

ISSN 1016-9288

제50권 2호

2023년 2월호

전자공학회지

The Magazine of the IEIE

vol.50. no.2

인공지능 컴퓨팅 최적화 기술

- 딥러닝 모델의 초매개변수 최적화
- 트랜스포머 모델의 경량화
- AI Inference를 위한 하드웨어 플랫폼과 소프트웨어 프레임워크의 동향
- 전자설계자동화 기술과 차세대 지능형반도체 소자 및 회로 설계 응용
- Kubernetes 기반 오픈소스 AI/ML 플랫폼



Boost your lab's performance

임의파형발생기

- 2.4 GSa/s, 16 bit, 750 MHz
- 4 또는 8 채널 이상
- 50 ns 이하의 트리거 딜레이

어플리케이션

반도체 테스트, 양자 컴퓨터, Phased array 레이더 설계 & 테스트, Lidar, 분광학, NMR

임피던스분석기

- DC ~ 5 MHz까지, 1mΩ ~ 1TΩ까지
- 0.05% 기본 정확도
- 측정 정확도 보상 및 측정 신뢰도 표시 기능

어플리케이션

높은 Q 값의 유전체, 정전용량형 센서, 슈퍼 커패시터, PV 소자, 소자 특성 분석

락인앰플리파이어

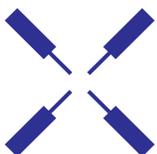
- 600 MHz까지 측정 가능
- 스코프, FFT, 주파수 응답 분석기, Sweeper, 이미징 툴
- 옵션: 임의파형발생기, PID, PLL, Boxcar, 주파수 카운터, AM & FM 변조

어플리케이션

AFM, LVP, CARS, SRS, SNOM, graphene, optical PLL, THz, pump-probe, RFID, MEMS, NEMS, gyros, NDT, MRFM

LabOne® 소프트웨어

취리히인스트루먼트의 모든 장비는 제어소프트웨어인 LabOne®을 사용할 수 있습니다. 다양한 기능, 효율적인 작업, 쉬운 사용자인터페이스를 제공합니다. 웹브라우저에서 장비를 액세스하거나 LabVIEW™, MATLAB®, Python, C 또는 .NET 프로그램들과 통합할 수 있습니다.



Zurich
Instruments

한국담당자 010-6456-3463
ilnam.yeom@zhinst.com
www.zhinst.com

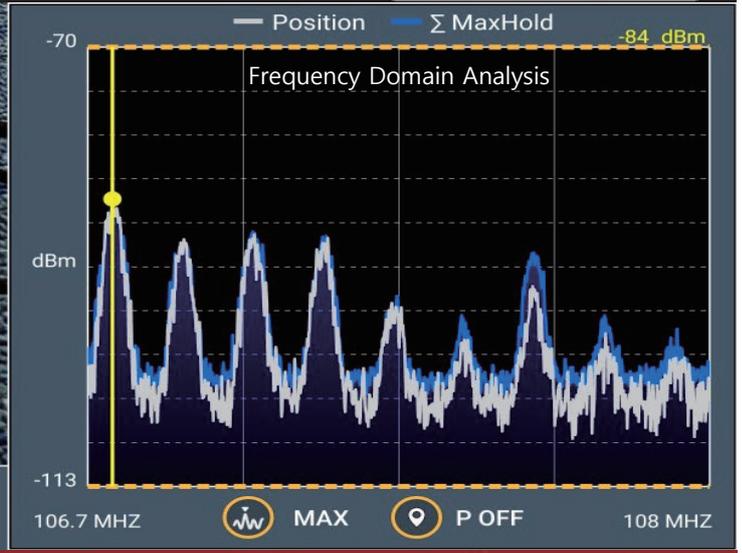
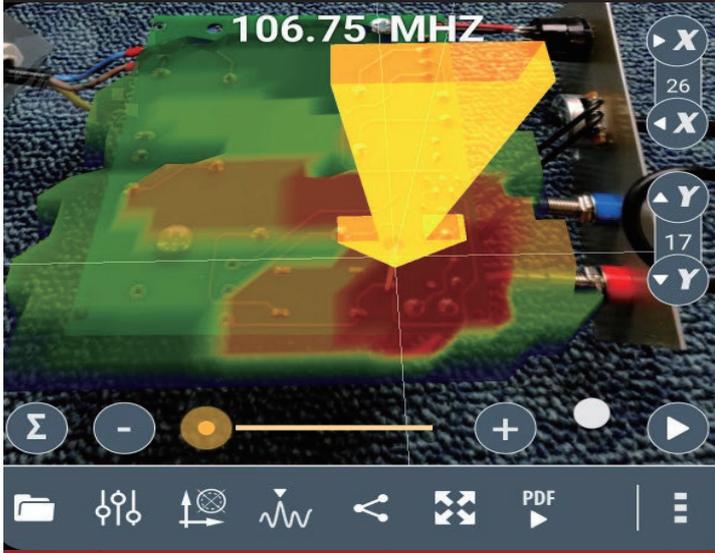
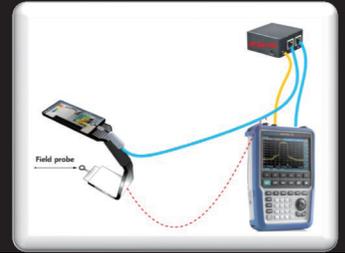
응용분야 솔루션에 대하여
기술문의 해주세요

SCANPHONE

증강현실 EMI/EMC 측정 솔루션!

Scanphone 특징

- 증강현실 기술 적용으로 실시간 측정 실현
- 대형 구조물의 EMI / EMC 테스트 (ex. 실드 박스, TV, 차량, 안테나 etc..)
- 별도의 전원이 필요 없는 휴대용 장비
- 측정 현장에서 DUT와 필드를 중첩하여 분석
- 노이즈 소스원 분석, 안테나 최적화, 방사 패턴
- Scanphone + Spectrum = 주파수 성분 분석
- PC용 SW를 통한 데이터 추출 및 분석



MCK

MCK (Material Characterization Kit) 고주파대역 Material의 유전율 측정 솔루션



Overview

MCK는 전기전자, 화학분야에서 고체, 액체, 파우더, Multi-layer 등 다양한 형태의 Material의 유전특성을 정확하게 측정할 수 있는 제품입니다.



