문턱 전압($V_{th,high}$)을 가질 때 +1의 시냅스 가중치를 나타낼 수 있다. 이 때, 반대로, 두 NAND 셀이 반대 상태를 가질 때 시냅스 가중치는 -1로 정의된다. 입력값의 경우, V_{in1} 과 V_{in2} 에 각각 턴온 전압(V_{on})과 턴오프 전압(V_{off})을 적용하여 +1을 정의한다. 반대로, -1의 입력값은 두 입력 전압의 반대 패턴으로 정의될 수 있다. 이러한 방법을 통해, XNOR 출력을 나타내는 스트링 전류(I_{SL})는 상보 입력 전압과 인접한 두 NAND 플래시 셀의 상태 조합에 따라 결정된다. <그림 5>의 (a)와 (b)는 각각 +1과 -1의 입력값일 때의 입력 전압을 보여준다.

제안된 방식에서, <그림 1(a)>에서 시냅스 스트링의 N 번째 행에 있는 NAND 셀은 N번째 시냅스 후 뉴런과 연결된다. 읽기 바이어스(V_{read})가 시냅스 스트링을 따라 워드 라인(WL)에 순차적으로 적용되면서, 각 시냅스 후 뉴런에 대한 출력이 순차적으로 생성된다. 이 방식은 모든 셀 스트링의 시냅스에 대해 전류 감지 회로(CSA)를

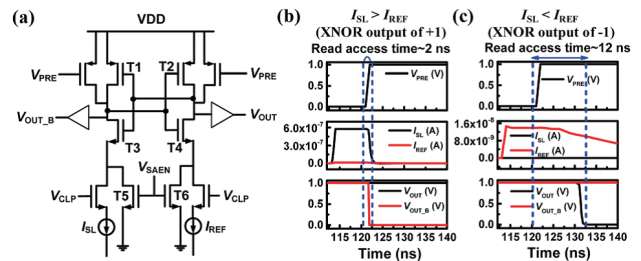


<그림 5> BNN에서 사용되는 2T2S의 구조 [3]

재사용할 수 있게 하여 회로 복잡성을 줄이고 집적도를 높여준다. 또한, 2T2S 설계는 디지털 도메인에서 작동하여, 일반적으로 아날로그 VMM에서 필수적인 대형 연산 증폭기나 아날로그-디지털 변환기(ADC)가 필요하지 않다.

<그림 6>의 (a)는 BNN을 위해 설계된 전류 래치 기반의 CSA 회로를 나타낸다. 이 회로는 ISL(두 개의 비트 라인 전류(IBL)로 구성됨)와 20nm FinFET 기반의 (BSIM-CMG)을 사용하여 시뮬레이션을 수행하였다. NAND 셀의 I_{on} 과 I_{off} 가 각각 590 nA와 0.1 pA일 때, <그림 6>의 (b)와 (c)는 각각 +1과 -1의 XNOR 출력에 대한 시뮬레이션 결과를 보여준다. CSA는 NAND 셀의 I_{on} 을 감지하여 XNOR 출력이 +1일 때 read access 시간이 2 ns가 된다. 반대로 XNOR 출력이 -1일 때 CSA는 IREF를 감지하며, 이로 인해 read access 시간이 12 ns로 연장되었다.

비트 오류의 감소를 위해^[4]에서는 BNN 내에서 XNOR 연산을 수행하기 위한 4T2S(네 개의 트랜지스터와 두 개의 NAND 셀 스트링) 시냅스 스트링 구조를 제안했다. 이 4T2S 구조는 NAND 스트링 내에서 differential sensing 방식을 포함한다. 각 시냅스 스트링은 두 개의 NAND 스트링으로 구성되며, 직렬로 연결된 시냅스 셀과 네 개의 입력 트랜지스터로 구성된다. 이 네 개의 입력 트랜지스터는 각각의 게이트에서 V_{in1} 과 V_{in2} 를 받는다. 이 네 개의 입력 트랜지스터가 단일 감지 증폭기와 결합되어 있으며, 이는 <그림 5>에 나타난 시냅스 스트링 및 감지 증폭기와 비교하여 시냅스 스트링 구조를 간소화한다는 점이 주목할 만하다.



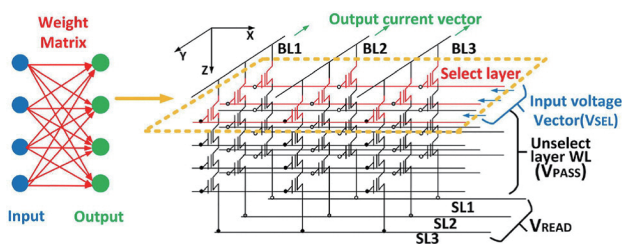
<그림 6> BNN에서 사용되는 Sense amplifier의 구조와 시뮬레이션 결과 [3]

2. QNN의 시냅스로 사용

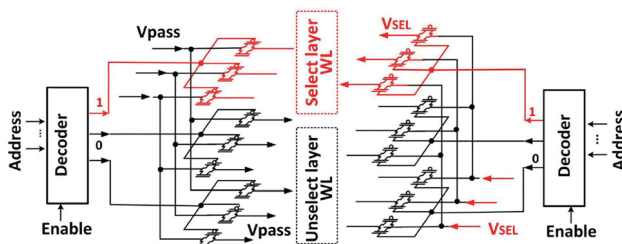
멀티비트 양자화 신경망(QNN)은 BNN보다 높은 추론 정확도를 달성할 수 있으며, 부동소수점 가중치를 가진 신경망과 비교해도 유사한 추론 정확도를 보여준다. [5]에서는 VMM 연산을 구현하기 위해 3D 수직 채널 NAND 어레이 아키텍처를 제안했다.

〈그림 7〉은 3D NAND 어레이에서 VMM 연산을 위한 바이어스 방식을 나타내며, 읽기 동작은 WL 층별로 수행된다. VMM 연산에서는 모든 SL에 동일한 Vread가 적용되고, BL은 접지된다. 선택되지 않은 WL 층에서는 WL에 패스 바이어스(VPASS)가 인가되어 선택되지 않은 트랜지스터가 임계 전압 V_{th} 값에 상관없이 패스 트랜지스터로 작동하도록 한다. 선택된 층의 WL에는 입력 벡터 패턴에 따라 선택 전압(VSEL) 또는 접지가 인가된다. 〈그림 8〉은 입력 전압 벡터와 VPASS를 인가하는 주변 회로를 보여준다. 두 개의 디코더 세트가 사용되며, 하나의 디코더는 선택된 층의 WL에 VSEL을 인가하고, 다른 디코더는 선택되지 않은 층의 WL에 VPASS를 인가한다. 각 스트링의 전류는 주로 선택된 층 내 NAND 셀의 컨덕턴스에 의해 결정된다.

이후, 모든 드레인 전류는 BL을 따라 집계되어 해당 열 내에서 가중합을 나타낸다. 이러한 아날로그 전류는 BL



〈그림 7〉 QNN에서 사용되는 낸드 플래시 메모리 구조 [5]

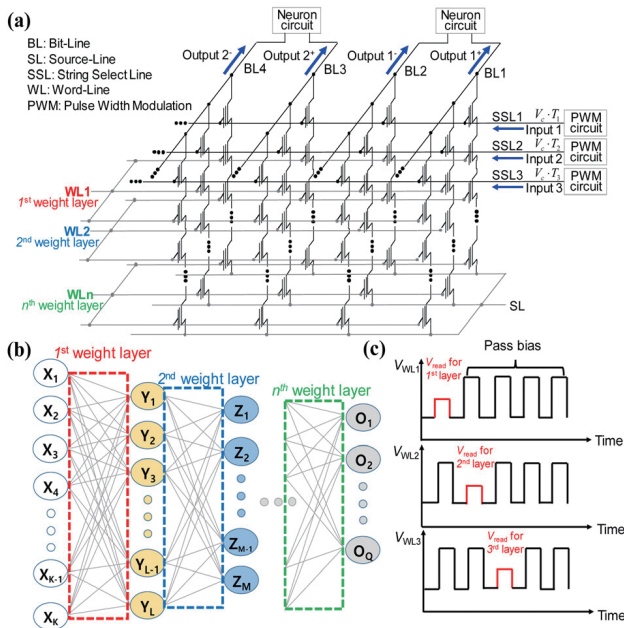


〈그림 8〉 낸드 플래시 메모리의 주변 회로 [5]

끝에 위치한 아날로그-디지털 변환기를 통해 감지될 수 있다. 본질적으로, 3D NAND 어레이에서 수행된 VMM 연산은 위에서 볼 때 2D NOR 플래시 배열에서 수행된 동일한 연산과 유사하게 보인다. 그러나 각 교차점에서 다른 선택되지 않은 층에 있는 패스 트랜지스터로 인한 직렬 채널 저항이 존재한다는 점을 유의할 필요가 있다. 이 아키텍처는 3D NAND 플래시 메모리를 활용하여 병렬 VMM을 구현할 수 있다. 그러나 각 층에서 WL이 분리되어 있어 기존 NAND 플래시 메모리 기술과는 차이가 있다.

[6]에서는 아날로그 입력과 4비트 가중치를 사용하는 NAND 플래시 메모리 아키텍처 기반의 고밀도, 높은 신뢰성을 가진 뉴로모픽 컴퓨팅을 위한 동작 방식을 제안했다. 〈그림 9〉는 PWM 회로를 사용하는 3차원 NAND 플래시 메모리를 활용한 뉴로모픽 시스템의 동작 방법도 식적으로 보여준다. PWM 회로는 조정 가능한 펄스 폭을 가진 입력 전압을 생성하여 SSL에 적용하며, 셀 전류는 〈그림 9〉의 (a)에 나타난 바와 같이 BL에서 더해진다. k th WL 내 NAND 셀은 〈그림 9〉의 (b)에서 신경망의 k th 시냅스 층 내 시냅스에 해당한다. 동작 방식은 〈그림 9〉의 (c)와 같이 선택된 WL과 선택되지 않은 WL에 각각 Vread와 VPASS를 적용하는 방식이다. 시냅스 스트링을 따라 WL에 Vread를 순차적으로 적용하여 각 시냅스 후 뉴런의 출력을 순차적으로 생성한다. 선택된 WL에 연결된 셀은 가중치를 저장하며, 각 가중치는 해당 셀 스트링의 전류를 결정한다.

제안된 동작 방식은 기존 NAND 플래시 메모리 아키텍처와 다르다는 점을 주목할 필요가 있다. 제안된 방식에서는 뉴런 활성화에 해당하는 입력 바이어스가 SSL에 적용되며, 전류 합산은 BL을 통해 이루어진다. 반면, 기존 NAND 플래시 메모리에서는 입력 주소로 선택된 셀이 BL을 통해 읽힌다. 또한, 제안된 방식에서는 입력 전압이 SSL에 동시에 인가되는 반면, 기존 NAND 플래시 메모리에서는 각 SSL에 읽기 바이어스가 순차적으로 인가된다. 이로 인해 이 방식은 기존 NAND 플래시 메모리 기술에 비해 지연 시간을 크게 줄인다.

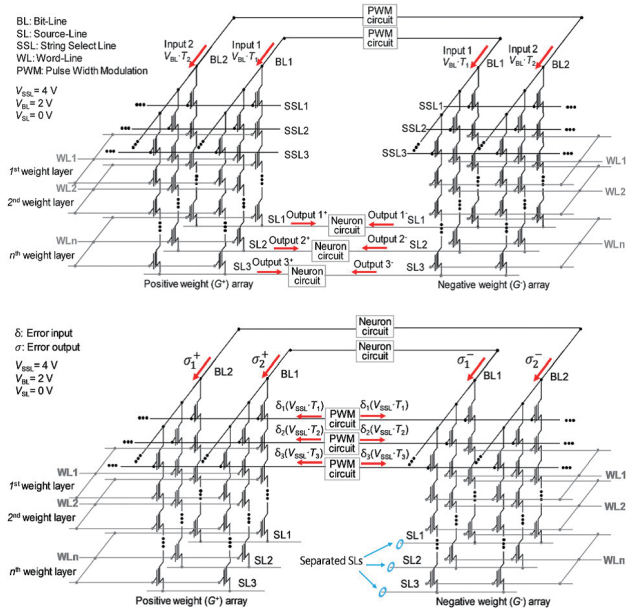


〈그림 9〉 QNN에서 PWM 회로와 같이 사용되는 시냅스 어레이 구조 [6]

3. 온칩 학습의 시냅스로 사용

온칩 학습은 오프칩 학습과 비교하여 학습 과정에서 소모되는 시간과 에너지를 상당히 줄일 수 있다. 또한 시냅스 소자의 산포의 효과로 인한 정확도 저하를 보정하고 실시간으로 변화하는 환경에 적응하는 장점이 있다. RRAM 크로스바 어레이의 경우, 순방향 패스에서는 행에 입력을 인가하고 역방향 패스에서는 열에 오류 입력을 인가하여 온칩 학습을 쉽게 구현할 수 있다. 그러나 NAND 플래시 셀을 시냅스 소자로 사용할 경우 셀 스트링 구조 때문에 이 방법을 사용할 수 없다. [7]에서는 NAND 플래시 메모리 어레이에서 순방향 전파(FP)와 역방향 전파(BP)를 모두 수행하여 온칩 학습을 가능하게 하는 새로운 방법을 제안했다.

〈그림 10〉은 NAND 플래시 메모리를 기반으로 한 제안된 시냅스 아키텍처에서 FP와 BP를 수행하여 온칩 학습을 가능하게 하는 동작을 각각 보여준다. 기존 NAND 플래시 메모리에서는 SL이 블록 내에서 연결되어 있어 BP 수행에 제한이 있다. 이러한 한계를 극복하기 위해, FP와 BP를 동일한 시냅스 배열에서 동작하도록 설계된 시냅스 아키텍처에서는 BL에 수직 방향으로 SL을 분리한다. 서로 다른 시냅스 가중치 어레이(G^+ 어레이, G^-



〈그림 10〉 온칩 학습을 위한 낸드 플래시 메모리 기반의 시냅스 아키텍처 [7]

어레이)에 위치한 두 개의 NAND 셀이 하나의 시냅스로 활용되어 음수 가중치를 나타낸다.

FP 동안에는 입력 바이어스가 BL에 인가되고, 각 가중합 전류는 분리된 SL에서 읽어낸다. BP에서는 오류 입력을 SSL에 적용하고, 각 가중합 전류는 BL에서 읽어낸다. RRAM 어레이에서 사용되는 방법처럼, 오류 입력이 SL에 적용되고 가중합 전류가 BL에서 읽어내는 경우, 셀 스트링 내 셀의 위치에 따라 셀 양쪽의 패스 셀의 저항이 다르기 때문에, FP와 BP에서 셀 전류가 달라질 수 있다.

IV. 전망과 결론

인공지능에 대한 관심이 높아지면서 컴퓨팅의 개념이 단순한 계산에서 사물과 사건을 인식하는 단계로 발전하고 있다. 이에 따라 인공지능의 적용 분야가 넓어지면서 소프트웨어뿐만 아니라 뉴로모픽 반도체도 주목받고 있다. 특히 저전력과 고속이 요구되는 인공지능 응용에서는 소프트웨어보다 반도체 칩을 통해 구현하는 것이 비용과 성능 면에서 훨씬 뛰어나다.

앞으로 뉴로모픽 반도체 연구는 클라우드에서는 인간이 해결하기 어려운 복잡한 문제를 풀어내는 슈퍼 인텔리전



스의 실현을 목표로 하며, 모바일에서는 인간의 능력을 크게 향상시켜 일종의 초능력을 부여하는 뉴로모픽 반도체로 발전할 것으로 예상된다. 이러한 맥락에서 AI는 단순히 ‘인공지능’이 아닌 ‘증강 지능’으로 해석될 수도 있다.

현재는 뉴로모픽 반도체 분야의 개념이 자리 잡고 다양한 연구자들이 연구와 개발에 박차를 가하기 시작한 중요한 시점이다. CPU에서 GPU로 이어진 발전 흐름에 따라 뉴로모픽 반도체가 주류가 될 것으로 기대하며, 한국에서 개발한 뉴로모픽 반도체가 세계적으로 널리 사용될 수 있는 기회가 되기를 바란다. 또한 이에 대한 연구와 지원이 더욱 확대되기를 희망한다.