Features

- 25GHz ~ 1.1THz
- Material의 유전율, Loss tangent 측정
- 유전율 측정 오차 범위 : ±1%
- Loss tangent 측정 오차 범위 : ±2%
- Plug & Play방식

Advantage

- 쉽고 빠른 Cal과 측정
- THz 대역의 높은 주파수까지 지원
- Broadband 측정
- 높은 재현성 및 빠른 측정 시간
- Solid, Liquid, Powder, Coating, Muti layer 구조 Material 측정 가능

Benefits

- 높은 주파수 대역 측정으로 5G, 6G 솔루션 지원
- Broadband 측정 및 분석
- 다양한 VNA 제품과의 호환성
- 다양한 Material 측정 가능



2023 하계종합학술대회

6.28(수)~30(금)
롯데호텔 제주(중문)



주요일정

논문제출 : 2023년 5월 10일(수)

심사통보 : 2023년 5월 24일(수)

사전등록 : 2023년 5월 16일(화)~6월 9일(금)



발표분야

소사이어티	연구회
통신(Communication)	통신, 마이크로파 및 전파전파, ITS, 무선PAN/BAN, 미래지능형네트워크, 군사전자
반도체(Semiconductor)	반도체소자 및 재료, SoC 설계, 광파 및 양자전자공학, PCB & Package, RF 집적회로, 정보보안시스템, 내방사선 반도체 설계 및 소자, ESD/EOS & Latchup, 인메모링 컴퓨팅
컴퓨터(Computer)	융합컴퓨팅, 멀티미디어, 인공지능/신경망/퍼지, M2M/IoT, 휴먼ICT, 인공지능 및 보안, 증강휴먼, AI응용, 블록체인
인공지능 신호처리 (AI Signal Processing)	영상처리, 음향 및 음성신호처리, 영상이해, 바이오영상신호처리, 딥러닝, 로봇지능
시스템 및 제어 (System and Control)	의용전자 및 생체공학, 제어계측, 회로 및 시스템, 전력전자, 지능로봇, 국방정보 및 제어, 자동차전자, 의료영상시스템, 스마트팩토리, 스마트 미터링
산업전자 (Industry Electronics)	산업전자제어, 임베디드시스템, 유비쿼터스 센서네트워크, 디지털통신 시스템
New Emerging Area	의료, 에너지, Software, 기타



“미래 기술 · 신산업 기술 정보의 보고” 해동일본기술정보센터, 최신 정보 한글요약 제공

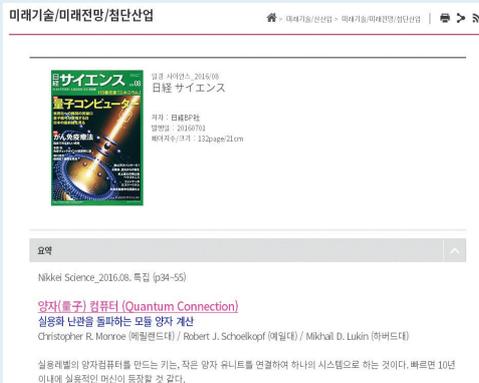
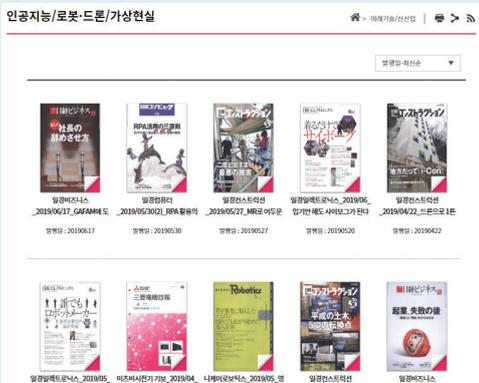


서울대학교 공과대학 해동일본기술정보센터는 대덕전자(故)김정식 회장의 열정과 지원에 의해 최신 일본 기술정보를 산업계와 학계에 널리 알리고자 2010년 3월에 설립하여 현재까지 운영해 오고 있습니다.

3천여권의 공학 및 신산업 관련 서적과 20여종의 Nikkei가 발행한 정기간행물과 40여개사의 기술보고서 등 4천여권의 도서를 통해 다양한 분야의 기술 정보를 제공하고 있습니다.

2016년부터는 소장 정보를 26개의 신산업 카테고리 구분하여 미래기술과 신산업 관련한 정기간행물의 특집기사와 신문기사의 한글요약 제공과 함께, 주간브리핑 등을 통해 매주 새로운 정보를 메일과 SNS 등으로 배포하고 있습니다.

상세한 사항은 로그인 없이 모든 정보와 이용이 가능한 홈페이지를 참조바랍니다.



해동일본기술정보센터
HAEDONG JAPAN TECHNOLOGY INFORMATION CENTER
HJTIC <http://hjt看.snu.ac.kr>
08826 서울특별시 관악구 관악로 1,
서울대학교 공과대학 35동
전화 : 02-880-8279

<https://www.facebook.com/snuhjt看>
<http://blog.naver.com/hjt看2010>
카카오톡 오픈채팅@HJTIC 브리핑룸

주간브리핑의 무료 이메일 구독을 원하시면,
['hjt看@snu.ac.kr'에 "구독"으로 신청.](mailto:hjt看@snu.ac.kr)
카카오톡으로 매일의 기사까지 받아 보시려면,
['오픈채팅@HJTIC브리핑룸' 가입 \(pw:2016\)](https://openchatting.com/join/pw2016)

IEEE DiscoveryPoint Communications(IDPC)



IEEE가 만든 첨단 통신산업 R&D를 위한 All-in-One Platform



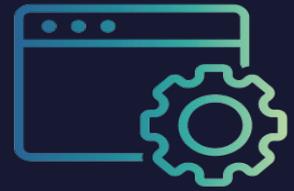
IEEE Full-text를
포함한 다양한 양질의
자료 제공



최첨단 연구를 위한
자료, 표준, 공급업체를
위한 솔루션까지
하나의 플랫폼에서
검색가능



5G, 6G, Edge,
Computing, IoT 등
다양한 통신기술
분야의 자료중
전문적으로
선별된 콘텐츠 제공



설계 문제에 대한
솔루션 제공하여
실무 엔지니어에게
최적화

Trial가능,
문의 02-3474-5290
이희진 과장 hjlee@kitis.co.kr
김은진 대리 ejkim@kitis.co.kr



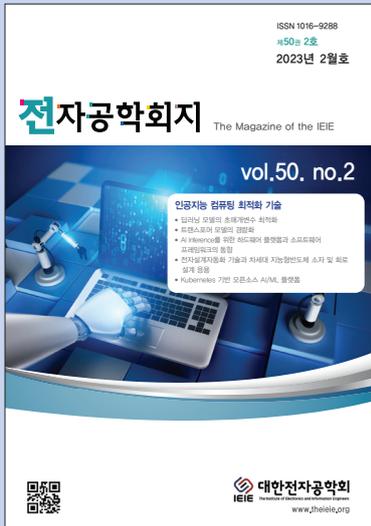
Authorized
Dealer



키티스 産學研情報(株)
KITIS Info. & Co., Ltd.

CONTENTS

제50권 2호 (2023년 2월)



※ 학회지 2월호 표지 (vol 50, No 2)

회지편집위원회

- 위 원 장 양 준 성 (연세대학교 교수)
- 위 원 박 관 서 (연세대학교 교수)
- 박 승 창 (㈜유오씨 사장)
- 안 진 호 (호서대학교 교수)
- 윤 석 현 (단국대학교 교수)
- 장 지 원 (연세대학교 교수)
- 정 재 용 (인천대학교 교수)
- 최 정 옥 (한양대학교 교수)
- 허 준 (고려대학교 교수)
- 한 태 화 (연세대학교 의료원 팀장)
- 사무국 편집담당
배 기 동 부장
TEL : (02)553-0255(내선 5)
FAX : (02)552-6093
- 학회 홈페이지
<http://www.theieie.org>

학회소식

12 학회소식 / 편집부

특집 : 인공지능 컴퓨팅 최적화 기술

17 특집편집기 / 김명선

18 딥러닝 모델의 초매개변수 최적화 / 김신규

27 트랜스포머 모델의 경량화 / 오지훈

37 AI Inference를 위한 하드웨어 플랫폼과 소프트웨어 프레임워크의 동향 / 이연복

47 전자설계자동화 기술과 차세대 지능형반도체 소자 및 회로 설계 응용 / 정성엽

57 Kubernetes 기반 오픈소스 AI/ML 플랫폼 / 박용석

회원광장

63 논문지 논문목차

65 박사학위 논문초록 / 이상학(티맥스소프트)