참 고 문 헌

- [1] D. Chao, et al., VLSI Symp. Tech. Dig, 2007.
- [2] D. Zhang, et al., IEEE Trans. Biomed. Circuits Syst., vol. 10, no. 4, pp858–836, 2015.
- [3] S. T. Lee, H. Kim, J. H. Bae, H. Yoo, N. Y. Choi, D. Kwon, S. Lim, B. G. Park and J. H. Lee, High-density and highlyreliable binary neural networks using NAND flash memory cells as synaptic devices, IEEE Int. Electron Devices Meet., 2019, 34–38. IEEE.
- [4] S. T. Lee, H. Kim, H. Yoo, D. Kwon and J. H. Lee, Novel, parallel and differential synaptic architecture based on NAND flash memory for high-density and highly-reliable binary neural networks, Neurocomputing, 2022, 498, 1–13.
- [5] P. Wang, F. Xu, B. Wang, B. Gao, H. Wu, H. Qian and S. Yu, Three-dimensional NAND flash for vector-matrix multiplication, IEEE Trans. Very Large Scale Integr. (VLSI) Syst., 2019, 27(4), 988–991.
- [6] S. T. Lee and J. H. Lee, Neuromorphic computing using NAND flash memory architecture with pulse width modulation scheme, Front. Neurosci., 2020, 14, 571292.
- [7] S. T. Lee, G. Yeom, H. Yoo, H. S. Kim, S. Lim, J. H. Bae, B. G. Park and J. H. Lee, Novel method enabling forward and backward propagations in NAND flash memory for onchip learning, IEEE Trans. Electron Devices, 2021, 68(7), 3365–3370.



이 성 태

- 2016년 2월 서울대학교 전기정보공학부 학사
- 2021년 8월 서울대학교 전기정보공학부 박사
- 2021년 9월 ~ 2021년 9월 서울대학교 ISRC 연구원
- 2021년 10월 ~ 2022년 2월 Georgia Tech
박사후연구원
- 2022년 3월 ~ 2023년 2월 가천대학교 조교수
- 2023년 3월 ~ 현재 홍익대학교 조교수

〈관심 분야〉

Hardware-based Neural Network, Charge Storage Memory, NAND Flash Memory

반도체설계를 위한 AI/ML 기술

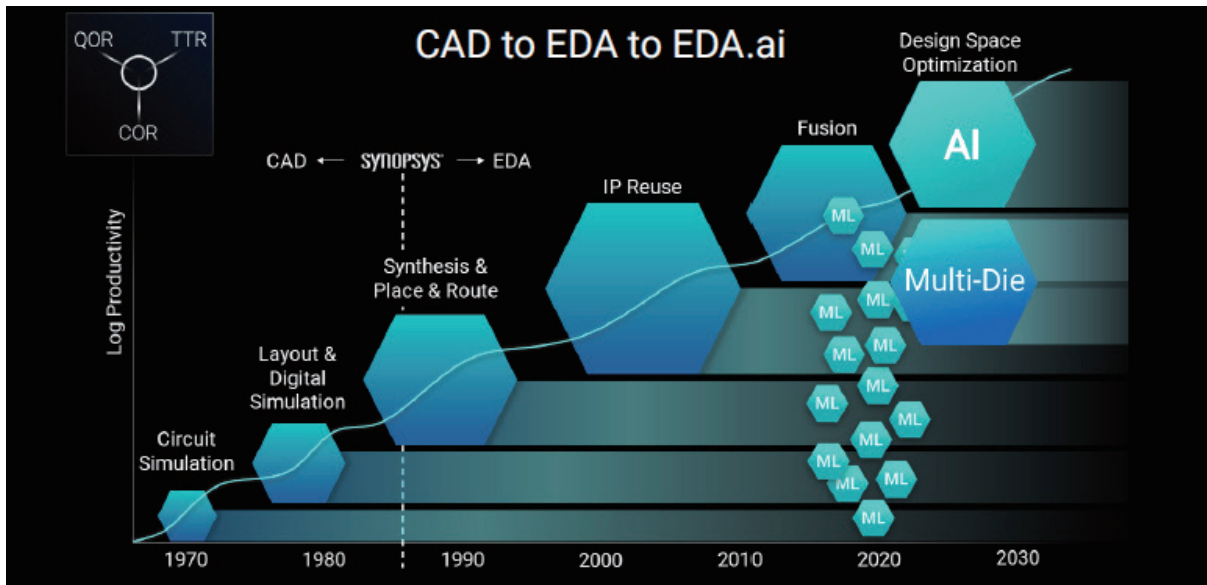
I. 서론

지난 반세기 반도체 공정/소자 기술의 급속한 발전은 범용 컴퓨터, 개인용 컴퓨터, 인터넷, 모바일기를 거쳐 인공지능(AI: Artificial Intelligence)으로 요약되는 현재의 정보기술 혁명을 가져왔다. 반면, 무어의 법칙(반도체 칩의 집적도가 1.5년에 2배씩 증가)으로 요약되는 반도체 공정/소자 기술의 급속한 발전에 대응하기 위한 반도체 설계 기술 역시 <그림 1>과 같이 발전을 거듭해 단순한 컴퓨터 이용 설계 (CAD: Computer-Aided Design)에서 반도체설계 자동화 (EDA: Electronic Design Automation)로 진보해 왔으나 공정/소자 기술 발전과 설계 기술 발전의 격차는 커지는 상황이다.

이러한 현재의 반도체설계 위기는 다음에 열거하는 다양한 원인에 효과적으로 대응하지 못한 한계에 의해 발생한다. (1) 공정/소자 기술의 혁신: 트랜지스터 구조가 2차원 Planar FET에서 3차원 FinFET으로 변화했고, 지속적으로 Nanosheet, CFET으로 변화 추진, 미세 Patterning 요구에 따른 Lithography 장비의 변화(ArF → EUV) 및 관련된 RET(Resolution Enhancement Technology) 기술의 변화, (2) 집적회로 구조 변화: 2차원에서 3차원, 특히 패키지 기술의 발전과 동반하여 Heterogeneous Integration 구조로 발전, (3) Technology Driver 다양화: 설계 기술 발전을 견인하는 대표 제품이 과거 CPU, AP(Application Processor) 등 한, 두 가지에서 이들 외에 AI, Automotive, IoT, AR/VR 등으로 다변화, (4) 설계 최적화 목표 복합화: 설계 최적화의 목표가 PPA(Performance, Power, Area) 최적화에서 이들 외에 설계 비용(Cost), 제품 수율(Yield) 등의 최적화로 다양화되고 있다. 이러한 사안에 효과적으로 대응하지 못한 이유로 현재까지의 반도체설계 자동화가 대부분 알고리즘에 의한 문제 해결에 의존한



최규명
서울대학교



〈그림 1〉 반도체설계 기술의 발전 (Source: Synopsys, 2023)

것을 주요 원인으로 지목할 수 있다.

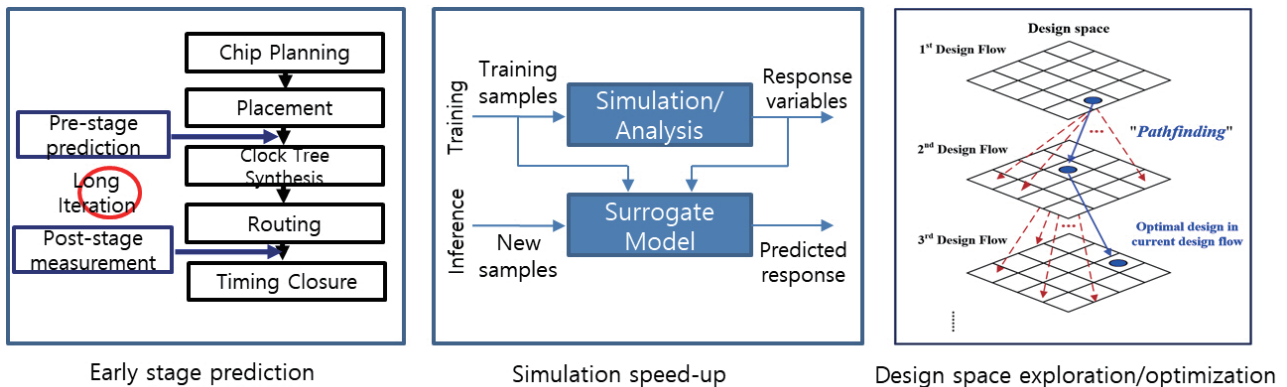
반면, 현재의 인공지능 기술을 견인한 동력인 기계학습(ML: Machine Learning) 기술은 데이터 분석에 의존하는 것으로 다음의 차별화된 특성에 의해 알고리즘으로 해결하지 못한 반도체설계 자동화 문제를 해결할 가능성을 가지고 있다. (1) 다양한 원인에 의해 요구되는 다양한 목적함수의 통합된 분석 가능, (2) 설계 매트릭(목적함수 값, 예: PPA 값)의 설계 초기 단계 예측 가능, (3) 대안 모델 설정으로 빠르고 정확한 시뮬레이션 및 분석 가능, (4) 빠른 설계영역 탐구 및 최적화 가능, (5) 알고리즘에 의한 해결로는 자동화가 불가능했던 아날로그 설계, 메모리 설계 등 순차적 설계 최적화 문제 자동화 가능. 단, 이는 기계학습 방식이 현재까지의 알고리즘 방식에 의한 반도체설계 자동화를 대체하는 것이 아니라 현재의 알고리즘 방식에 의한 반도체설계 자동화에 기계학습 방식을 접목한 새로운 방식이 현재의 반도체설계 위기를 해결할 방법으로 부상하게 되었다.

II. 반도체설계를 위한 AI/ML 기술

기계학습은 반도체설계뿐만 아니라 반도체 공정/소자 개발, 제조, 패키지/테스트 등 전 단계에 적용되거나, 적

용 준비 중이다. 공정/소자 개발 단계의 예측에 의한 측정 및 분석, 제조 단계의 수율 분석, 패키지/테스트 단계의 오류 분석 등이 대표적인 예이다. 반도체설계에는 시스템/아키텍처/RTL설계의 경우 최적의 설계영역 탐구/최적화 및 기능 검증에 적용 가능하고, 반도체 칩 구현의 경우 RTL2GDS 전 단계의 합성 및 검증에 적용 가능하다. 또한, 마스크 생성 단계의 검증, 모델링, 합성에도 적용 가능하다. 이렇듯 기계학습은 반도체설계뿐만 아니라 반도체 공정/소자 개발, 제조, 패키지/테스트 등 전반에 적용되고 있으나 여기서는 반도체설계 적용에 한정해 기술하고자 한다. 이러한 기계학습 적용은 적용 방식에 따라 (1) 알고리즘에 의한 자동화와 혼합되어 적용되는 방식 (Inside tool), (2) 알고리즘에 의한 자동화의 주변 또는 외부에서 적용되는 방식 (Outside/Around tool)로 나눌 수 있는데, (1)을 통해서는 PPA 개선, 설계 시간 단축 등 설계 자동화의 질을 향상할 수 있고, (2)를 통해서는 분석 및 의사결정에 조력, 설계영역 탐구 등으로 설계 자동화 완성도 재고 및 제품 경쟁력 향상을 꾀할 수 있다.

반도체설계에 적용 가능한 기계학습을 기계학습 분야에 따라 분류하면 지도학습 (Supervised Learning)에 의한 예측 및 분석, 대안모델 (Surrogate Model)에 의한 시뮬레이션/분석 시간 단축, 강화학습 (Reinforcement



〈그림 2〉 기계학습 분야별 반도체설계 자동화 적용 사례

Learning)에 의한 설계영역 탐구/최적화 및 순차적설계 자동화 등으로 나눌 수 있다. 〈그림 2〉는 기계학습의 대표적인 반도체설계 자동화 적용 사례이다. 단, 이는 응용 분야별 대표적인 기계학습 방안이나 실제적으로는 보다 다양한 기계학습 방안이 적용 될 수 있다. 아울러 최근 활발히 연구되고 있는 생성형 인공지능(Generative AI)의 반도체설계 자동화 적용 가능성 연구도 최근 시작되고 있다.

기계학습에 의한 반도체설계 자동화 기술 연구개발은 2016년 이래 선진 대학/ EDA기업/ 반도체기업에서 본격화되어 왔으나 실 설계 적용 측면에서는 아직도 적용 초기 단계라 말할 수 있다. 미국은 이 분야 연구개발에 가장 앞서 나가, 1990년대 말 이래로 거의 사라졌던 반도체설계 자동화 분야의 국가적인 연구지원이 재개되면서 CAEML(NFS: 2017.1~), MAGESTIC(DARPA: 2018.7~) 등을 위시한 대규모의 국가적인 연구지원 프로그램이 구성되어 (1) 반도체설계를 위해 특화된 기계학습 방법론 개발, (2) 일체화된 SOC(System On Chip), SIP(System In Package), PCB(Printed Circuit Board) 설계 플랫폼 개발을 위한 최신의 기계학습 기술 사용 등을 목표로 주요 대학과 EDA기업이 연구개발하고 있으며 반도체기업도 직간접으로 참여하고 있다. 중국의 경우 중국을 중심으로 홍콩, 대만, 미국 내 중국인 연구진이 유기적으로 연결되어 활발한 연구개발을 추진하고 있다. 미국에 비해 시작은 늦었으나 최근의 연구개발은 양적인 면에서 관련 학회에서 발표되는 논문의 양으로 미국을 능가하고 가장 앞서 나가는 상황이다. 특히, 최근의 미/중 간 반

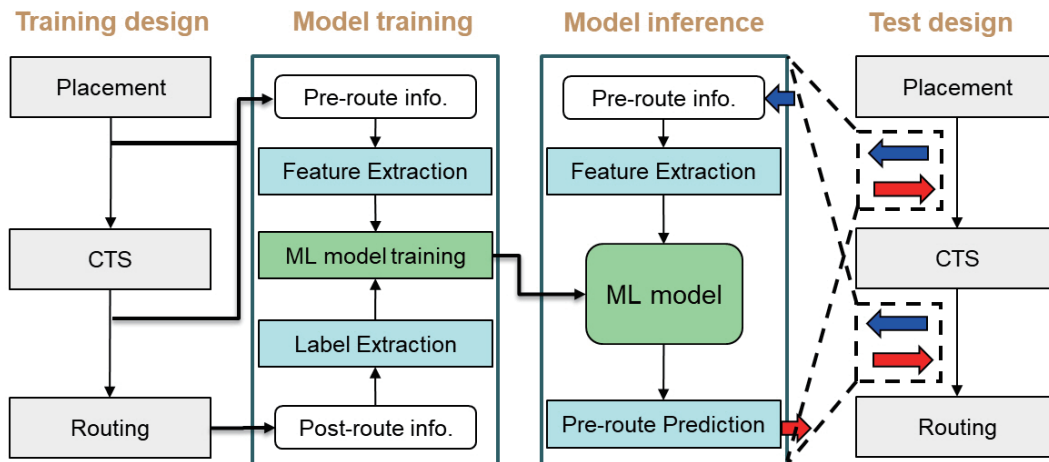
도체 전쟁에 선제적으로 대응하기 위해 중국 내 반도체설계를 위한 일괄된 자동화 시스템을 자체 개발하고 있으며 이의 일환으로 iEDA(an open-source infrastructure of EDA)^[1], iPD(an open-source intelligent physical design toolchain)^[2] 과제를 진행해 최근 1차 연구개발 결과를 발표한 바 있다.

한국의 경우 반도체설계 자동화 연구개발 활동이 미국 및 중국 대비 절대적으로 취약한데, 다행히 지능형 반도체 개발을 위한 국가 주도 사업의 일환으로 과학기술정보통신부 산하 정보통신기획평가원(IITP)이 지원하는 과제인 “지능형 반도체설계를 위한 소프트웨어 시스템” 개발이 추진되고 있다. 이는 KAIST 신영수 교수와 본인이 주도하여 국내 8개 대학 연구진 및 2개 EDA기업이 참여한 과제로 약 4년(2021년4월~2024년12월) 간 기계학습에 의한 반도체설계 자동화 기술을 연구개발하고 있다. 본 과제의 연구개발 활동에 대한 보다 자세한 사항은 KAIST 신영수 교수가 MLCAD 학회에서 소개한 과제 소개 논문^[3] 및 논문 내의 참고자료들을 참조하길 바란다.