정보교차로

66 국내외 학술행사 안내 / 편집부

80 특별회원사 및 후원사 명단

2023년도 임원 및 각 위원회 위원

회 장	이 혁 재 (서울대학교 교수)	권 호 열 (정보통신정책연구원 원장)
수석부회장	이 총 용 (연세대학교 교수) - 총괄 / AI	김 영 재 (해동과학문화재단 이사장)
고 문	권 오 경 (한양대학교 석좌교수)	방 승 찬 (한국전자통신연구원 원장)
	김 기 남 (삼성전자 종합기술원 회장)	안 승 권 (연암공과대학교 총장)
	박 성 옥 (SK하이닉스(주) 부회장)	전 영 현 (삼성SDI(주) 부회장)
	신 회 동 (한국전자기술연구원 원장)	최 창 식 (주DB하이텍 부회장)
	윤 석 진 (한국과학기술연구원 원장)	인 치 호 (세명대학교 교수)
	천 경 준 (주씨젠 회장)	김 종 옥 (고려대학교 교수) - 국제협력 / ICCE-Asia / 영문논문
감 사	백 광 현 (중앙대학교 교수) - ITC-CSCC	노 태 문 (한국전자통신연구원 센터장) - 연구소
부 회 장	강 문 식 (강릉원주대학교 교수) - 교육 / 표준화	심 동 규 (광운대학교 교수) - SPC 영문지
	노 원 우 (연세대학교 교수) - 하계 및 추계학술 총괄 / 국문논문	이 용 옥 (한화시스템 부사장) - 산학연
	류 수 정 (주사피온코리아 대표이사) - 회원 / 여성	이 재 훈 (유정시스템(주) 대표이사) - 산학연
	이 승 호 (한밭대학교 교수) - 지부	황 인 철 (경원대학교 교수) - 사업 / 대외협력
	이 재 관 (한국자동차연구원 소장) - 산학연	김 진 상 (경희대학교 교수) - 반도체
	정 영 모 (한성대학교 교수) - 홍보 총괄 / 정보화 / 학회지	송 병 철 (인하대학교 교수) - 인공지능신호처리
소사이머티 회장	유 명 식 (송실대학교 교수) - 통신	김 은 원 (대림대학교 교수) - 산업전자
	최 용 수 (신한대학교 교수) - 컴퓨터	강 성 원 (한국전자통신연구원 부원장)
	김 영 진 (한국생산기술연구원 수석연구원) - 시스템 및 제어	김 달 수 (주티엘아이 대표이사)
협동부회장	강 민 식 (LGI노텍(주) 부사장)	김 상 태 (한국산업기술평가관리원 연구위원)
	김 강 태 (삼성전자(주) 부사장)	김 형 준 (한국과학기술연구원 소장)
	김 부 군 (송실대학교 교수)	남 궁 선 (주유니트론텍 부회장)
	김 영 한 (UC San Diego / 가우스랩스 대표이사)	송 문 섭 (주심텍 회장)
	김 후 식 (주뷰웍스 대표이사)	오 윤 제 (정보통신기획평가원 PM)
	손 보 의 (주LX세미콘 대표이사)	이 광 영 (서경대학교 교수)
	엄 낙 응 (한국전자통신연구원 책임연구원)	이 병 선 (김포대학교 교수)
	유 창 동 (한국과학기술원 교수)	이 장 규 (주텔레칩스 대표이사)
	이 동 규 (주카카오빌리티 부사장)	전 병 우 (상관대학교 교수)
	이 서 규 (한국팹리스산업협회 회장)	정 준 (주솔리드 대표이사)
	이 창 한 (한국반도체산업협회 상근부회장)	
	전 선 익 (파이낸셜뉴스 회장)	
	정은 승 (삼성전자(주) 사장)	
상 임 이 사	강 명 곤 (한국교통대학교 교수) - 회원	강 석 주 (서경대학교 교수) - 재무
	구 본 태 (한국전자통신연구원 본부장) - 하계학술	권 영 수 (한국전자통신연구원 책임연구원) - 국제협력(ICCE-Asia)
	권 혁 인 (중앙대학교 교수) - 국제협력(ITC-CSCC)	김 동 순 (세종대학교 교수) - 대외협력(정책)
	김 수 연 (동국대학교 교수) - 홍보	김 용 신 (고려대학교 교수) - 국제협력 총괄(ICCE-Asia)
	김 원 중 (한국전자통신연구원 실장) - 표준화	김 익 균 (한국전자통신연구원 본부장) - 사업
	김 중 현 (고려대학교 교수) - 사업	김 재 준 (서울대학교 교수) - JSTS 영문지
	김 지 훈 (이화여자대학교 교수) - 기획	김 철 우 (고려대학교 교수) - 사업
	김 현 (서울과학기술대학교 교수) - 총무(대외협력)	김 훈 (인천대학교 교수) - 대외협력 총괄
	류 현 석 (서울대학교 교수) - 교육(산업체)	변 대 석 (삼성전자(주) 마스터) - 교육 총괄
	선 우 경 (서울대학교 교수) - 회원 총괄	손 교 민 (삼성전자(주) 마스터) - 산학연
	양 준 성 (연세대학교 교수) - 학회지	연 규 봉 (한국자동차연구원 팀장) - 산학연 총괄
	유 찬 세 (한국전자기술연구원 센터장) - 사업 총괄	이 남 윤 (고려대학교 교수) - 사업
	이 채 은 (인하대학교 교수) - 총무 총괄	장 익 준 (경희대학교 교수) - 하계학술
	전 세 영 (서울대학교 교수) - AI	정 진 곤 (중앙대학교 교수) - 국문논문
	제 민 규 (한국과학기술원 교수) - 사업	조 성 현 (한양대학교 교수) - 정보화
	최 기 장 (서울대학교 교수) - 산학연	최 병 호 (한국전자기술연구원 본부장) - 산학연
	최 재 혁 (한국과학기술원 교수) - 홍보	한 재 호 (고려대학교 교수) - 추계학술
산업체이사	강 석 판 (LG전자(주) 상무)	고 용 남 (하나미이크론(주) 전무)
	김 녹 원 (주딴엑스 대표이사)	김 동 현 (ICT(주) 대표이사)
	김 태 진 (주더즈텍 대표이사)	김 현 수 (삼성전자(주) 상무)
	배 순 민 (주케이티 연구소장)	오 의 열 (LG디스플레이(주) 연구위원)
	우 정 호 (비전넥스트(주) 대표이사)	원 제 형 (도쿄일렉트론코리아(주) 대표이사)
	윤 영 권 (삼성전자(주) 마스터)	이 도 훈 (국가보안기술연구소 부소장)
	이상 만 (주시스메이트 대표이사)	이상 훈 (주웨이브피아 대표이사)
	이수 민 (한국센서연구소 대표이사)	이수 인 (주텔레칩스 상무)
	조영 민 (SkyMirr CEO)	조혜 정 (삼성물산(주) 상무)
	천이우 (주넥스트칩 연구소장)	최성 민 (주해치텍 대표이사)

이 사

한은혜 (에스에스앤씨 대표이사)
 홍국태 (㈜X세미콘 연구위원)
 고병철 (계명대학교 교수) - 학술(하계)
 권기룡 (부경대학교 교수) - 학술(하계)
 김민규 (LGI노텍(주) 상무) - 산학연
 김성우 (서울대학교 교수) - 대외협력
 김유철 (LG AI연구원 부문장) - AI
 김재욱 (한국과학기술연구원 선임연구원) - 학술(하계)
 박성정 (건국대학교 교수) - 국제협력
 백종덕 (연세대학교 교수) - AI
 심현정 (한국과학기술원 교수) - AI
 안상철 (한국과학기술연구원 책임연구원) - AI
 원용욱 (명지대학교 교수) - 학술(추계)
 윤종윤 (주 파두 사장) - 교육
 이구순 (파이낸셜뉴스 부국장) - 홍보
 이창우 (가톨릭대학교 교수) - 홍보
 임동구 (전북대학교 교수) - 정보화
 장지원 (연세대학교 교수) - 학회지
 정무경 (㈜사피온코리아 CTO) - AI
 정해준 (경희대학교 교수) - 국문논문
 차혁규 (서울과학기술대학교 교수) - 정보화
 최광성 (한국전자통신연구원 실장) - 대외협력
 최정욱 (한양대학교 교수) - 학회지
 하태준 (광운대학교 교수) - 학술(추계)
 한정환 (충남대학교 교수) - 정보화
 한태희 (성균관대학교 교수) - 국문논문
 허재두 (한국전자통신연구원 책임연구원) - 사업
 황태호 (한국전자기술연구원 부문장) - 학술(하계)

협동이사

구민석 (인천대학교 교수) - 기획
 권경하 (한국과학기술원 교수) - 사업
 김대영 (순천향대학교 교수) - 호서지부
 김사혁 (KISDI 책임연구원) - 대외협력
 김승환 (한국전자통신연구원 책임연구원) - 학술(추계)
 김형진 (인하대학교 교수) - 회원
 민경식 (국민대학교 교수) - 회원
 박성욱 (강릉원주대학교 교수) - 강원지부
 배준호 (가천대학교 교수) - 표준화
 서종열 (LG전자(주) 그룹장) - 산학연
 송준영 (인천대학교 교수) - 홍보
 신세운 (UNIST 교수) - 사업
 안진호 (호서대학교 교수) - 학회지
 유경창 (삼성전자(주) 수석연구원) - 회원
 유호영 (충남대학교 교수) - 대전·충남지부
 윤상훈 (한국전자기술연구원 책임연구원) - 사업
 이권형 (LG전자(주) 책임연구원) - 대외협력
 이왕상 (경상국립대학교 교수) - 부산·경남·울산지부
 이재규 (삼성전자(주) 마스터) - 산학연
 이지훈 (전북대학교 교수) - 전북지부
 임승찬 (한경국립대학교 교수) - 국문논문
 정민채 (세종대학교 교수) - 국문논문
 정성엽 (차세대융합기술연구원 선임연구원) - 학술(하계)
 정재용 (인천대학교 교수) - 학회지
 채관엽 (삼성전자(주) 마스터) - 국제협력
 최강선 (한국기술교육대학교 교수) - SPC 영문지
 최웅 (숙명여자대학교 교수) - 회원
 추상현 (현대자동차(주) 책임매니저) - 정보화

함철희 (삼성전자(주) 마스터)

권구덕 (강원대학교 교수) - 기획
 권태수 (서울과학기술대학교 교수) - 정보화
 김선욱 (고려대학교 교수) - 교육
 김소영 (성균관대학교 교수) - SPC 영문지
 김윤 (서울시립대학교 교수) - 기획/회원
 동성수 (용인예술과학대학교 교수) - 교육
 배현철 (한국전자통신연구원 책임연구원) - 학술(추계)
 변영재 (UNIST 교수) - 교육
 안광호 (한국전자기술연구원 센터장) - 대외협력
 안호균 (한국전자통신연구원 실장) - 사업
 윤상민 (국민대학교 교수) - 학술(추계)
 이강윤 (성균관대학교 교수) - SPC 영문지
 이영택 (ASML 전무) - 교육
 이형민 (고려대학교 교수) - 대외협력
 장성욱 (㈜카카오모빌리티 부사장) - AI
 전동석 (서울대학교 교수) - 학술(하계)
 정일권 (한국전자통신연구원 부문장) - 산학연
 조성재 (이화여자대학교 교수) - 기획
 채영철 (연세대학교 교수) - 회원
 최영규 (인하대학교 교수) - 학술(하계)
 하정우 (네이버 AI연구소장) - AI
 한동국 (국민대학교 교수) - 사업
 한진호 (한국전자통신연구원 책임연구원) - 국제협력
 함범섭 (연세대학교 교수) - 학술(하계)
 황진영 (한국항공대학교 교수) - 홍보

권건우 (홍익대학교 교수) - 학술(하계)
 권종원 (한국산업기술시험원 책임연구원) - 학술(추계)
 김범현 (한양대학교 교수) - 정보화
 김시준 (강원대학교 교수) - 사업
 김정석 (가천대학교 교수) - 국제협력
 류성주 (숭실대학교 교수) - 학술(하계)
 박관서 (연세대학교 교수) - 학술(추계)
 배준성 (강원대학교 교수) - 정보화
 서민재 (가천대학교 교수) - 기획/홍보
 송익현 (한양대학교 교수) - 회원
 송철 (DGIST 교수) - 학술(추계)
 심용 (중앙대학교 교수) - 홍보
 오윤호 (고려대학교 교수) - 학술(추계)
 유동훈 (디시일로 연구소장) - 회원
 윤명국 (이화여자대학교 교수) - AI
 윤희인 (UNIST 교수) - 홍보
 이성학 (경북대학교 교수) - 대구·경북지부
 이인영 (조선대학교 교수) - 광주·전남지부
 이정원 (서울대학교 교수) - 회원/사업
 임매순 (한국과학기술연구원 책임연구원) - 사업
 장지수 (삼성전자(주) PE) - 사업
 정방철 (충남대학교 교수) - 정보화
 정완영 (한국과학기술원 교수) - 홍보
 차철웅 (한국전자기술연구원 센터장) - 표준화
 채주형 (광운대학교 교수) - 총무
 최병수 (부경대학교 교수) - 교육
 추민성 (한양대학교 교수) - 학술(하계)/홍보
 황원준 (아주대학교 교수) - AI

지부장 명단

강원지부	강문식 (강릉원주대학교 교수)	광주·전남지부	최수일 (전남대학교 교수)
대구·경북지부	공성호 (경북대학교 교수)	대전·충남지부	김철영 (충남대학교 교수)
부산·경남·울산지부	고진환 (경상대학교 교수)	전북지부	이주연 (전주비전대학교 교수)
제주지부	고석준 (제주대학교 교수)	충북지부	최영규 (한국교통대학교 교수)
호서지부	강윤희 (백석대학교 교수)	일본	백인천 (AIZU대학교 교수)
미국	최명준 (텔레다인 박사)	러시아지부	Prof. Edis B. TEN (National University of Science and Technology)

위원회 명단

자문위원회

위원장 김도현 (국민대학교 명예교수)
 부위원장 박항구 (소암시스텔 회장)
 위원 고성제 (고려대학교 교수)
 김덕진 (고려대학교 명예교수)
 김영권 (건국대학교 명예교수)
 문영식 (한양대학교 교수)
 박진옥 (육군사관학교 명예교수)
 성광모 (서울대학교 명예교수)
 이상실 (한양대학교 명예교수)
 이충웅 (서울대학교 명예교수)
 임혜숙 (이화여자대학교 교수)
 정정화 (한양대학교 명예교수)

기획위원회

위원장 김지훈 (이화여자대학교 교수)
 부위원장 조성재 (이화여자대학교 교수)
 위원 권구덕 (강원대학교 교수)
 김윤 (서울시립대학교 교수)

학술연구위원회 - 하계

위원장 노원우 (연세대학교 교수)
 부위원장 장익준 (경희대학교 교수)
 위원 고병철 (계명대학교 교수)
 김윤 (서울시립대학교 교수)
 변대석 (삼성전자(주) 마스터)
 전동석 (서울대학교 교수)
 최영규 (인하대학교 교수)
 함법섭 (연세대학교 교수)

학술연구위원회 - 추계

위원장 노원우 (연세대학교 교수)
 위원 권종원 (한국산업기술시험원 책임연구원)
 기관서 (연세대학교 교수)
 오윤호 (고려대학교 교수)
 하태준 (광운대학교 교수)