III. 분야별 연구개발 사례

1. 지도학습에 의한 예측 및 분석

지도학습에 의한 예측 방법론은 반도체설계 자동화를 위한 다양한 분야에 적용을 위해 연구가 진행되어왔다. 특히 〈그림 2〉의 원편에 기술된 Physical Design(Placement & Routing) 단계는 자동화 완성



〈그림 3〉 Pre-Route 예측 및 결과 반영을 위한 Framework

도를 극대화하여 설계 기간의 단축을 이룩해야 하는 단계이나 반도체에 미세공정 적용이 심화 됨에 따라 Physical Design 기간은 점점 증가하는 추세에 있다. 이는 첨단 미세공정이 될수록 Gate(Cell)의 Pin을 연결하는 Interconnect Delay가 전체 Delay에서 차지하는 비중이 증가하여 Timing 맞추기가 어려워지고, 설계 규칙(Design Rule)이 점점 복잡해져 특히 Interconnect에서의 설계 규칙 오류(DRV: Design Rule Violation)가 증가함에 기인한다. 이의 해결을 위해서는 〈그림 2〉 Physical Design 후단에 있는 Routing 만을 최적화하여 달성되는 것이 아니라, 순차적 설계인 Physical Design 특성상 앞단에서부터 후단의 Routing을 고려하여 설계되어야 하는데, 앞단에서 고려가 제대로 되지 않은 경우 〈그림 2〉에서 지직한 Physical Design의 여러 단계를 반복적으로 수행하면서 최적화를 해야 한다. 특히 Routing 기간이 Physical Design 기간의 50~70%를 점유하는 최근의 미세공정 설계 시 Routing 전 단계인 Placement 또는 CTS(Clock Tree Synthesis)에서 Routing 후의 결과를 정확히 예측하는 방법론이 필요하다.

알고리즘에 의한 방법론이 주인 과거의 반도체 설계 자동화 Tool의 경우 Routing 후의 Interconnect를 Routing 전 단계(Pre-route)에서 정확하게 예측하는 것은 불가능하였다. 반면 Pre-route 단계에서 Routing에 영향을 주는 요소들을 입력 인자로 정의하고 Routing 후(Post-route) 결과를 Label로 하여 지도학습을 진행하여

기계학습 모델을 구성하고 이를 이용하여 Pre-route에서 예측하는 방법론 〈그림 3〉에 대한 다양한 연구개발이 최근 활발히 이루어지고 있다.

[4]의 연구에서는 Pre-route에서 Path Delay 예측을 정확하게 하기 위한 1단계로 Pre-route R(resistance)/C(capacitance)/Net congestion 예측, 이를 바탕으로 2 단계로 Arc(2-pin net) Delay/Arc Output Slew 예측을 유기적으로 하는 CNN과 ANN이 혼용된 계층적 예측 Model을 개발하여 궁극적으로 Pre-route Arc Delay와 Arc Output Slew 예측 정확도를 개선하였다. 특히, 예측 모델 만을 개발하는 타 그룹 연구와 달리 개선된 예측 결과를 상용화 P&R Tool(Synopsys, Cadence사 제품)에 접목하여 실제적인 Timing Closure 개선 및 반복적 설계 문제 감소를 가져오게 하였다.

[5]의 연구에서는 Pre-route에서 DRV를 정확하게 예측하는 방법론을 개발하였다. 미세 공정 반도체 구현설계에서 DRV의 원인은 기존 반도체 칩 내의 국소적인 배선 혼잡도(Global Routing Congestion) 외에 Cell의 Pin에 Interconnect Metal 연결 난이도 (Pin Accessibility) 2가지로 대변될 수 있는데, 이 연구에서는 Pin Accessibility 문제는 Pin Proximity Graph 모델링 후 GNN(Graph Neural Network) 적용, 배선 혼잡도 문제는 칩을 Grid로 나눈 후 Grid 내 배선 혼잡도에 영향을 주는 요소를 입력 인자로 정의하고 U-net 적용, 이후 이들을 ANN 구조로 연결하는 기계학습 구조를

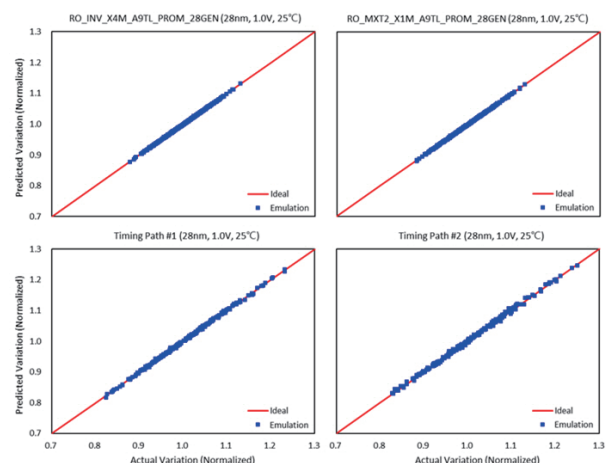
개발하였다. 아울러 [6]에서는 예측 결과 DRV 발생 원인에 따라 Pin Accessibility 문제에 연유된 DRV는 기계학습을 적용한 배치개선 툴을 개발하여 개선하였고, 배선 혼잡도 문제에 연유된 DRV는 상용화 P&R Tool의 Global Routing 조건을 조절하는 방법론을 개발하여 해결하였다.

한편 기계학습에 의한 분석 방법론은 설계 자동화 Tool 적용 결과를 정확히 분석하고 해결책을 제시하여 설계자의 주관적 분석에서 오는 오류를 크게 감소시킬 수 있게 하였다. [7]은 이 분야의 대표적인 연구로 반도체 신뢰성 문제인 EM(Electro-migration) 문제 해결에 기계학습에 의한 분석 방법론을 적용하였다. EM은 반도체 칩의 Metal에 과도한 전류가 흐르면 Metal에 균열이 생겨 Metal 저항이 증가하여 이로 인한 전기적인 문제가 발생하고, 심한 경우 Metal의 Open 또는 Short가 발생하는 문제이다. 이의 해결을 위해 설계 자동화 Tool 개발 기업인 Ansys는 Physical Design 후 EM 발생 여부를 판단하는 Tool을 개발 제공하고, 반도체설계 기업은 이를 이용 EM 발생 여부를 분석해 왔다. 문제는 과거의 설계 이력 관리 부실, 설계자의 숙련도 미흡에 따른 주관적 분석으로 분석 오류가 발생할 수 있고 이는 반도체 칩 오류를 유발한다는 것이었다. 반도체설계 기업인 Nvidia는 Ansys와 협력하여 이를 해결하는 기계학습에 의한 분석 방법론을 개발하였다. 개발된 분석 방법론에서는 과거의 EM Tool 수행 결과와 제조 후 실제 EM 발생 여부 data를 이용하여 반지도학습인 K-means Clustering 기법 및 지도학습인 KNN(K-Nearest Neighborhood) 기법을 이용한 분석 방법론을 개발 적용하여 EM 분석 오류를 감소시킬 수 있었다. 이러한 기계학습에 의한 분석 방법론은 다양한 설계 검증 Tool 결과 분석에 적용될 수 있는데, 특히 복수의 Tool 결과에 대한 종합적인 분석이 요구되는 ESD(Electro-Static Damage: 정전기에 의한 설계 오류) 분석 등에 효과적으로 적용될 수 있다.

2. 대안모델에 의한 시뮬레이션/분석 시간 단축

반도체설계의 검증 완성도를 높이기 위해서는 긴 시뮬레이션 시간 문제 해결이 필요하다. 예로 아날로그 설계

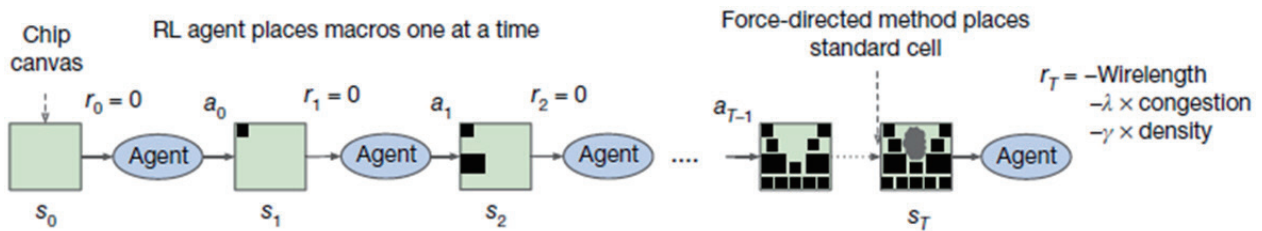
에는 회로 시뮬레이션이 필수적이며 다양한 공정 변화 조건을 고려 시 Monte Carlo 회로 시뮬레이션이 필요한데 수행 시간 문제로 원하는 전체 회로의 시뮬레이션이 현실적으로 불가능하여 회로의 많은 부분을 모델링하고 일부 분만을 시뮬레이션하는 방법을 사용하는데 이는 검증 완성도 저하를 가져오고 있다. 이의 해결을 위해 다양한 방법으로 대안모델 (Surrogate Model)을 만들고 이를 이용 시뮬레이션 수행 시간을 단축하는 방안이 개발되어 사용되고 있다. <그림 2>의 중간 그림은 이를 개념적으로 나타낸 것으로 예로 모든 경우에 대해 Monte Carlo 방식으로 회로 시뮬레이션 (Spice)을 수행하는 대신에 일부에 대해 수행한 후 Training Sample과 Response Variable 간 관계를 이용 지도학습 등의 방법으로 대안모델을 구축하고 나머지 경우는 대안모델에 의해 결과를 예측한다. EDA기업은 Monte Carlo 회로 시뮬레이션이 필요한 High-sigma 분석 분야 및 Cell Library 및 IP Characterization 분야에 이러한 방식을 적용하여 높은 검증 완성도를 상대적으로 단기간의 검증 수행 시간 내에 달성하는 Tool을 제공하고 있다. 이러한 High-sigma 분석에 적용되는 Stochastic Modeling을 위한 기계학습 방안은 UCLA의 Lei He 교수를 중심으로 Classification에 의한 방법^[8], Important Sampling에 의한 방법^[9] 등으로



500 training samples / 250 validation samples / Total 19+20 parameters
Average and standard deviation of relative errors (on average)

- RO surrogate models: -0.0333% / 0.0852%
- Timing path surrogate models: -0.0201% / 0.255%

<그림 4> 대안모델 정확도 테스트 결과



〈그림 5〉 Google의 칩 Floorplan을 위한 강화학습 적용 방안

나누어 많은 연구가 이루어지고 있으니 이를 참조하기 바란다.

[10]의 연구에서는 미세공정 및 문턱전압 부근의 저전력(NTV: Near Threshold Voltage) 동작에서 점점 심화되고 있는 공정 변이를 극복할 수 있는 설계방법론 개발을 추진하였는데, 구체적으로 기술하면 공정 변이 중 Die-to-Die Variation (공정 변이의 원인이 칩 간 공정 변이에서 연유된 것)을 구분하기 위해 임의의 Die(칩)의 Die-to-Die Variation을 HPM (Hardware Performance Monitoring) 방법론으로 찾아 칩 내 CP(critical path)의 정확한 Timing을 예측하고 Die-to-Die Variation 고려 시 필요한 마진을 줄임으로서 경쟁력 있는 칩을 생산하는 설계 방법론을 개발하였다. 방법론은 2단계 Flow로 되어 있는데, HPM2PV 단계에서는 HPM 회로의 delay 변화 측정을 통해 다양한 FET소자 (Front-end) 및 Metal layer (Back-end)의 PV(Process Variation) 변화를 통계적 방법으로 추출하고, PV2CPT 단계에서는 PV의 변화가 가져오는 회로 내 CP Timing 변화를 예측하게 되어 있다. 이를 위해서는 다양한 PV Parameter (〈그림 4〉 테스트에 적용된 PV Parameter는 Front-end 19개, Back-end 20개)의 변화를 고려한 Monte Carlo 회로 시뮬레이션이 필요하고 이는 오랜 수행 시간을 요구한다. [10]의 연구에서는 이의 해결을 위해 [11]에서 제안한 Tensor Recovery에 의한 불확실성 정량화 (Uncertainty Quantification) 방안을 도입하여 Monte-Carlo 회로 시뮬레이션을 대체하는 대안모델을 개발하였다. 〈그림 4〉는 대안모델에 의한 예측 결과를 Monte-Carlo 회로 시뮬레이션 결과와 비교한 것으로 대안모델에 의한 예측 결과가 Monte-Carlo 회로 시뮬레이션 대비 유사한 결과를 얻을 수 있음을 확인하

였다. 이 테스트에서는 10K Sample의 Monte-Carlo 회로 시뮬레이션 수행 시간 대비, [11]의 방안으로 추출된 750 Sample의 Monte-Carlo 회로 시뮬레이션 수행과 training에 의한 대안 모델 개발로 약 12.5배의 수행 시간 단축을 달성하였다.

3. 기계학습에 의한 설계 최적화

기계학습에 의한 설계 최적화는 알고리즘에 의한 자동화로는 좋은 결과를 가져오지 못했던 순차적 설계 최적화 문제 자동화와 빠른 설계영역 탐구에 의한 최적화 문제 해결을 위해 연구되고 있다.

Google은 대표적인 순차적 설계 최적화 문제인 칩 Floorplan(Macro-cell Placement) 자동화를 위해 강화학습을 적용한 방법론을 개발하였다^[12]. Floorplan은 과거 수작업에 의한 설계가 주를 이루었고 최근에는 Analytic Solver에 의한 방법이 적용되는 분야이다. Analytic Solver는 미분 가능한 cost function을 정의하고 이를 Gradient-based Optimization으로 해결하는 방법으로 결과의 우수성은 인정되나 상수의 Cost Function 대신에 미분 가능한 Cost Function을 정의하고 이를 Gradient-based Optimization으로 해결하여 과도한 수행 시간을 감수해야 한다. 〈그림 5〉는 Floorplan에 강화학습을 적용하는 방안으로 Cost Function을 Wirelength, Congestion, Density의 합(Weighted Sum)으로 설정하고 이를 Reward로 강화학습을 수행하여 Analytic Solver와 같은 Convex Approximation 없이 빠른 시간 내에 목적함수를 최적화하였다. Google은 이를 그들의 TPU Accelerator 설계에 적용하여 설계자의 수작업 및 Analytic Solver Tool 적용 대비 단기간에 우수한 결과를 가져왔다고 하였다. 그러

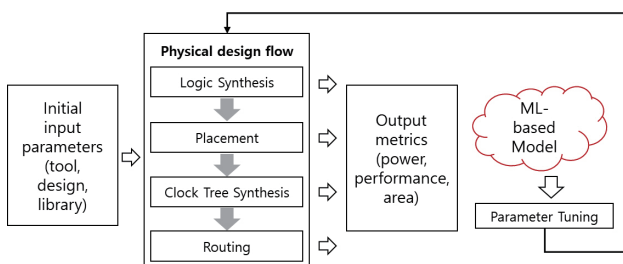
나 Google의 TPU Accelerator를 구성하는 Macro Cell은 일반적인 Floorplan의 Macro Cell보다 강한 동질성을 갖고 있으며, 이의 Training에 많은 유사 Data가 사용되었다는 점, Google 자체도 이 결과를 현재 제대로 이용하지 못하고 있다는 소문 등을 미루어 볼 때 강화학습에 의한 순차적 설계 최적화는 일반적으로 반도체기업에서 손쉽게 실 설계에 적용될 수 있는 단계에 있지는 않다.

설계영역 탐구에 의한 최적화 문제 해결을 위해 최근 기계학습 내의 설계영역 탐구/최적화 (DSE/DSO: Design Space Exploration/ Optimization) 기술을 적용하려는 연구가 상위레벨 설계 (시스템 /아키텍처 / RTL설계)와 RTL2GDS 설계 (Physical Design 설계) 분야별로 나누어 이루어지고 있다. 이들에 적용되는 기계학습 방법론은 개념적으로 유사하나 적용 영역의 특성에 맞추어 최적화되어야 하는 이슈가 있다. 여기서는 이 중 Physical Design에 적용되는 DSE/DSO 기술에 대해 설명한다. 반도체설계의 Physical Design (Logic Synthesis 및 Place & Route) 최적화를 위해서는 <그림 6>의 Physical Design Flow를 이루는 각 단계 수행 시 단계별 Tool, Design, Library 관련 입력변수 조절을 통해 최종적인 PPA 최적화를 달성해야 하나 탐구할 설계영역이 너무 광대하여(예로, Placement 단계에 국한한 관련 Tool, Design 입력변수 12개의 설계영역이 대략 4,860,000임) 설계영역 탐구에 의한 Physical Design 최적화가 불가능한 실정이었다. 반면, 최근 DSE/DSO 기술을 적용하려는 연구가 다양하게 이루어지고 있는데 이를 위한 쉬운 방안은 인공지능 분야 적용을 위해 개발된 DSO Tool(예, AutoTuner, AutoML)을 적용하는 것이나 이는 적용 영역 차이로 인해 일반적으로 만족할 만

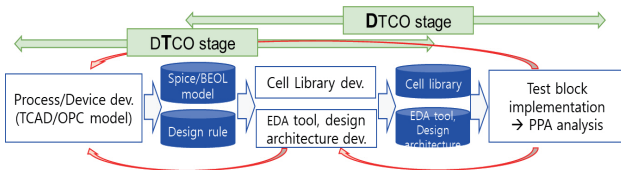
한 결과를 가져오지 못하고 있다. 이에 반도체 Physical Design 설계 자동화에 특화된 DSO Tool의 연구개발이 이루어지고 있다.