논문편집위원회

위원장 정진곤 (중앙대학교 교수)
 부위원장 정해준 (경희대학교 교수)
 위원 강성복 (한국생산기술연구원 수석연구원)
 권구덕 (강원대학교 교수)
 김명선 (한성대학교 교수)
 김영로 (명지전문대학교 교수)
 김학구 (중앙대학교 교수)
 임민중 (동국대학교 교수)
 조성인 (동국대학교 교수)

국제협력위원회

위원장 김용신 (고려대학교 교수)
 부위원장 권영수 (한국전자통신연구원 책임연구원)
 위원 김정석 (가천대학교 교수)
 최웅 (숙명여대 교수)

대외협력위원회

위원장 김훈 (인천대학교 교수)
 부위원장 김성우 (서울대학교 교수)
 위원 김동순 (세종대학교 교수)

공준진(삼성전자공과대학교 주임교수) 김수중 (경북대학교 명예교수)
 김성대 (한국과학기술원 명예교수) 나정웅 (한국과학기술원 명예교수)
 김재희 (연세대학교 명예교수) 박성한 (한양대학교 명예교수)
 박규태 (연세대학교 명예교수) 박성욱 ((전) 과학기술부 장관)
 백준기 (중앙대학교 교수) 서정욱 (연세대학교 명예교수)
 윤중웅 (한국공학교육인증원 이사장) 이문기 (연세대학교 명예교수)
 이재홍 (서울대학교 명예교수) 이진구 (동국대학교 명예교수)
 이태원 (고려대학교 명예교수) 임제택 (한양대학교 명예교수)
 전국진 (서울대학교 명예교수) 전홍태 (중앙대학교 명예교수)
 홍대식 (연세대학교 교수) 홍승홍 (인하대학교 명예교수)

구민석 (인천대학교 교수) 서민재 (가천대학교 교수)

구분태 (한국전자통신연구원 본부장) 권기룡 (부경대학교 교수)
 권건우 (홍익대학교 교수) 류성주 (서강대학교 교수)
 김재욱 (KIST 그룹장) 이종호 (숭실대학교 교수)
 양준성 (연세대학교 교수) 최민석 (경희대학교 교수)
 정성엽 (차세대융합기술연구원 실장) 추민성 (한양대학교 교수)
 최정욱 (한양대학교 교수)
 황태호 (한국전자기술연구원 본부장)

한재호 (고려대학교 교수) 김진영 (광운대학교 교수)
 김승환 (한국전자통신연구원 책임연구원) 송철 (DGIST 교수)
 배현철 (한국전자통신연구원 책임연구원) 윤상민 (국민대학교 교수)
 원용욱 (명지대학교 교수)

강제원 (이화여자대학교 교수) 공규열 (한성대학교 교수)
 권종원 (한국산업기술시험원 책임연구원) 김기연 (한국산업기술시험원 박사)
 김상범 (한국폴리텍대학교 교수) 김선웅 (건국대학교 교수)
 김영선 (대림대학교 교수) 김태환 (한국항공대학교 교수)
 심규성 (한경대학교 교수) 유재준 (UNIST 교수)
 임승찬 (한경국립대학교 교수) 정민채 (세종대학교 교수)
 추민성 (한양대학교 교수)

김종욱 (고려대학교 교수) 채관엽 (삼성전자(주) 마스터)
 권혁인 (중앙대학교 교수)
 박성정 (건국대학교 교수)
 한진호 (한국전자통신연구원 책임연구원)

연규봉 (한국자동차연구원 책임연구원) 김원중 (한국전자통신연구원 실장)
 김사혁 (KISDI 책임연구원)

김형준 (한국전자통신연구원 센터장)	안광호 (한국전자기술연구원 센터장)	이권형 (LG전자㈜ 책임연구원)
이형민 (고려대학교 교수)	최광성 (한국전자통신연구원 실장)	한상욱 (삼성전자㈜ 책임연구원)
황성운 (가천대학교 교수)		

산학연협동위원회

위원장	연구봉 (한국자동차연구원 실장)	김원종 (한국전자통신연구원 실장)	서종열 (LG전자㈜ 그룹장)
위원	김민규 (LGI노텍㈜ 상무)	이용욱 (한화시스템 부사장)	이재관 (한국자동차연구원 소장)
	손교민 (삼성전자㈜ 마스터)	이재훈 (유정시스템 대표이사)	정일권 (한국전자통신연구원 본부장)
	이재규 (삼성전자㈜ 마스터)	최병호 (한국전자기술연구원 본부장)	
	최기창 (서울대학교 교수)		

회원관리위원회

위원장	선우경 (서울대학교 교수)	김형진 (인하대학교 교수)	민경식 (국민대학교 교수)
부위원장	강명곤 (한국교통대학교 교수)	송익현 (한양대학교 교수)	유경창 (삼성전자㈜ 수석연구원)
위원	김윤 (서울시립대학교 교수)	이상훈 (㈜웨이브피아 대표이사)	이정원 (서울대학교 교수)
	배종호 (국민대학교 교수)	최웅 (숙명여자대학교 교수)	
	유동훈 (디사일로 연구소장)		
	채영철 (연세대학교 교수)		

회지편집위원회

위원장	양준성 (연세대학교 교수)	박승창 (㈜유오씨 사장)	안진호 (호서대학교 교수)
위원	박관서 (연세대학교 교수)	장지원 (연세대학교 교수)	정재용 (인천대학교 교수)
	윤석현 (단국대학교 교수)	한태화 (연세대학교 의료원 팀장)	허준 (고려대학교 교수)
	최정욱 (한양대학교 교수)		

사업위원회

위원장(총괄)	유찬세 (한국전자기술연구원 센터장)	김중현 (고려대학교 교수)	김철우 (고려대학교 교수)
위원	김익균 (한국전자통신연구원 본부장)	이남운 (고려대학교 교수)	제민규 (한국과학기술원 교수)
	선우경 (서울대학교 교수)	김시준 (강원대학교 교수)	송준영 (인천대학교 교수)
위원	권경하 (한국과학기술원 교수)	안호균 (한국전자통신연구원 실장)	윤상훈 (한국전자기술연구원 책임연구원)
	신세운 (UNIST 교수)	임매순 (한국과학기술연구원 책임연구원)	장지수 (삼성전자㈜ PE)
	이정원 (서울대학교 교수)	허재두 (한국전자통신연구원 책임연구원)	
	한동국 (국민대학교 교수)		