[13]의 연구에서는 Physical Design 영역에 특화된 대표적인 DSO Tool 선행 연구들의 구현을 통한 분석 결과를 바탕으로 보다 효과적인 DSO Tool을 개발하였다. DSO를 위한 DSE 과정은 Sampling 기법의 효율성에 따라 결과 성능 차이를 가져오는데 개략적인 Idea로 기계학습의 Sampling 기법 중 가장 우수한 결과를 가져오는 XGBoost^[14]을 DSO Tool의 Exploitation Engine으로 선정하고, 초기 Training Dataset에 따라 결과 성능 차이가 발생하는 XGBoost의 문제를 BO(Bayesian Optimization)^[15]를 Exploration Engine으로 사용하여 해결하였다. 또한 Physical Design의 최적화 대상이 되는 입력변수의 특성인 많은 변수와 변수 당 적은 Option을 BO가 효과적으로 Exploration 하기 위한 전 처리 과정으로 PCA(Principal Component Analysis)^[16]에 의한 차원 감소를 진행하였으며, XGBoost 적용 시 초기 반복수행에서 발생하는 예측 오차를 Anomaly Detection^[17]을 통해 해결하였다. 개발된 Tool은 적용 실험 결과 선행 연구 대비 우수한 최적화 개선 결과를 가져왔으며, 0.002% Sampling에 의한 입력변수 조절 DSO 수행으로 기본 입력변수 사용 시 대비 5~30% PPA 개선을 가져와 개발된 방법론의 유효성을 입증하였다.

반도체설계를 최적화하는 앞서 기술한 방법과 다르게 DTCO(Design Technology Co- Optimization)는 반도체 공정/소자와 설계 최적화를 동시에 고려하는 개념이다. <그림 7>은 DTCO의 개념을 나타낸 것으로 공정/소자 개발 결과인 Spice/BEOL Model/Design Rule, 이들을 이용 개발되는 Cell Library, Design 구조(PDN: Power Delivery Network, Metal Stack)의 결과 최적화는 Block-level PPA 최적화를 이루는 방향으로 동시에 최적화되어야 한다는 것이다. 특히 최근 미세공정 기술(7nm 이하)에서 Pitch Scaling에 의한 칩 집적도 향상이 한계에 직면한 상황에서 이러한 Design Scaling 개념이 포함된 DTCO는 칩 집적도 향상의 가장 중요한 요소가 되었고 이는 3차원 Heterogeneous Integration



〈그림 6〉 설계영역 탐구/최적화에 의한 Physical Design 최적화



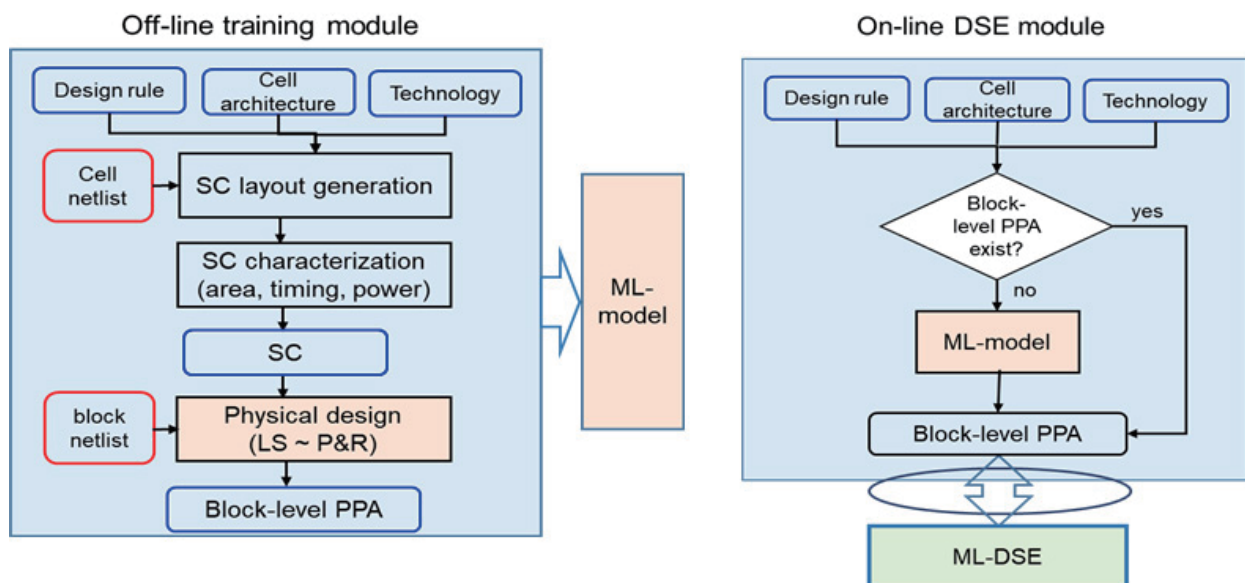
〈그림 7〉 DTCO 개념도

을 효과적으로 하는 공정/소자, 설계, 시스템의 동시 최적화를 목표로 하는 STCO(System Technology Co-Optimization)으로 개념 확대가 이루어지고 있다.

DTCO의 효과적인 수행을 위해서는 〈그림 7〉의 각 단계 자동화가 극대화되고, 이들에 다양한 조건의 입력을 넣었을 때 DSE/DSO가 효과적으로 이루어져야 한다. 특히 Design Centric DTCO (〈그림 7〉 flow 중 Design Rule, Spice/BEOL Model 등을 입력으로 Cell Library, Design Architecture 생성을 거쳐 Block-level PPA로 이들의 성능을 평가하는 DTCO) 수행을 위해서는 SC(standard Cell)의 자동생성 및 특성추출 그리고 이들의 DSE/DSO를 효과적으로 하는 Framework이 요구된다. 서울대에서는 Design Centric DTCO의 핵심 요소인 SC Layout 자동생성기 연구를 진행하여 첨단공정의 FET구조(FinFET, NanoSheet FET, CFET), 다양한 Cell 구조(Multi-row SC 포함), Design Rule,

Technology 조건에서 경쟁력 있는 SC Layout을 자동생성하고 특성추출이 가능한 Framework을 개발하고^{[18][19]}^[20], 이를 바탕으로 기계학습에 의해 실시간으로 Design-Centric DTCO 분석이 가능한 ML-assisted DTCO Framework을 개발하였다. 〈그림 8〉은 이를 나타낸 것으로, 개발된 Framework은 Cell Architecture, Design Rule, Technology Parameter의 변화에 따른 block-level PPA 분석을 실시간으로 수행하여 공정/소자 및 설계자의 DTCO 작업을 효과적으로 수행할 수 있게 한다. Framework은 Off-line Training Module을 통해 다양한 Cell Architecture, Design Rule, Technology Parameter 조건 입력 시 이에 맞는 SC 생성 그리고 생성된 SC들을 이용한 block-level 칩 구현 결과 PPA를 알아내는 과정을 통해 입력 Cell Architecture, Design Rule, Technology Parameter 조건에 따른 결과 PPA를 Label로 이들 간의 ML-model을 구축하고, DTCO 단계에서는 On-line DSE Module을 통해 실시간으로 공정/소자 및 설계자가 원하는 DTCO 분석이 가능하게 하였다. 이는 현재 다양한 DTCO Case Study를 통해 사용성 검증 단계에 있다.

IV. 기업의 연구개발 동향



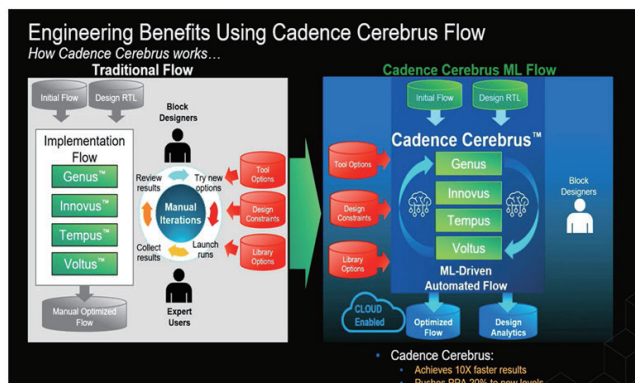
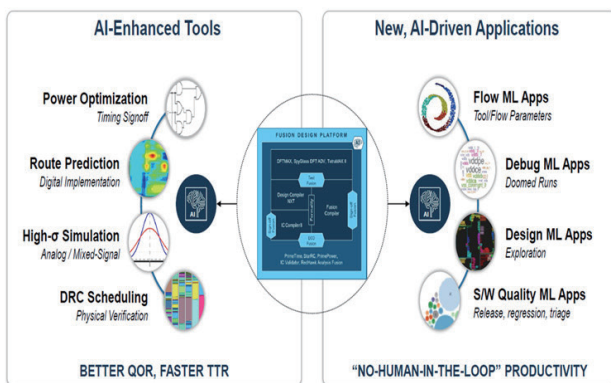
〈그림 8〉 ML-assisted DTCO Framework

EDA기업은 자신이 제공하는 Tool 및 Solution에 AI/ML 기술을 적용하여 성능을 보완하는 작업을 2018년 이후 발 빠르게 진행해 오고 있다. <그림 9>의 왼편 그림은 Synopsys의 AI/ML 기술 소개 자료로 그들이 제공하는 Tool을 AI로 개선하는 방식(AI-Enhanced Tools)은 II에서 언급한 Inside Tool에 해당하고, Flow/Solution을 AI로 자동화하는 방식(AI-Driven Application)은 II에서 언급한 Outside/Around Tool에 해당한다. 결론적으로 기존에 제공하는 Tool/Solution을 다양한 분야에서 AI/ML 기술로 보완했고 이는 현재 진행형이며 Cadence도 비슷하다. <그림 9>의 오른편 그림은 III-3에서 기술한 Physical Design DSO를 Cadence가 Tool화 한 것이고 Synopsys도 DSO.ai라는 비슷한 Solution이 준비되어 있다. 이렇듯 EDA기업은 최근 수년간 AI/ML에 의한 기존 Tool/Solution 보완에 그들의 역량을 집중하고 있는 것으로 보인다. 단, 실제 반도체기업에서 이를 얼마나 사용하고 있는지에 대해 다양한 각도로 조사한 바로는 아직 적용 초기 단계인 것으로 추측 된다.

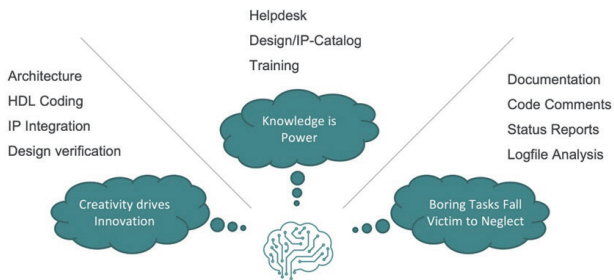
반도체기업은 설계환경 구축을 위해 1990년대 중반까지는 대기업의 경우(IBM, Intel, 삼성 등) 자체적인 EDA Tool 개발을 일정 부분 해 왔으나 이후 EDA Tool은 전문 EDA기업에서 대부분 도입해 왔다. 그러나 새로운 AI/ML 방식인, 알고리즘에 의한 현 EDA Tool(이는 EDA기업이 대부분 제공)에 Data Driven에 의한 AI/ML 기술을 접목하는 방식의 경우, 구축되는 설계환경의 우수성은 반도체설계 Data를 소유하고 있는 반도체기업의 능력에 따라 많은 차이를 가져올 것으로 예상된다. 실제로 세계적

인 일류 반도체기업은 AI/ML 기술로 반도체설계 환경을 개선시키는 내부 조직을 구성해 대처하고 있다.

Nvidia의 이 분야 전략은 CTO인 Bill Dally의 2021년 DAC Keynote에 나타나 있는데^[21], 핵심적인 EDA 알고리즘을 GPU와 AI/ML 기술로 보완하는 연구전략을 수립, 수행하고 있다. 첫째, EDA 알고리즘을 그들의 AI Accelerated Computing Platform으로 가속화 하는 것으로, 예로 Global Placement의 경우 최신의 방법론인 Analytic Solver 적용을 위해 그들의 GPU AI Hardware System과 AI SW Library (CUDA, PyTorch, TensorFlow로 이루어진 DL Tool kits)를 사용 자체적인 Solution을 구축하였다. 이러한 AI Accelerated Computing Platform으로 가속화는 Global Placement 외에 Logic Simulation 분야에도 적용하고 있다. 둘째, 앞서 III-1에서 기술한 지도학습에 의한 예측 및 분석 방법론을 자체 개발하는 것으로 AI를 위해 개발한 최신의 NN(neural network) Model을 EDA 알고리즘에 적용하는 것으로 방법론 개발은 앞서 언급한 대학의 연구를 앞서가지 않으나 실제 Data를 가지고 실 설계에 적용한다는 면에서 강점을 갖는다. 셋째, 강화학습 등의 AI/ML 기술을 이용하여 새로운 최적화 EDA 알고리즘을 개발하는 것이다. 이 분야 연구가 과거의 최적화 알고리즘인 Simulated Annealing, Genetic 알고리즘 대비 우수성을 보이기 위해서는 양적, 질적인 면에서 충분한 Data가 필요한데 대학의 경우 이 점이 여의치 않아 이 분야 연구는 많은 경우 형식적으로 마무리될 가능성이 있다. 반면 한 분야의 개발에 오랜 기술 자산을 가지고 있는 반도체



<그림 9> 왼편: Synopsys AI/ML 기술, 오른편: Cadence's Cerebrus Physical Design



〈그림 10〉 생성형 AI 기술의 반도체설계 적용 예

기업의 경우 이의 해결이 가능할 것인데, Nvidia는 1차로 Standard Cell Layout 생성 분야에 집중하여 이 분야 연구개발을 하고 있다.

V. 생성형 AI

최근 LLM(Large Language Model)에 의한 생성형 AI(Generative AI) 기술 발전은 AI 관련 분야 기업 생태계를 바꿀 정도로 급속도로 변화하고 있으며, 이의 반도체설계 분야 적용을 위한 연구도 초기 단계이나 시작되고 있다. 반도체설계에 생성형 AI 기술은 〈그림 10〉^[22] 나타낸 바와 같이 (1) HDL(예, Verilog RTL) Code 생성 및 검증 등의 자동화 영역, (2) 존재하는 반도체 칩 및 IP 설계 정보 등을 제공하는 일종의 Helpdesk 영역, (3) Documentation과 Code Comments 등 설계자의 단순 업무 대행 영역에 활용을 목표로 1차적으로 연구개발 되고 있다. EDA기업들은 발 빠르게 관련 기술개발 결과를 홍보하고 있어 Synopsys는 Microsoft의 Azure OpenAI Service를 이용하여 Synopsys.ai Copilot이라는 Framework를 구축하였고, Cadence도 이와 유사한 Cadence.AI를 구축한 것으로 발표하고 있다. 그러나 이들은 상기 기술한 (2), (3)에 해당하는 것으로 EDA기업이 그들이 제공하는 Tool과 Solution에 대한 내용을 고객에 제공하는 것이고, 반도체기업이 (2), (3)을 구축하고자 하는 경우 이를 Framework으로 활용할 수 있으나 이외에 반도체기업 자체 인력에 의한 많은 사전 준비 작업이 필요할 것으로 생각된다.

한편, 상기 기술된 (1)은 현재까지의 반도체설계 자동화 기술로는 해결되지 않은 부분이다. 참고로 TLM

(Transaction-Level Model) 레벨에서 SystemC, SystemVerilog 등으로 반도체설계를 위한 Coding을 하고 Verilog RTL을 자동 생성하는 High-level Synthesis (Architectural Synthesis) 기술 관련 오랜 연구가 있었고 현재 부분적으로 적용되고 있으나 반도체설계에 전면 적용되지는 않고 있다. 반면, 생성형 AI 기술로 (1)을 해결하고자 하는 연구가 최근에 대학 및 기업에서 이루어지고 있는데 현재까지 생성되는 결과는 실사용이 불가한 수준으로 향후 많은 추가적인 연구가 필요하여 생성형 AI 기술의 반도체설계 적용은 이제 시작 단계로 볼 수 있고, 앞으로 많은 발전이 필요하다고 볼 수 있다.

VI. 결론, 문제점 및 제언

최근 반도체설계를 위한 AI/ML 기술 연구개발이 활발히 진행되고 있으며 이는 반도체설계의 오랜 꿈인 “Silicon Compilation” (반도체설계 사양을 기술하면 반도체 칩 제조의 입력인 마스크 Data가 자동으로 생성되는 반도체설계의 완전 자동화)의 미래에 실현 가능성을 높여주고 있다. 그러나 이는 현재의 알고리즘 방식에 의한 반도체설계 자동화를 AI/ML 기술로 대체하는 것이 아니라 현재의 반도체설계 자동화에 AI/ML 기술이 접목된 새로운 방식으로 가능할 것이다.

한편, 지난 10년 남짓한 기간 동안 대학, EDA기업, 반도체기업에서 진행된 많은 연구에 비해 실제 반도체설계에 AI/ML 기술 적용은 아직 초기 단계이다. 원인으로 반도체설계 자동화에 AI/ML 기술 적용을 위해서는 반도체설계에 필요한 다양한 Data(예: 공정/소자, 회로, 패키지, 시스템 Data)에 AI/ML 기술이 적용되어야 하는데 이들을 소유하고 있는 주체가 상이(Foundry, Fabless 설계기업)하고 이들 Data는 기업의 민감한 기술 정보를 포함하고 있어 극도의 보안으로 관리되므로 AI/ML 기술에 의한 반도체설계 자동화를 연구개발하는 대학, EDA기업과 이들 반도체기업 간 공유하는 것이 용이하지 않다. 또한, 이들 반도체 관련 Data는 반도체공정/소자 기술, 다양한 단계의 반도체 모델링 기술, 반도체설계 기술, EDA 기술 발전에 따라 지속적으로 변화되고 있어 기계학습이

필요로 하는 충분한 양과 우수한 질의 Data를 지속적으로 유지보수 하는 것이 쉽지 않다. 따라서 대학, EDA기업, 반도체기업(Foundry, Fabless 설계기업) 간 유기적인 협력 체제 구축이 요구되어 진다. 구체적으로 국내 상황을 기준으로 제언하면 삼성 등 대규모 선진 반도체기업은 자체 AI/ML 기술 연구개발 조직으로 자체 개발만 아니라 EDA기업 및 대학과의 협력을 리드할 수 있어야 하고, 다수의 Fabless 벤처기업을 위해서는 AI/ML 기술의 반도체설계 적용을 위한 관련 Data 및 기술 공유체제를 국가/사회적 차원에서 구축해야 한다고 생각한다.

참고 문헌

- [1] Xingquan Li et al., "iEDA: An Open-source infrastructure of EDA", ASPDAC 2024.
- [2] Xingquan Li et al., "iPD: An Open-source Physical Design Toolchain", ASPDAC 2024.
- [3] Youngsoo Shin, "AI-EDA: Toward a Holistic Approach to AI-Powered EDA", MLCAD 2023.
- [4] Kyungjoon Chang et al., "DTC: integrating Deep-learning driven Timing Optimization into the state-of-art Commercial EDA tool," DATE 2023.
- [5] Kyeonghyeon Baek et al., "Pin Accessibility and Routing Congestion Aware DRC Hotspot Prediction using Graph Neural Network and U-Net", ICCAD 2022
- [6] Suwan Kim et al., "Methodology of Resolving Design Rule Checking Violations Coupled with Fully Compatible Prediction Model", ISPD 2024.
- [7] Normal Chang et al., "Machine learning based generic violation waiver system with application on electromigration sign-off", ASPDAC 2018.
- [8] Wei Wu et al., "REscope: High-dimensional statistical Circuit Simulation towards Full Failure Region Coverage", DAC 2014.
- [9] Xiao Shi et al., "A Fast and Robust Failure Analysis of Memory Circuits Using Adaptive Importance Sampling Method", DAC 2018.
- [10] Jeongwoo Heo et al., "Hardware Performance Monitoring Methodology at Near-Threshold Computing and Advanced Technology Nodes: From Design to Postsilicon", Trans-on CAD 2022.
- [11] Z. Zhang et al., "Big-data tensor recovery for high-dimensional uncertainty quantification of process variations", Trans-on components, packaging, manufacturing technology 2017.
- [12] Azalia Mirhoseini et al., "A graph placement methodology for fast chip design", Nature, 10 June 2021.
- [13] Chanhee Jeon et al., "BOXGB: Design Parameter Optimization with Systematic Integration of Bayesian Optimization and XGBoost", DATE 2024.
- [14] T. Chen et al., "XGBoost: A scalable tree boosting system", in KDD, 2016, pp.785-794.
- [15] J. Bergstra et al., "Algorithms for hyper-parameter optimization", Advances in neural information processing system, vol. 24, 2011.
- [16] R. Bro et al., "Principal component analysis", Analytical methods, vol.6, no.9, pp.2812-2831, 2014.
- [17] J. Henriques et al., "Combining k-means and xgboost models for anomaly detection using log datasets", Electronics, 2020.
- [18] Kyeonghyeon Baek et al., "Simultaneous transistor folding and placement in standard cell layout synthesis", ICCAD 2021.
- [19] Handong Cho et al., "Standard Cell Layout Generator Amendable to Design Technology Co-Optimization in Advanced Process Nodes", DATE 2024.
- [20] Sehyeon Chung et al., "Optimal Layout Synthesis of Multi-Row Standard Cells for Advanced Technology Nodes", ICCAD 20204.
- [21] Bill Dally, "GPUs, Machine Learning, and EDA", Keynote at DAC 2021.
- [22] Hans Bouwmeester, "Fully automating chip design", @ <https://semiwiki.com/eda/345291-fully-automating-chip-design/>



최규명

- 1983년 2월 한양대학교 전자공학과 학사
- 1985년 2월 한양대학교 전자공학과 석사
- 1995년 10월 미국 피츠버그대학교 전기컴퓨터공학과 박사
- 1985년 1월 ~ 2004년 12월 삼성전자 반도체부문 연구원~수석연구원
- 2005년 1월 ~ 2010년 12월 삼성전자 S.LSI사업부 설계기술팀장, 연구위원 (상무)
- 2011년 1월 ~ 2013년 12월 삼성전자 S.LSI사업부 기반설계센터장, 연구위원 (전무)
- 2014년 1월 ~ 2015년 12월 삼성전자 S.LSI사업부 LSI 개발실장, 연구위원(전무)
- 2016년 3월 ~ 2024년 8월 서울대학교 전기정보공학부 객원/ 산학협력중점교수
- 2024년 9월 ~ 현재 서울대학교 공학전문대학원 책임연구원

〈관심 분야〉

반도체설계 자동화, 인공지능/머신러닝에 의한 반도체설계 자동화, 반도체 설계/공정/시스템 동시 최적화

통신 분야

[통신]

- IRS 탑재 UAV를 활용한 ISAC 시스템에서 공동 빔포밍 기법
문상미, 황인태

[방송통신융합기술]

- 3차원 포인트 클라우드의 저지연 스트리밍을 위한 최적 전송 데이터 생성 방식에 관한 연구
박소민, 최민지, 유철우

반도체 분야

[반도체 소자 및 재료]

- 양성자 조사에 의한 β -Ga₂O₃ 쇼트키 배리어 다이오드의 변위 손상 효과
김영조, 김형우, 윤영준, 서재화
- 차세대 금속-나노탄소 도금기반 복합시트 열전도도 향상 연구
최소영, 이철승, 배준호

[SoC 설계]

- Questa를 이용한 Code Coverage 검증 방법
엄유진, 양희훈, 심용기, 배득우, 오민식, 유호영
- 고속 FHD 이미지 스티칭을 위한 실시간 Mapping 회로 설계
이효연, 조정순

컴퓨터 분야

[융합컴퓨팅]

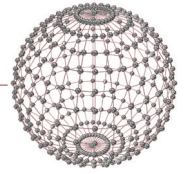
- VDO(Virtual Data Optimizer) 기반 중복제거 파일 시스템의 성능 분석
권진영, 박정민, 노재춘

[인공지능, 신경망 및 퍼지시스템]

- 순차 확률비 검정과 정규화 흐름을 사용한 시계열 데이터 조기 분류
조준희, 최계원

[인공지능 및 보안]

- 멀티 헤드 어텐션의 재설계와 헤드 믹싱 기법으로 인한 메모리 및 연산량 감소
김규동, 김진현, 나혁주, 장현성, 박재민, 황재기, 하남구, 김승룡, 김영근
- 수면 소리 데이터를 활용한 1D ResNet 딥러닝 모델 기반 수면무호흡증 탐지 연구
박태원, 윤승원, 윤혜원, 이규철



인공지능 신호처리 분야

[영상 신호처리]

- 금속활자본 내 동일 글자에 대한 활자별 분류
박채호, 최강선
- 강건한 자기주도 학습 기반 다중 프레임 깊이 추정: 코스트 볼륨 대체제를 찾아서
김진현, 김규동, 나혁주, 장현성, 박재민, 황재기, 하남구, 김영근, 김승룡

[음향 및 신호처리]

- 재난 현장 인명 구조를 위한 딥러닝 기반 잡음 제거 및 음향 분류
김준휘, 최준규, 임성빈

[영상 이해]

- 초분광 얼굴 데이터셋 증강을 통한 얼굴 인식 시스템 성능 향상
최영인, 강주성, 이지오, 신영학, 이흥노

시스템 및 제어 분야

[회로 및 시스템]

- 선형 위상 특성을 갖는 필터 함수 근사를 이용한 준 가우시안 펄스 성형기의 설계 및 제작
채규성

산업전자 분야

[신호처리 및 시스템]

- 레이더 신호처리 장치 검증 및 분석을 위한 자동 분석 프로그램 구현
김윤지, 송우진, 송준호

[컴퓨터 응용]

- 품질인식 기반 센서 데이터 관리 시스템 설계 및 구현
노태현, 이민구, 박용국, 정경권

국 내 외 학 술 행 사 안 내

국·내외에서 개최되는 각종 학술대회/전시회를 소개합니다.
 게재를 희망하시는 분은 간략한 학술대회 정보를 이메일로 보내주시면 게재하겠습니다.
 연락처: ieie@theieie.org

>>2025년 1월

일자	학술대회명	개최장소	홈페이지/연락처
01.07 - 01.09.	2025 IEEE Power Electronics Society International Decentralized Energy Access Symposium (IDEAS)	Bali, Indonesia	http://ideas-2024.ieee.org/
01.09. - 01.10.	2025 Fifth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT)	Bhilai, India	https://icaect.com/
01.10. - 01.13.	2025 IEEE 22nd Consumer Communications & Networking Conference (CCNC)	Las Vegas, Nevada, USA	https://ccnc2025.ieee-ccnc.org/
01.11. - 01.14.	2025 IEEE International Conference on Consumer Electronics (ICCE)	Las Vegas, Nevada, USA	https://icce.org/2025/
01.13. - 01.17.	2025 Joint MMM-Intermag Conference (INTERMAG)	New Orleans, Louisiana, USA	https://2025-joint.magnetism.org/
01.15. - 01.17.	2025 Fourth International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)	Bangkok, Thailand	https://ica-symp-2025.ecti-thailand.org/
01.16. - 01.17.	2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET)	Pune, India	http://cumminscollege.org/icaet/
01.16. - 01.17.	2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE)	Bangalore, India	https://www.iciitcee.in/
01.19. - 01.23.	2025 IEEE 38th International Conference on Micro Electro Mechanical Systems (MEMS)	Kaohsiung, Taiwan	https://www.mems25.org/
01.19. - 01.22.	2025 IEEE Radio and Wireless Symposium (RWS)	San Juan, Puerto Rico, USA	https://www.radiowirelessweek.org/
01.19. - 01.22.	2025 104th ARFTG Microwave Measurement Conference (ARFTG)	San Juan, Puerto Rico, USA	https://arftg.org/
01.19. - 01.22.	2025 IEEE 24th Topical Meeting on Silicon Monolithic Integrated Circuits in RF Systems (SiRF)	San Juan, Puerto Rico, USA	https://www.radiowirelessweek.org/
01.19. - 01.22.	2025 IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications (PAWR)	San Juan, Puerto Rico, USA	https://www.radiowirelessweek.org/
01.19. - 01.22.	2025 IEEE Space Hardware Radio Conference (SHaRC)	San Juan, Puerto Rico, USA	https://www.radiowirelessweek.org/
01.19. - 01.22.	2025 IEEE Topical Conference on Wireless Sensors and Sensor Networks (WiSNeT)	San Juan, Puerto Rico, USA	https://www.radiowirelessweek.org/
01.20. - 01.23.	2025 30th Asia and South Pacific Design Automation Conference (ASP-DAC)	Tokyo, Japan	https://www.aspdac.com/aspdac2025/
01.20. - 01.21.	2025 IEEE Electrical Energy Storage Applications and Technologies Conference (EESAT)	Charlotte, North Carolina, USA	https://cmte.ieee.org/pes-eesat/

일 자	학술대회명	개최장소	홈페이지/연락처
01.20. - 01.22.	2025 IEEE Applied Sensing Conference (APSCON)	Hyderabad, India	https://2025.ieee-apscon.org/
01.20. - 01.22.	2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T)	Raipur, India	http://icpc2t.nitr.ac.in/
01.21. - 01.23.	2025 IEEE PES Grid Edge Technologies Conference & Exposition (Grid Edge)	San Diego, California, USA	https://pes-gradedge.org/
01.23. - 01.25.	2025 IEEE 23rd World Symposium on Applied Machine Intelligence and Informatics (SAMI)	Stará Lesná, Slovakia	https://conf.uni-obuda.hu/sami2025/
01.23. - 01.24.	2025 International Conference on Next Generation Communication & Information Processing (INCIP)	Bangalore, India	https://incip.in/
01.27. - 01.29.	2025 IEEE International Conference on Artificial Intelligence and eXtended and Virtual Reality (AIxVR)	Lisbon, Portugal	https://aixvr.tecnico.ulisboa.pt/
01.27. - 01.30.	2025 Annual Reliability and Maintainability Symposium (RAMS)	Destin, Florida, USA	https://rams.org/
01.28. - 01.30.	2025 Institute for the Future of Education Conference (IFE)	Monterrey, Mexico	https://ciie.itesm.mx/en/

>>2025년 2월

02.02. - 02.04.	2025 15th International Renewable Energy Congress (IREC)	Hamamet, Tunisia	https://irec-conference.com/
02.02. - 02.05.	2025 Cybernetics & Informatics (K&I)	Mikulov na Morave, Czech Republic	https://ki2025.sski.sk/
02.03. - 02.05.	2025 Conference on Artificial Intelligence x Multimedia (AIxMM)	Laguna Hills, California, USA	https://www.aixmm.org/
02.05. - 02.07.	2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)	Bengaluru, India	http://icoici.org/2025/
02.06. - 02.07.	2025 International Conference on Artificial Intelligence and Data Engineering (AIDE)	Nitte, India	http://aide2025.in/
02.06. - 02.11.	2025 13th International Conference on Intelligent Control and Information Processing (ICICIP)	Muscat, Oman	https://conference.cs.cityu.edu.hk/icicip/
02.06. - 02.07.	2025 International Conference on Advances in Renewable Energy & Electric Vehicles (AREEV)	Karkala, India	http://www.areev2025.in/
02.07. - 02.09.	2025 Emerging Technologies for Intelligent Systems (ETIS)	Trivandrum, India	https://www.etis-2025.org/
02.07. - 02.08.	2025 International Conference on Computational, Communication and Information Technology (ICCCIT)	Indore, India	https://ies.ipsacademy.org/incccit-2024/
02.08. - 02.09.	2025 International Conference on Pervasive Computational Technologies (ICPCT)	Greater Noida, India	https://www.glbitm.org/icpct-2025/
02.09. - 02.12.	2025 IEEE Conference on Advances in Magnetism (AIM)	Bressanone, Italy	https://www.aim2025.it/
02.09. - 02.12.	2025 IEEE International Conference on Big Data and Smart Computing (BigComp)	Kota Kinabalu, Malaysia	http://www.bigcomputing.org/conf2025/
02.09. - 02.11.	2025 2nd International Conference on Advanced Innovations in Smart Cities (ICAISC)	Jeddah, Saudi Arabia	https://www.jicollge.edu.sa/en/conferences-and-events/
02.10. - 02.11.	2025 6th Australian Microwave Symposium (AMS)	Gold Coast, Australia	http://www.amsymp.org/index_s.php
02.10. - 02.11.	2025 IEEE Texas Power and Energy Conference (TPEC)	College Station, Texas, USA	https://tpec.engr.tamu.edu/
02.10. - 02.12.	2025 International Conference on Innovation in Artificial Intelligence and Internet of Things (AIIT)	Jeddah, Saudi Arabia	https://aiit.daasite.online/