교육연구위원회

위원장	강문식 (강릉원주대 교수)	동성수 (용인예술과학대학교 교수)	류현석 (서울대학교 교수)
부위원장	변대석 (삼성전자㈜ 마스터)	변영재 (UNIST 교수)	윤종윤 (㈜파두 사장)
위원	김선욱 (고려대학교 교수)	이후진 (한성대학교 교수)	최병수 (한국전자통신연구원 실장)
	박영우 (TEL 부사장)		
	이영택 (ASML 전무)		

홍보위원회

위원장	정영모 (한성대학교 교수)	최재혁 (한국과학기술원 교수)	심용 (중앙대학교 교수)
부위원장	김수연 (동국대학교 교수)	송준영 (인천대학교 교수)	이창우 (가톨릭대학교 교수)
위원	서민재 (가천대학교 교수)	이구순 (파이낸셜뉴스 부국장)	황진영 (한국항공대학교 교수)
	윤희인 (UNIST 교수)	추민성 (한양대학교 교수)	
	정완영 (한국과학기술원 교수)		

표준화위원회

위원장	김원종 (한국전자통신연구원 실장)	김성동 (서울과학기술대학교 교수)	박재영 (광운대학교 교수)
부위원장	연구봉 (한국자동차연구원 팀장)	이상근 (성균관대학교 교수)	이종묵 (SOL 대표)
위원	권기원 (성균관대학교 교수)	차철웅 (한국전자기술연구원 센터장)	한태수 (한국전자기술연구원 연구위원)
	배준호 (가천대학교 교수)		
	좌성훈 (서울과학기술대학교 교수)		

정보화위원회

위원장	조성현 (한양대학교 교수)	김범현 (한양대학교 교수)	김중현 (고려대학교 교수)
위원	권태수 (서울과학기술대학교 교수)	임동구 (전북대학교 교수)	정방철 (충남대학교 교수)
	배준성 (강원대학교 교수)	추상혁 (현대자동차 책임매니저)	한정환 (충남대학교 교수)
	차혁규 (서울과학기술대학교 교수)		

AI위원회

위원장	이 총 용 (연세대학교 교수)	전 세 영 (서울대학교 교수)	
위원	강 석 주 (서강대학교 교수)	김 유 철 (LG AI연구원 부부장)	백 종 덕 (연세대학교 교수)
	심 현 정 (한국과학기술원 교수)	안 상 철 (KIST 책임연구원)	윤 명 국 (이화여자대학교 교수)
	정 무 경 (㈜사피온코리아 CTO)	한 재 호 (고려대학교 교수)	황 원 준 (이주대학교 교수)

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 (경상대학교 교수)	공 성 호 (경북대학교 교수)	김 철 영 (충남대학교 교수)
	이 주 연 (전주비전대학교 교수)	최 수 일 (전남대학교 교수)	최 영 규 (한국교통대학교 교수)

선거관리위원회

위원장	이 재 홍 (서울대학교 명예교수)		
위원	강 석 주 (서강대학교 교수)	김 현 (서울과학기술대학교 교수)	김 지 훈 (이화여자대학교 교수)
	이 채 은 (인하대학교 교수)	장 의 준 (경희대학교 교수)	정 진 곤 (중앙대학교 교수)

포상위원회

위원장	최 천 원 (단국대학교 교수)		
위원	김 종 옥 (고려대학교 교수)	김 지 훈 (이화여자대학교 교수)	노 원 우 (연세대학교 교수)
	이 총 용 (연세대학교 교수)	황 인 철 (강원대학교 교수)	
위원 및 간사겸임	이 채 은 (인하대학교 교수)		

재정위원회

위원장	이 혁 재 (서울대학교 교수)		
위원	강 석 주 (서강대학교 교수)	박 성 한 (명예회장)	박 영 기 (㈜싸인텔레콤 대표이사)
	원 제 형 (도쿄일렉트론코리아㈜ 대표이사)	유 창 동 (한국과학기술원 교수)	이 윤 종 (전) (㈜DB하이텍 부사장)
	이 총 용 (연세대학교 교수)	인 치 호 (세명대학교 교수)	조 중 휘 (인천대학교 교수)

인사위원회

위원장	이 혁 재 (서울대학교 교수)		
위원	김 현 (서울과학기술대학교 교수)	강 석 주 (서강대학교 교수)	이 채 은 (인하대학교 교수)
	이 총 용 (연세대학교 교수)	전 세 영 (서울대학교 교수)	

JSTS 편집위원회

위원장	김 재 준 (서울대학교 교수)		
위원	강 인 만 (경북대학교 교수)	권 혁 인 (중앙대학교 교수)	김 상 범 (서울대학교 교수)
	김 소 영 (성균관대학교 교수)	김 재 준 (포항공과대학교 교수)	김 주 성 (한밭대학교 교수)
	김 지 훈 (이화여자대학교 교수)	남 일 구 (부산대학교 교수)	민 경 식 (국민대학교 교수)
	박 성 주 (한양대학교 교수)	백 광 현 (중앙대학교 교수)	신 민 철 (한국과학기술원 교수)
	신 창 환 (고려대학교 교수)	오 정 우 (연세대학교 교수)	이 가 원 (충남대학교 교수)
	이 강 윤 (성균관대학교 교수)	장 호 원 (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 재 덕 (한양대학교 교수)