일 자	학술대회명	개최장소	홈페이지/연락처
02.11. - 02.13.	2025 International Conference on Control, Automation, and Instrumentation (IC2AI)	Beirut, Lebanon	http://lreee.org/ic2ai/
02.13. - 02.14.	2025 International Conference on Intelligent Control, Computing and Communications (IC3)	Mathura, India	https://www.glbajajgroup.org/ICCC-2025
02.16. - 02.20.	2025 IEEE International Solid-State Circuits Conference (ISSCC)	San Francisco, California, USA	https://www.isscc.org/
02.18. - 02.21.	2025 International Conference on Artificial Intelligence in Information and Communication (ICAIC)	Fukuoka, Japan	https://icaic.org/
02.19. - 02.21.	2025 IEEE International Workshop on Antenna Technology (IAT)	Cocoa Beach, Florida, USA	https://attend.ieee.org/iwat-2025/
02.19. - 02.22.	2025 29th International Conference on Information Technology (IT)	Zabljak, Montenegro	https://www.it.ucg.ac.me/en/
02.20. - 02.22.	2025 18th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC)	Porto, Portugal	https://biostec.scitevents.org/
02.21. - 02.22.	2025 First International Conference on Advances in Computer Science, Electrical, Electronics, and Communication Technologies (CE2CT)	Bhimtal, Nainital, India	http://ce2ct.gehu.ac.in/
02.23. - 02.26.	2025 International Conference on Mobile and Miniaturized Terahertz Systems (ICMMTS)	Dubai, United Arab Emirates	http://typo3-202311301006.p415500.webspaceconfig.de/home
02.23. - 02.25.	2025 World Utility Summit (WUS)	Greater Noida, India	https://www.worldutilitysummit.org/
02.25. - 02.28.	2025 IEEE 16th Latin America Symposium on Circuits and Systems (LASCAS)	Bento Gonçalves, Brazil	https://www.ufrgs.br/lascas/index.php
02.26 - 03.06.	2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)	Tucson, Arizona, USA	https://wacv2025.thecvf.com/
02.28 - 03.01.	2025 International Conference on Innovation in Computing and Engineering (ICE)	Greater Noida, India	https://snuice.in/
02.28 - 03.02	2025 4th International Conference on Smart Grid and Green Energy (ICSGGE)	Sydney, Australia	http://www.icsgge.org/
02.28 - 03.02	2025 IEEE International Conference on Mechatronics (ICM)	Wollongong, Australia	https://icm2025.ieee-ies.org/

>>2025년 3월

03.01. - 03.08	2025 IEEE Aerospace Conference	Big Sky, Montana, USA	https://www.aeroconf.org/
03.01. - 03.05	2025 IEEE International Symposium on High Performance Computer Architecture (HPCA)	Las Vegas, Nevada, USA	https://hpca-conf.org/
03.02. - 03.05	2025 IEEE Underwater Technology (UT)	Taipei, Taiwan	https://ut2025.org/
03.03. - 03.07.	2025 IEEE IAS Electrical Safety Workshop (ESW)	Jacksonville, Florida, USA	https://www.ewh.ieee.org/cmte/ias-esw/
03.03. - 03.06.	2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)	Melbourne, Australia	https://humanrobotinteraction.org/2025/
03.04. - 03.07.	2025 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)	Montreal, Quebec, Canada	https://conf.researchr.org/home/saner-2025
03.04. - 03.06.	2025 International Conference on Automation and Computation (AUTOCOM)	Dehradun, India	https://autocom.org.in/
03.05. - 03.07.	2025 International Conference on Emerging Smart Computing and Informatics (ESCI)	Pune, India*	https://escioit.org/

일 자	학술대회명	개최장소	홈페이지/연락처
03.06. - 03.08.	2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)	Gwalior, India*	https://iatmsi.iitm.ac.in/
03.07. - 03.09.	2025 International Conference on Power Electronics and Electric Drives (PEED)	Dali, China	http://www.icpeed.com/
03.07. - 03.08.	2025 3rd International Conference on Disruptive Technologies (ICDT)	Greater Noida, India	https://www.glbitm.org/icdt-2025/
03.07. - 03.09.	2025 9th International Conference on Green Energy and Applications (ICGEA)	Singapore, Singapore	https://www.icgea.org/index.html
03.07. - 03.08.	2025 3rd International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)	Tumakuru, India*	http://icsses.sit.ac.in/
03.09. - 03.12.	2025 9th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)	Hong Kong, Hong Kong	https://edtm2025.com/
03.10. - 03.12.	2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)	Prawet, Thailand *	http://icmlas.com/
03.11. - 03.14.	2025 28th Conference on Innovation in Clouds, Internet and Networks (ICIN)	Paris, France	https://www.icin-conference.org/
03.12. - 03.14.	2025 Systems of Signals Generating and Processing in the Field of on Board Communications	Moscow, Russia	http://media-publisher.ru/en/2025-on-board/
03.14. - 03.16.	2025 14th International Conference on Educational and Information Technology (ICEIT)	Guangzhou, China	https://www.iceit.org/index.html
03.14. - 03.16.	2025 IEEE 8th Information Technology and Mechatronics Engineering Conference (ITOEC)	Chongqing, China	http://www.itoec.org/
03.15.	2025 IEEE Integrated STEM Education Conference (ISEC)	Princeton, New Jersey, USA	https://ewh.ieee.org/conf/stem/
03.16. - 03.20.	2025 IEEE Applied Power Electronics Conference and Exposition (APEC)	Atlanta, Georgia, USA	https://apec-conf.org/
03.16. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Artificial Life and Cooperative Intelligent Systems Companion (ALIFE-CIS Companion)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.19.	2025 IEEE International Conference on Electronics, Energy Systems and Power Engineering (EESPE)	Shenyang, China	http://www.eespe.org/
03.17. - 03.21.	2025 IEEE International Conference on Pervasive Computing and Communications (PerCom)	Washington DC, District of Columbia, USA "	https://www.percom.org/
03.17. - 03.20.	2025 IEEE Symposium Series on Computational Intelligence (SSCI)	Trondheim, Norway	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposia on Computational Intelligence for Energy, Transport and Environmental Sustainability (CIETES)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence on Engineering/Cyber Physical Systems (CIES)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Image, Signal Processing and Synthetic Media (CISM)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Artificial Life and Cooperative Intelligent Systems (ALIFE-CIS)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Security, Defence and Biometrics (CISDB)	Trondheim, Norway*	https://ieee-ssci.org/

일 자	학술대회명	개최장소	홈페이지/연락처
03.17. - 03.20.	2025 IEEE Symposium for Multidisciplinary Computational Intelligence Incubators (MCII)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Natural Language Processing and Social Media (CI-NLPSoMe)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CiFer)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Health and Medicine (CIHM)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Image, Signal Processing and Synthetic Media Companion (CISM Companion)	Trondheim, Norway*	https://ieee-ssci.org/
03.17. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Security, Defence and Biometrics Companion (CISDB Companion)	Trondheim, Norway*	https://ieee-ssci.org/
03.18. - 03.21.	2025 Data Compression Conference (DCC)	Snowbird, Utah, USA	https://datacompressionconference.org/
03.19. - 03.21.	2025 24th International Symposium INFOTEH-JAHORINA (INFOTEH)	East Sarajevo, Bosnia and Herzegovina*	https://infoteh.etf.ues.rs.ba/indexe.php
03.21. - 03.23.	2025 5th International Conference on Advances in Electrical, Electronics and Computing Technology (EECT)	Guangzhou, China	http://www.eect-conf.net/
03.21. - 03.23.	2025 2nd International Conference on Smart Grid and Artificial Intelligence (SGAI)	Changsha, China	https://www.icsgai.com/
03.21. - 03.22	2025 3rd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)	Dehradun, India	https://dicct.geu.ac.in/
03.23. - 03.26	2025 IEEE Engineering Education World Conference (EDUNINE)	Montevideo, Uruguay	https://edunine.eu/edunine2025/eng/index.php
03.24. - 03.27.	2025 IEEE Wireless Communications and Networking Conference (WCNC)	Milan, Italy	https://wcnc2025.ieee-wcnc.org/
03.24. - 03.27.	2025 IEEE 37th International Conference on Microelectronic Test Structures (ICMTS)	San Antonio, Texas, USA	https://icmts.net/
03.26. - 03.28.	2025 IEEE Green Technologies Conference (GreenTech)	Wichita, Kansas, USA	https://ieeegreentech.org/
03.26. - 03.28.	2025 27th International Conference on Digital Signal Processing and its Applications (DSPA)	Moscow, Russia	http://dspa-conf.org/pages/home
03.28. - 03.31	2025 7th Asia Energy and Electrical Engineering Symposium (AEEES)	Chengdu, China	https://www.aeees.org/
03.28. - 03.30.	2025 IEEE Region 5 Annual Meeting (R5)	Wichita, Kansas, USA	https://r5conferences.org/
03.28. - 03.29.	2025 Next Generation Information System Engineering Conference (NGISE)	Ghaziabad, Delhi (NCR), India	https://ngise.org/
03.29. - 03.31.	2025 1st International Conference on Consumer Technology (ICCT-Pacific)	Matsue, Shimane, Japan	https://2025.icct-pacific.org/2025
03.30. - 04.04.	2025 19th European Conference on Antennas and Propagation (EuCAP)	Stockholm, Sweden	https://www.eucap.org/
03.30. - 04.03.	2025 Optical Fiber Communications Conference and Exhibition (OFC)	San Francisco, California, USA	https://www.ofcconference.org/en-us/home/
03.30. - 04.03.	2025 IEEE International Reliability Physics Symposium (IRPS)	Monterey, California, USA	https://www.irps.org/

일 자	학술대회명	개최장소	홈페이지/연락처
03.31. - 04.04.	2025 IEEE Conference on Software Testing, Verification and Validation (ICST)	Napoli, Italy	https://conf.researchr.org/home/icst-2025
03.31. - 04.04.	2025 IEEE 22nd International Conference on Software Architecture (ICSA)	Odense, Denmark	https://conf.researchr.org/home/icsa-2025/
03.31. - 04.04.	2025 IEEE 22nd International Conference on Software Architecture Companion (ICSA-C)	Odense, Denmark	https://conf.researchr.org/home/icsa-2025

>>2025년 4월

04.02. - 04.04.	2025 IEEE International Conference on Cybernetics and Innovations (ICCI)	Chonburi, Thailand	http://icci2025.smc-thailand.org/
04.03. - 04.05.	2025 2nd International Conference on Trends in Engineering Systems and Technologies (ICTEST)	Ernakulam, India	https://ictest.mec.ac.in/
04.04. - 04.06.	2025 International Conference on Advancement in Communication & Computing in Technology (INOACC)	Bengluru, India	https://inoacc.in/
04.04. - 04.06.	2025 27th International Conference on Enterprise Information Systems (ICEIS)	Porto, Portugal	https://iceis.scitevents.org/
04.04. - 04.06.	2025 20th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)	Porto, Portugal	https://enase.scitevents.org/
04.05. - 04.06.	2025 Devices for Integrated Circuit (DevIC)	Kalyani, India	https://edu.ieee.org/in-kgec/devic-2025-home-page/
04.05. - 04.06.	2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI)	Michigan, USA	http://www.icmiconf.com/
04.06. - 04.11.	ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	Hyderabad, India	https://2025.ieeeicassp.org/
04.06. - 04.11.	2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)	Hyderabad, India	https://2025.ieeeicassp.org/
04.07. - 04.10.	2025 IEEE International systems Conference (SysCon)	Montreal, Quebec, Canada	https://ieeesystemscouncil.org/event/conference/ieee-international-systems-conference
04.08. - 04.10.	2025 IEEE International Workshop on Integrated Power Packaging (IWIPP)	Tuscaloosa, Alabama, USA	http://iwipp.org/
04.08. - 04.09.	2025 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)	Waco, Texas, USA	https://texassymposium.org/
04.08. - 04.10.	2025 7th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)	Moscow, Russia	https://reepe.mpei.ru/IEEE/Pages/default.aspx
04.09. - 04.11.	2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)	Copenhagen, Denmark	https://satml.org/
04.10. - 04.11.	2025 International Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits (INMMIC)	Torino, Italy	https://www.inmmic.org/
04.12.	2025 IEEE International Conference on Robotics and Technologies for Industrial Automation (ROBOTHIA)	Kuala Lumpur, Malaysia	https://sites.google.com/view/robothia-malaysia/home
04.13. - 04.15.	2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD)	Manama, Bahrain	http://itikd.ahlia.edu.bh/
04.13. - 04.17.	2025 IEEE Custom Integrated Circuits Conference (CICC)	Boston, Massachusetts, USA	https://www.ieee-cicc.org/
04.14. - 04.17.	2025 26th International Vacuum Electronics Conference (IVEC)	Rotterdam, Netherlands	https://atpi.eventsair.com/ivec-2025

일 자	학술대회명	개최장소	홈페이지/연락처
04.14. - 04.17.	2025 IEEE Silicon Photonics Conference (SiPhotonics)	London, United Kingdom	http://www.ieee-siphotonics.org/
04.14. - 04.17.	2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)	Houston, Texas, USA	https://biomedicalimaging.org/2025/
04.14. - 04.15.	2025 IEEE Wireless and Microwave Technology Conference (WAMICON)	Cocoa Beach, Florida, USA	https://www.ieeewamicon.org/
04.15. - 04.17.	2025 IEEE International Microwave Biomedical Conference (IMBiOC)	Kaohsiung, Taiwan	https://www.imbioc2025.org/site/page.aspx?pid=901&sid=1585&lang=en
04.15. - 04.19.	2025 10th Asia Conference on Power and Electrical Engineering (ACPEE)	Beijing, China	https://acpee.net/
04.16. - 04.17.	2025 2nd International Conference on Microwave, Antennas & Circuits (ICMAC)	Islamabad, Pakistan	http://icmac.seecs.edu.pk/
04.16. - 04.17.	2025 11th International Conference on Web Research (ICWR)	Tehran, Iran	https://iranwebconf.ir/
04.18. - 04.20.	2025 11th International Conference on Control, Automation and Robotics (ICCAR)	Kyoto, Japan	https://www.iccar.org/
04.18. - 04.20.	2025 6th International Conference on Electrical, Electronic Information and Communication Engineering (EEICE)	Shenzhen, China	https://www.eeice.net/
04.18. - 04.20.	2025 International Conference on Sensor-Cloud and Edge Computing System (SCECS)	Zhuhai, China	https://www.scecs.org/
04.20. - 04.23.	2025 IEEE Conference on Technologies for Sustainability (SusTech)	Los Angeles, California, USA	https://ieee-sustech.org/
04.22. - 04.25.	2025 IEEE Global Engineering Education Conference (EDUCON)	London, United Kingdom	https://educon-conference.org/
04.22. - 04.25.	2025 IEEE 18th Pacific Visualization Conference (PacificVis)	Taipei City, Taiwan	https://pacificvis2025.github.io/pages/index.html
04.22. - 04.24.	2025 IEEE 5th International Conference on Digital Twins and Parallel Intelligence (DTPPI)	Atlanta, Georgia, USA	http://2025.ieee-dtpi.org/
04.22. - 04.24.	2025 IEEE International Conference on RFID (RFID)	Atlanta, Georgia, USA	https://2025.ieee-rfid.org/
04.23. - 04.24.	2025 IEEE Vision, Innovation, and Challenges Summit and Honors Ceremony (VIC Summit)	Tokyo, Japan	https://corporate-awards.ieee.org/event/vic-summit-honors-ceremony-gala/
04.23. - 04.25.	2025 26th International Symposium on Quality Electronic Design (ISQED)	San Francisco, California, USA	https://www.isqed.org/
04.24. - 04.25.	2025 International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology (5NANO)	Ernakulam, Kerala, India	https://www.5nano2025.com/
04.24. - 04.25.	2025 13th International Workshop on Biometrics and Forensics (IWBF)	Munich, Germany	https://www.unibw.de/iwbf2025
04.25. - 04.26.	2025 IEEE Bangalore Humanitarian Technology Conference (B-HTC)	Belagavi, India	http://bhct-2025.ieeebangalore.org/
04.25. - 04.27.	2025 10th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)	Chengdu, China	https://www.icccbda.com/index.html
04.25. - 04.27.	2025 IEEE 3rd International Conference on Control, Electronics and Computer Technology (ICCECT)	Jilin, China	http://www.iccect.com/
04.25. - 04.27.	2025 11th International Conference on Electrical Engineering, Control and Robotics (EECR)	Changzhou, China	https://eecr.org/index.html

일 자	학술대회명	개최장소	홈페이지/연락처
04.25. - 04.27.	2025 International Conference on Intelligent Transportation and New Energy Technology (ITNET)	Nanning, China	https://www.ic-itnet.com/
04.26. - 05.03.	2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)	Ottawa, Ontario, Canada	http://www.icse-conferences.org/
04.28. - 05.01.	2025 IEEE Rural Electric Power Conference (REPC)	Denver, Colorado, USA	https://ieeerepc.org/
04.28. - 05.01.	2025 IEEE/ION Position, Location and Navigation Symposium (PLANS)	Salt Lake City, Utah, USA	https://www.ion.org/plans/index.cfm
04.28. - 04.30.	2025 IEEE International Conference on Consumer Technology-Europe (ICCT-Europe)	Algarve, Portugal	https://www.it.pt/Events/Event/5632
04.28. - 05.02.	2025 International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAN)	Palermo, Italy	https://www.simpar2025.org/

>>2025년 5월

05.01. - 05.02.	2025 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC)	Bengaluru, India	https://sites.google.com/view/iceccc-2025/home
05.02.	2025 Systems and Information Engineering Design Symposium (SIEDS)	Charlottesville, Virginia, USA	http://engineering.virginia.edu/sieds
05.02. - 05.03.	2025 International Conference on Intelligent and Cloud Computing (ICICC)	Bhubaneswar, India	http://www.icicc.org.in/
05.03. - 05.09.	2025 IEEE International Radar Conference (RADAR)	Atlanta, Georgia, USA	https://2025.ieee-radar.org/
05.04. - 05.08.	2025 IEEE 33rd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)	Fayetteville, Arkansas, USA	https://www.fccm.org/
05.04. - 05.09.	2025 Conference on Lasers and Electro-Optics (CLEO)	Long Beach, California, USA	https://cleoconference.org/
05.04. - 05.07.	2025 IEEE 32nd Symposium on Computer Arithmetic (ARITH)	El Paso, Texas, USA	https://www.arith2025.org/
05.04. - 05.06.	2025 IEEE 3rd International Conference on Mobility, Operations, Services and Technologies (MOST)	Newark, Delaware, USA	http://ieeemobility.org/MOST2025/index.php
05.04. - 05.08.	2025 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA)	Birmingham, Alabama, USA	https://cementconference.org/
05.05. - 05.07.	2025 IEEE 28th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)	Lyon, France	https://cas.polito.it/DDECS2025/
05.05. - 05.07.	2025 IEEE Conference on Artificial Intelligence (CAI)	Santa Clara, California, USA	https://cai.ieee.org/2025/
05.05. - 05.08.	2025 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)	San Jose, California, USA	http://www.hostsymposium.org/
05.05. - 05.07.	2025 Conference on AI x Software Engineering	Laguna Hills, California, USA	https://www.aixse.org/
05.05. - 05.08.	IEEE INFOCOM 2025 - IEEE Conference on Computer Communications	London, United Kingdom	https://infocom2025.ieee-infocom.org/
05.05. - 05.09.	2025 II International Congress of Engineering with a Social Sense (CIISS)	Bogotá D.C., Colombia	https://eventos.uniminuto.edu/107918
05.06. - 05.09.	2025 11th International Conference on Engineering, Applied Sciences, and Technology (ICEAST)	Phuket, Thailand	https://iceast.kmitl.ac.th/2025/
05.06. - 05.09.	2025 IEEE 31st Real-Time and Embedded Technology and Applications Symposium (RTAS)	Irvine, California, USA	http://2025.rtas.org/

일 자	학술대회명	개최장소	홈페이지/연락처
05.08. - 05.09.	2025 IEEE Build-Up Substrate Symposium (BUSS)	Milpitas, California, USA	http://ieee-buss.org/
05.09.	2025 IEEE Sensors in Spotlight (SenSIS)	Boston, Massachusetts, USA	https://2025.ieee-sensorsinspotlight.org/
05.09. - 05.11.	2025 Global Conference in Emerging Technology (GINOTECH)	PUNE, India	https://ginotech.org/
05.09. - 05.10.	2025 Intermountain Engineering, Technology and Computing (IETC)	Orem, Utah, USA	https://www.uvu.edu/cet/i-etec/
05.09. - 05.11.	2025 IEEE 2nd International Conference on Big Data Science and Engineering (ICBDSE)	Kunming, China	https://www.icbdse.org/
05.09. - 05.11.	2025 4th International Symposium on New Energy Technology Innovation and Low Carbon Development (NET-LC)	Hangzhou, China	https://www.net-lc.net/
05.09. - 05.11.	2025 5th International Conference on Mechanical, Electronics and Electrical and Automation Control (METMS)	Chongqing, China	https://www.met-ms.com/
05.09. - 05.11.	2025 10th International Conference on Control and Robotics Engineering (ICCRE)	Nagoya, Japan	https://www.iccre.org/
05.11. - 05.14.	2025 IEEE 29th Workshop on Signal and Power Integrity (SPI)	Gaeta, Italy	https://spi-workshop.org/
05.11. - 05.13	2025 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)	Ghent, Belgium	http://ispass.org/ispass2025/
05.11. - 05.14.	2025 IEEE Energy Conversion Congress & Exposition Asia (ECCE-Asia)	Bengaluru, India	https://ecceasia2025.org/
05.11. - 05.14.	2025 IEEE/IAS 61st Industrial and Commercial Power Systems Technical Conference (I&CPS)	Montreal, Quebec, Canada	https://site.ieee.org/ias-icps/main-conference/
05.11. - 05.14.	2025 29th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC)	Portland, Oregon, USA	https://asynctsymposium.org/
05.12. - 05.15.	2025 IEEE Symposium on Security and Privacy (SP)	San Francisco, California, USA	https://sp2025.ieee-security.org/
05.12. - 05.16.	2025 Joint Conference of the European Frequency and Time Forum and IEEE International Frequency Control Symposium (EFTF/IFCS)	Querétaro, Mexico	https://2025.ieee-ifcs-efrf.org/
05.12. - 05.16.	2025 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)	Sochi, Russia	https://icieam.su-ieee.ru/
05.12. - 05.14.	2025 ESA Workshop on Aerospace EMC (Aerospace EMC)	Seville, Spain	https://atpi.eventsair.com/emc-2025
05.12. - 05.16.	2025 International Wireless Communications and Mobile Computing (IWCMC)	Abu Dhabi, United Arab Emirates	https://iwcmc.net/2025/
05.12. - 05.16.	NOMS 2025-2025 IEEE Network Operations and Management Symposium	Honolulu, Hawaii, USA	https://noms2025.ieee-noms.org/
05.12. - 05.16.	2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)	St. Petersburg, Russia	http://media-publisher.ru/en/about-weconf-2025/
05.12. - 05.16.	2025 IEEE Magnetic Frontiers: Altermagnetism - New Opportunities in Magnetism	Liblice, Czech Republic	https://www.frontiers2025.cz/
05.13. - 05.14.	2025 International Conference on Military Communication and Information Systems (ICMCIS)	Oerias, Portugal	https://www.icmcis.eu/
05.14. - 05.16.	2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST)	Astana, Kazakhstan	https://sist.astanait.edu.kz/

일 자	학술대회명	개최장소	홈페이지/연락처
05.14. - 05.16.	2025 5th International Conference on Artificial Intelligence and Education (ICAIE)	Suzhou, China	https://www.icaie.org/index.html
05.14. - 05.18.	2025 International Spring Seminar on Electronics Technology (ISSE)	Budapest, Hungary	https://isse-conf.eu/
05.14. - 05.16.	2025 5th International Conference on Pervasive Computing and Social Networking (ICPCSN)	Salem, India	https://icpsn.com/
05.15. - 05.16.	2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)	Fez, Morocco	http://www.iraset.org/2025/
05.15. - 05.17.	2025 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2025 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)	Timisoara, Romania	https://info-optim.ro/
05.16. - 05.19.	2025 37th Chinese Control and Decision Conference (CCDC)	Xiamen, China	http://www.ccdc.neu.edu.cn/
05.16. - 05.18.	2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)	Bhubaneswar, India	http://assic.info/index.php
05.16. - 05.18.	2025 IEEE 3rd International Conference on Power Science and Technology (ICPST)	Kunming, China	https://www.icpst.org/
05.16. - 05.18.	2025 The 16th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT)	Cape Town, South Africa	https://www.mimt.us/
05.16. - 05.17.	2025 3rd International Conference on Data Science and Information System (ICDSIS)	Hassan, India	https://icdsis.in/
05.16. - 05.18.	2025 IEEE 8th International Electrical and Energy Conference (CIEEC)	Changsha, China	https://www.cieec.com.cn/
05.16. - 05.18.	2025 5th International Conference on Computer, Control and Robotics (ICCCR)	Hangzhou, China	https://www.icccr.org/
05.16. - 05.18.	2025 2nd International Conference on Intelligent Computing and Robotics (ICICR)	Dalian, China	http://www.icrconf.com/
05.16. - 05.18.	2025 10th International Conference on Intelligent Computing and Signal Processing (ICSP)	Xi'an, China	https://www.ic-csp.org/
05.16. - 05.18.	2025 IEEE International Annual Conference on Complex Systems and Intelligent Science (CSIS-IAC)	Shenzhen, China	http://www.conference123.org/csisiac2025/
05.16. - 05.18.	2025 8th International Conference on Communication Engineering and Technology (ICCET)	Guangzhou, China	https://www.iccet.org/
05.16. - 05.18.	2025 4th International Conference on Electronics, Integrated Circuits and Communication Technology (EICCT)	Chengdu, China	https://www.ic-eict.net/
05.17. - 05.20.	2025 8th International Conference on Electronics Technology (ICET)	Chengdu, China	https://www.icet.net/
05.18. - 05.21.	2025 IEEE International Memory Workshop (IMW)	Monterey, California, USA	https://www.ewh.ieee.org/soc/eds/imw/
05.18. - 05.21.	2025 IEEE International Electric Machines & Drives Conference (IEMDC)	Houston, Texas, USA	https://www.iemdc.org/
05.18. - 05.21.	2025 15th International Symposium on Linear Drivers for Industry Applications (LDIA)	Daejeon, Korea (South)	http://www.ldia2025.com/
05.19. - 05.23.	2025 IEEE International Conference on Robotics and Automation (ICRA)	Atlanta, Georgia, USA	https://2025.ieee-icra.org/

일 자	학술대회명	개최장소	홈페이지/연락처
05.19. - 05.22.	2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)	Chemnitz, Germany	https://i2mtc2025.ieee-ims.org/
05.19. - 05.23.	2025 Asia-Pacific International Symposium and Exhibition on Electromagnetic Compatibility (APEMC)	Taipei, Taiwan	https://www.apemc2025.org/site/page.aspx?pid=901&sid=1565&lang=en
05.19. - 05.21.	2025 IEEE 34th Microelectronics Design & Test Symposium (MDTS)	Albany, New York, USA	https://mdts.ieee.org/
05.19. - 05.22.	2025 10th International Conference on Fog and Mobile Edge Computing (FMEC)	Tampa, Florida, USA	https://emergingtechnet.org/FMEC2025/index.php
05.19. - 05.22.	2025 International Conference on Modern Computing, Networking and Applications (MCNA)	Tampa, Florida, USA	http://mcna-conference.org/2025/index.php
05.19. - 05.22.	2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing (CCGrid)	Tromsø, Norway	https://site.uit.no/ccgrid2025/
05.19. - 05.22.	2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)	Tromsø, Norway	https://site.uit.no/ccgrid2025/
05.19. - 05.22.	2025 IEEE Conference on Education and Training in Optics and Photonics (ETOP)	Glasgow, United Kingdom	http://www.etop-conference.org/
05.19. - 05.22.	2025 IEEE Emerging Technology Reliability Roundtable (ETR-RT)	Prague, Czech Republic	https://cqr.committees.comsoc.org/etr-rt-2025/
05.19. - 05.22.	2025 IEEE 9th International Conference on Fog and Edge Computing (ICFEC)	Tromsø, Norway	https://icfec2025.ontariotechu.ca/
05.19.	IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)	London, United Kingdom	https://infocom2025.ieee-infocom.org/authors/call-workshop-proposals
05.19. - 05.23.	2025 IEEE 41st International Conference on Data Engineering (ICDE)	Hong Kong, Hong Kong	https://iee-icde.org/2025/
05.20. - 05.22.	2025 IEEE 19th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)	Antalya, Türkiye	https://cpepowereng2025.org/
05.20. - 05.22.	2025 IEEE 26th International Conference on High Performance Switching and Routing (HPSR)	Suita, Osaka, Japan	https://hpsr2025.ieee-hpsr.org/
05.20. - 05.23.	2025 22nd International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)	Bangkok, Thailand	https://ecti-con2025.eng.chula.ac.th/
05.21. - 05.23.	2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)	Trichy, India	http://icaiss.com/
05.21. - 05.23.	2025 26th International Radar Symposium (IRS)	Hamburg, Germany	https://www.dgon-irs.org/home
05.23. - 05.25.	2025 International Conference in Advances in Power, Signal, and Information Technology (APSIT)	Bhubaneswar, India	https://apsit.co.in/
05.23. - 05.25.	2025 6th International Conference for Emerging Technology (INCET)	BELGAUM, India	http://www.incet.org/
05.23. - 05.25.	2025 IEEE 12th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)	Chongqing, China	http://www.itaic.org/
05.23. - 05.25.	2025 2nd International Symposium on New Energy Technologies and Power Systems (NETPS)	Hangzhou, China	https://www.netps.org/
05.23. - 05.25.	2025 IEEE International Conference on Electrical Energy Conversion Systems and Control IEEESC	Chongqing, China	https://iee-ieesc.net/2025/

일 자	학술대회명	개최장소	홈페이지/연락처
05.23. - 05.26.	2025 8th International Conference on Artificial Intelligence and Big Data (ICAIBD)	Chengdu, China	https://www.icaibd.org/
05.23. - 05.25.	2025 6th International Conference on Control, Communication and Computing (ICCC)	Thiruvananthapuram, India	https://iccc2025.cet.ac.in/
05.23. - 05.25.	2025 IEEE 5th International Conference on Electronic Technology, Communication and Information (ICETCI)	Changchun, China	http://www.icetci.org/
05.23. - 05.24.	2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)	Ankara, Turkiye	http://www.ichoracongress.com/
05.23. - 05.26.	2025 IEEE International Workshop on Radio Frequency and Antenna Technologies (IWRF&AT)	Shenzhen, China	https://rfat.szu.edu.cn/rfat2025/
05.23. - 05.25.	2025 6th International Conference on Computing, Networks and Internet of Things (CNIOT)	Shanghai, China	http://www.cniot-conf.net/
05.23. - 05.25.	2025 IEEE 5th International Conference on Computer Communication and Artificial Intelligence (CCAI)	Haikou, China	http://www.ccai.net/
05.23. - 05.26.	2025 17th International Conference on Bioinformatics and Biomedical Technology (ICBBT)	Hangzhou, China	https://www.icbbt.org/index.htm
05.23. - 05.25.	2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL)	Ningbo, China	https://www.cvidl.org/
05.24. - 05.25.	2025 IEEE 15th Symposium on Computer Applications & Industrial Electronics (ISCAIE)	Penang, Malaysia	https://www.iscaie.org/home
05.25. - 05.28.	2025 IEEE International Symposium on Circuits and Systems (ISCAS)	London, United Kingdom	https://iscas2025.org/lander
05.26. - 05.31.	2025 IEEE 75th Electronic Components and Technology Conference (ECTC)	Dallas, Texas, USA	https://www.ectc.net/
05.26. - 05.30.	2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG)	Tampa/Clearwater, Florida, USA	https://fg2025.ieee-biometrics.org/
05.26. - 05.29.	2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)	Barcelona, Spain	https://icmlcn2025.ieee-icmlcn.org/
05.26. - 05.28	2025 IEEE 5th International Conference on Human-Machine Systems (ICHMS)	Abu Dhabi, United Arab Emirates	http://ieee-ichms.org/2025/
05.26. - 05.28	2025 IEEE European Technology and Engineering Management Summit (E-TEMS)	Bruges, Belgium	https://etems.digital/2025/
05.26. - 05.29.	2025 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)	Vancouver, British Columbia, Canada	https://attend.ieee.org/ccece-2025/
05.27. - 05.30.	2025 24th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)	Dallas, Texas, USA	https://www.ieee-itherm.net/
05.27. - 05.30.	2025 International Conference on Military Technologies (ICMT)	Brno, Czech Republic	https://icmt2025.cz/
05.27. - 05.30.	2025 12th International Conference on Information Technology (ICIT)	Amman, Jordan	https://icit.zuj.edu.jo/Home/
05.28. - 05.30.	2025 7th International Conference on Inventive Material Science and Applications (ICIMA)	Namakkal, India	https://icoimsa.com/2025/
05.28. - 05.29.	2025 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)	Novi Sad, Serbia	https://www.gozinc.org/
05.28. - 05.30.	2025 IEEE Medical Measurements & Applications (MeMeA)	Chania, Greece	http://www.memea2024.ieee-ims.org/

일 자	학술대회명	개최장소	홈페이지/연락처
05.29. - 05.30.	2025 IEEE Engineering Reliable Autonomous Systems (ERAS)	Worcester, Massachusetts, USA	https://erasrobotics.org/
05.29. - 05.31	2025 IEEE International Conference on Electro Information Technology (eIT)	Valparaiso, Indiana, USA	https://eit-conference.org/eit2025/
05.30. - 06.01.	2025 International Conference on Artificial Intelligence and Digital Ethics (ICAIDE)	Guangzhou, China	https://www.icaide.net/

The Magazine of the IEIE

특별회원사 명단

회원사	대표자	주 소	전 화	홈페이지
(주)디비하이텍	조기석	경기도 부천시 수도로 90(도당동)	032-680-4700	www.dbhitek.com
(주)레티널	김재혁	경기도 안양시 동안구 부림로170번지 41-10, 4층	02-6959-7007	https://letinar.com
(주)마르시스	박용규	서울시 강남구 언주로 85길 7	02-3445-3999	http://www.marusys.com
(주)세미파이브	조명현	경기도 성남시 분당구 양현로 322, 코리아디자인센터 2층		http://www.semifive.com
(주)센서위드유	이윤식	울산광역시 울주군 언양읍 유니스트길 50, 106동 501-4호	052-912-4282	http://www.sensorwyou.com
(주)에스비솔루션	변영재	울산광역시 울주군 언양읍 유니스트길 50, 106동 401-3호	052-217-7343	http://www.sb-solutions.co.kr
(주)와이솔	김지호	경기도 오산시 가창로 531-7	070-7837-2730	http://www.wisol.co.kr
(주)웨이브피아	이상훈	경기도 화성시 동탄기흥로 557 금강펜테리움IT타워 1301호	031-8058-3384	http://www.wavepia.com
KT	김영섭	경기도 성남시 분당구 정자동 206	031-727-0114	http://www.kt.com
LG이노텍(주)	문혁수	서울시 강서구 마곡중앙10로 30	02-3777-1114	www.lginnotek.com
LG전자(주)	조주완	서울시 영등포구 여의도동 30	02-3777-1114	http://www.lge.co.kr
LIG넥스원	김지찬	서울시 서초구 강남대로 369(서초동, 나라빌딩)	02-1644-2005	http://www.lignex1.com
LPKF Laser&Electronics	이용상, 벤델레과츠마티아스	경기도 안양시 동안구 흥안대로 427번길	031-689-3660	www.lpkf.com/kr
SK텔레콤(주)	유영상	서울시 중구 을지로65(을지로2가) SK T-타워	02-2121-2114	http://www.sktelecom.com
SK하이닉스(주)	곽노정	경기도 이천시 부발읍 아미리 산 136-1	031-630-4114	http://www.skhynix.com
네이버(주)	최수연	경기도 성남시 분당구 불정로 6 (정자동 그린팩토리)	031-784-2560	https://www.navercorp.com
누리미디어	최순일	서울시 영등포구 선유로 63, 4층(문래동 6가)	02-710-5300	http://www.nurimedia.co.kr
대덕전자(주)	신영환	경기도 안산시 단원구 강촌로230 (목내동 475)	031-8040-8000	http://www.daeduck.com
대전테크노파크	김우연	대전시 유성구 테크로9로	042-930-4300	www.djtp.or.kr
도쿄อิเล็ก트론코리아(주)	원제형	경기도 화성시 장안면 장안공단 6길 51	031-260-5000	https://www.tel.com
(주)동인시스템	곽동달	부산광역시 해운대구 센텀북대로 60,	051-787-7288	http://www.donginsm.com
롯데렌탈(주)	최진환	경기도 안양시 동안구 전파로88 (신원비전타워 8층)	02-3453-8970	https://www.lotterental.com
리얼테크코리아 주식회사	팅치창	서울시 서초구 사임당로 18, 석오빌딩 5층	070-4120-7966	www.realtek.cpm/en
비전테크	이원복	대전 유성구 테크노2로 187, 미건테크노월드2차 1층 118호	042-934-0236	http://www.visiontechkorea.com
머플 주식회사	안수남	경기 성남시 분당구 판교로 289번길 20, 2동		http://www.murple.ai/
삼성전자(주)	한종희	서울시 서초구 서초2동 1320-10 삼성전자빌딩	02-1588-3366	https://www.samsung.com
스카이칩스	이강윤	수원시 장안구 서부로 2066, 산학협력센터 85511호	031-299-6848	http://www.skaichips.co.kr
스테코(주)	최기환	충청남도 천안시 서북구 3공단1로 20(백석동)	041-629-7480	http://www.steco.co.kr
에스에스앤씨(주)	한은혜	서울시 영등포구 당산로171, 1301	02-6925-2550	http://www.secnc.co.kr
에어스메디컬	이진구	서울시 관악구 남부순환로 1838	070-7777-3186	www.airsmmed.com
오토아이티(주)	정명환	대구시 수성구 알파시티1로 117	053-795-6303	www.auto-it.co.kr
유정시스템(주)	이재훈	서울시 구로구 디지털로26길 110	02-852-8721	www.yjsys.co.kr

회원사	대표자	주 소	전 화	홈페이지
정보통신정책연구원	배경율	충북 진천군 덕산읍 정통로 18	043-531-4389	www.kisdi.re.kr
(주)LX세미콘	이윤태	대전시 유성구 탑립동 707	042-712-7700	www.lxsemicon.com
(주)넥스틴	박태훈	경기도 화성시 동탄면 동탄산단9길 23-12	031-629-2300	http://www.nextinsol.com
(주)더즈텍	김태진	경기도 안양시 동안구 학의로 292 금강펜테리움IT타워 A동 1061호	031-450-6300	http://www.doestek.co.kr
HL만도(주)	조성현	경기도 평택시 포승읍 하만호길 32	02-6244-2114	https://www.hlmando.com/
(주)빅텍	임만규	경기도 이천시 마장면 덕이로 180-31	031-631-7301	http://www.vitek.co.kr
(주)스프링클라우드	송영기	경기도 성남시 창업로 42	031-778-8328	www.aspringcloud.com
(주)시스메이트	이상만	대전시 유성구 유성대로 1184길 41	042-486-6135	http://www.sysmate.com
주식회사 뷰웍스	김후식	경기도 안양시 동안구 부림로 170번길 41-3	070-7011-6161	https://www.vieworks.com
(주)실리콘마이터스	허염	경기도 성남시 분당구 대왕판교로 660 유스페이스-1 A동 8층	1670-7665	http://www.siliconitus.com
(주)싸이몬	정창호	경기도 성남시 분당구 별말로48(구 야탑동 272-1 케이디티빌딩)	02-480-8580	http://www.cimon.com
(주)인텔레콤	박영기	서울시 영등포구 경인로 775, 문래동 3가 에이스하이테크시티 1동 119호	02-3439-0033	http://www.signtelecom.com
(주)솔리드	정준, 이승희	경기도 성남시 분당구 판교역로 220 솔리드스페이스	031-627-6000	http://www.st.co.kr
(주)유니트론텍	남궁 선	서울시 강남구 영동대로 638(삼도빌딩) 9층	02-573-6800	http://unitrontech.com
(주)코클리어닷에이아이	한윤창	서울시 강남구 봉은사로 51길 26		www.cochl.ai
(주)크레셈	오상민	대전시 유성구 대덕대로 582, 4층 402호(도룡동, 옥토빌딩)	031-427-3445	http://www.cressem.com
(주)텔레칩스	이장규	서울시 송파구 올림픽로 35다길 42(신천동 한국루터회관)19~23층	02-3443-6792	www.telechips.com
(주)티에이치엔	이광연, 채승훈	대구시 달서구 갈산동 973-3	053-583-3001	http://www.th-net.co.kr
(주)티엘아이	홍세경	경기도 성남시 중원구 양현로 405번길 12 티엘아이 빌딩	031-784-6800	http://www.tli.co.kr
(주)해치텍	최성민	충북 청주시 청원구 오창읍 연구단지로 40, 스타기업관 207호	043-715-9034	http://www.haechitech.com
중소벤처기업진흥공단	강석진	경상남도 진주시 동진로 430	055-751-9380	www.kosmes.or.kr
케이케이테크(주)	김경하	경기도 안성시 대덕면 무능로132	031-678-1586	http://www.k-ktch.co.kr
코어인사이트(주)	유용훈	경기도 성남시 중원구 갈마치로 186 반포테크노피아 5층	031-750-9200	http://www.coreinsight.co.kr
한국알박(주)	김선길	경기도 평택시 청북읍 한산길5	031-683-2922	http://www.ulvackora.co.kr
한국전자기술연구원	신희동	경기도 성남시 분당구 새나리로 25 (야탑동)	031-789-7740	http://www.keti.re.kr
한국전자통신연구원	방승찬	대전시 유성구 가정로 218	042-860-6114	http://www.etri.re.kr
한화시스템(주)	김연철	서울시 중구 청계천로 86 (장교동) 한화비딩 (19,20층)	02-729-3030	http://www.hanwhasystems.com
현대로템(주)	이용배	경기도 의왕시 철도박물관로 37	031-596-9114	http://www.hyundai-rotem.co.kr
현대모비스(주)	정의선, 이규석	서울시 강남구 테헤란로 203	02-2018-5114	http://www.mobis.co.kr
현대자동차(주)	정의선, 장재훈, 이동석	경기도 화성시 장덕동 772-1	02-3464-1114	http://www.hyundai-motor.com
호리바에스텍코리아(주)	김성환 외 1명	경기도 용인시 수지구 디지털벨리로 98 호리바빌딩	031-6520-6500	http://www.horiba.com
히로세코리아(주)	이상엽	경기도 시흥시 정왕동 희망공원로 250	031-496-7000	http://www.hirose.co.kr
히타치하이테크코리아(주)	MIYOSHI KEITA	경기도 성남시 분당구 정자일로 155, 엔16층(정자동, 분당두산타워)	031-725-4201	https://www.hitachi-hightech.com

박사학위 논문초록 게재 안내

본 학회에서는 전자공학회지에 국내외에서 박사학위를 취득한 회원의 학위 논문초록을 게재하고 있으니 해당 회원 여러분의 적극적인 참여를 바랍니다.(단, 박사학위 취득후 1년 이내에 제출해 주시는 것에 한함.)

성 명	(국문)	(한문)	(영문)
학위취득	학 교 명	대학교	학과
	취득년월	년 월	지도교수
현 근무처 (또는 연락처)	주 소	(우편번호 :)	
	전화번호	FAX번호	
학위논문 제목	국 문		
	영 문		
KEY WORD			

국문 초록(요약) : 1000자 이내

보내실 곳 _ 06130

서울특별시 강남구 테헤란로 7길 22(역삼동, 과학기술회관 제1관 907호)

사무국 회지담당자앞

E-mail : ieie@theieie.org

TEL : (02)553-0255(내선 6번) FAX : (02)562-4753



대한전자공학회

The Institute of Electronics and Information Engineers

전자공학회지 <월간>

제51권 제12호(통권 제487호)

The Magazine of the IEIE

2024년 12월 20일 인쇄

2024년 12월 25일 발행

발행및

편집인

인쇄인

발행인

(사) 대한전자공학회

한림원(주)

사 단 법 인 대 한 전 자 공 학 회

회장 이 총 용

대표 김 흥 중

(우)06130 서울 강남구 테헤란로 7길 22(역삼동, 과학기술회관 제1관 907호)

TEL.(02)553-0255~7 FAX.(02)562-4753

E-mail : ieie@theieie.org

Homepage : http://www.theieie.org



2025년도 회비납부 안내

1. 회비의 납부 및 유효기간

2025년도 회원 연회비는 2024년과 동일함을 알려드리며, 아직 2025년도 회비를 납부하지 않으신 회원님께서서는 납부하여 주시기 바라며, 연회비의 유효기간은 회비를 납부한 당해연도에 한합니다.

◆ 2025년도 회원 연회비는 다음과 같습니다.

- 정 회 원 : 70,000원 (입회비 : 10,000원)
- 학생회원 : 30,000원 (입회비 면제)
- 평생회원 : 700,000원
 - 평생회비 할인 제도 : 학회 홈페이지 안내 참조
 - 평생회비 분납 제도(1년 한) : 평생회비 분할 납부를 원하시는 회원께서는 회원 담당에게 요청하여 주시기 바랍니다.
 - 7월 1일부터 연회비 50% 할인 적용

2. 논문지(eBook) 제공

학회지와 논문지(국·영문)가 eBook으로 발간되어 학회 홈페이지(<http://www.theieie.org>)를 통해 제공되고 있습니다.

3. 회비의 납부방법

신용카드(홈페이지 전자결제) 및 계좌이체(한국씨티은행, 102-53125-258)를 이용하여 학회 연회비, 심사비 및 논문게재료 등 납부 가능합니다.

4. 석·박사 신입생 및 재학생 다년 학생회원 가입 및 회비 할인 제도 안내

우리 학회에서는 석·박사 신입생 및 재학생을 위하여 다년 학생회원 가입 제도 및 회비 할인 제도를 마련하였습니다. 한 번의 회원가입으로 졸업 및 수료 때까지 학회 활동에 참여하실 수 있는 기회가 되시기 바라며 회비 할인 혜택까지 받으시길 바랍니다.

◎ 가입 대상 및 할인 혜택

- 가입 대상 : 2025년 석·박사 신입생 및 재학생
- 할인 내용 : 2년 60,000원(1년당 30,000원) → 2년 50,000원(16.7% 할인)
3년 90,000원(1년당 30,000원) → 3년 70,000원(22.2% 할인)
4년 120,000원(1년당 30,000원) → 4년 90,000원(25% 할인)
5년 150,000원(1년당 30,000원) → 5년 110,000원(26.7% 할인)

5. 문의처

◆ 대한전자공학회 사무국 변은정 부장(회원담당)

Tel : 02-553-0255(내선 3번) / E-mail : edit@theieie.org

전자공학회논문지

단편논문 신설 안내

전자공학회논문지에서 단편논문을 신설하여 투고를 받습니다. 기존의 정규논문과 함께 단편논문을 신설하여, 투고 논문 형식의 다변화와 함께 신속한 논문심사 및 게재를 추진하고자 합니다.

논문투고시스템 투고구분 선택

☞ 정규논문(기존) : 긴급 / 일반 중 택일

▶ 단편논문(신설) : 특급

단편논문 양식

☞ 투고규정 : https://www.theieie.org/pages_journal/journal_info.vm

☞ 논문양식 : <https://www.theieie.org/board/?ncode=a008>

* 심사본 : 3쪽 이내, 최종본 : 4쪽 이내

단편논문 심사비 : 10만원/편당

단편논문 게재료 : 게재논문 면당 5만원, 최대 4쪽 이내

☞ 지원 문구 추가 시 10만원 추가

☞ 교신저자가 비회원인 경우 산정된 게재료의 150% 부과

단편논문 심사 기간 : 2주 이내 1차 심사를 원칙으로 함

시 행 : 2022년 10월 이후