SPC위원회

위원장	심 동 규 (광운대학교 교수)		
위원	강 석 주 (서강대학교 교수)	김 선 옥 (고려대학교 교수)	김 소 영 (성균관대학교 교수)
	김 영 민 (홍익대학교 교수)	김 원 준 (건국대학교 교수)	김 재 곤 (한국항공대학교 교수)
	김 종 옥 (고려대학교 교수)	박 철 수 (광운대학교 교수)	백 준 기 (중앙대학교 교수)
	서 용 호 (광운대학교 교수)	송 병 철 (인하대학교 교수)	유 양 모 (서강대학교 교수)
	이 강 윤 (성균관대학교 교수)	이 채 은 (인하대학교 교수)	전 병 우 (성균관대학교 교수)
	정 승 원 (동국대학교 교수)	조 남 익 (서울대학교 교수)	조 민 호 (고려대학교 교수)
	진 훈 (경기대학교 교수)	최 강 선 (한국기술교육대학교 교수)	황 원 준 (이주대학교 교수)
	황 인 철 (강원대학교 교수)		

Society 명단

통신소사이어티

회장	유명식 (송실대학교 교수)	김재현 (아주대학교 교수)	김진영 (광운대학교 교수)
부회장	김선용 (건국대학교 교수)	오정근 (㈜ATNS 대표이사)	유명식 (송실대학교 교수)
	홍현 (인천대학교 교수)	이정우 (중원대학교 교수)	최천엽 (단국대학교 교수)
	유명석 (단국대학교 교수)		
	해준 (고려대학교 교수)		
감사	이재진 (송실대학교 교수)	이흥노 (광주과학기술원 교수)	김영한 (송실대학교 교수)
합동부회장	김병남 (에이스테크놀로지 연구소장)	김연은 (㈜브로던 대표이사)	류승문 ((사)개인공간서비스협회 수석부회장)
	김용석 (㈜답스 대표이사)	김인경 (LG전자㈜ 상무)	연철흠 (LG텔레콤 상무)
	박용석 (㈜ICT 대표이사)	방승찬 (한국전자통신연구원 부장)	정진섭 (이노와이어리스 부사장)
	이승호 (㈜하이게인 부사장)	이재훈 (유정시스템㈜ 대표이사)	
	정현규 (한국전자통신연구원 부장)		
이사	김광순 (연세대학교 교수)	김성훈 (한국전자통신연구원 박사)	김정호 (이화여자대학교 교수)
	노윤섭 (한국전자통신연구원 박사)	방성일 (단국대학교 교수)	서철현 (송실대학교 교수)
	상원진 (서강대학교 교수)	신오순 (송실대학교 교수)	신요안 (송실대학교 교수)
	윤종호 (한국항공대학교 교수)	윤지훈 (서울과학기술대학교 교수)	이재훈 (동국대학교 교수)
	이종호 (송실대학교 교수)	이호경 (홍익대학교 교수)	임종태 (홍익대학교 교수)
	장병수 (이노벨루네트웍스 부사장)	조성현 (한양대학교 교수)	조인호 (에이스테크놀로지 박사)
	최진식 (한양대학교 교수)	허서원 (홍익대학교 교수)	
연구위원	최지웅 (DGIST 교수) - 통신	윤상민 (국민대학교 교수) - 지능형네트워크	
	조춘식 (한국항공대학교 교수) - 마이크로파 및 전파전자	이철기 (아주대학교 교수) - ITS	
	김강욱 (경북대학교 교수) - 군사전자	허재두 (한국전자통신연구원 본부장) - 무선 PAN/BAN	
	김중현 (고려대학교 교수)		

반도체소사이어티

회장	김진상 (경희대학교 교수)	권오경 (한양대학교 교수)	김영환 (포항공과대학교 교수)
자문위원	공준진 (삼성전자공과대학교 주임교수)	김희석 (청주대학교 교수)	박홍준 (포항공과대학교 교수)
	김재석 (연세대학교 교수)	손보익 (㈜LX세미콘 대표이사)	신윤승 (반소 전임회장)
	신우영 (아주대학교 교수)	우남성 (반소 전임회장)	이승훈 (서강대학교 교수)
	신현철 (한양대학교 교수)	임형규 (반소 전임회장)	장성진 (삼성전자㈜ 부사장)
	임신일 (서경대학교 교수)	정연모 (경희대학교 교수)	정항근 (전북대학교 교수)
	전영현 (삼성SDI㈜ 부회장)	조경순 (한국외국어대학교 교수)	조상복 (울산대학교 교수)
	정해수 (Synopsis 사장)	최기영 (서울대학교 교수)	최승중 (LG전자㈜ 부사장)
	조중희 (인천대학교 교수)		
	허영범 (실리콘마이트스 대표이사)		
감사	이강윤 (성균관대학교 교수)	이광엽 (서경대학교 교수)	
부회장	김동규 (한양대학교 교수)	안기현 (한국반도체산업협회 전무)	이한호 (인하대학교 교수)
	이희덕 (충남대학교 교수)	장성진 (삼성전자㈜ 부사장)	최중호 (서울시립대학교 교수)
총무이사	고형호 (충남대학교 교수)	김지훈 (이화여자대학교 교수)	류연석 (서울대학교 교수)
	박종선 (고려대학교 교수)	윤찬호 (삼성전자 마스터)	황상준 (삼성전자㈜ 부사장)
편집이사	노정진 (한양대학교 교수)	유창식 (삼성전자 부사장)	조성재 (가천대학교 교수)
	한태희 (성균관대학교 교수)		
학술이사	강진구 (인하대학교 교수)	김철우 (고려대학교 교수)	범진욱 (서강대학교 교수)
	변영재 (UNIST 교수)	송민규 (동국대학교 교수)	이병훈 (포항공과대학교 교수)
	이승호 (한밭대학교 교수)	이희재 (서울대학교 교수)	이희덕 (충남대학교 교수)
	인천호 (세명대학교 교수)	정진균 (전북대학교 교수)	차호영 (홍익대학교 교수)
	최우영 (연세대학교 교수)	최창환 (한양대학교 교수)	
사업이사	강우병 (삼성전자㈜ 마스터)	공배선 (성균관대학교 교수)	공정택 (성균관대학교 교수)
	김동순 (세종대학교 교수)	김소영 (성균관대학교 교수)	김시호 (연세대학교 교수)
	김용석 (성균관대학교 교수)	김원중 (한국전자통신연구원 실장)	김중선 (홍익대학교 교수)
	백광현 (중앙대학교 교수)	변대석 (삼성전자㈜ 마스터)	손교민 (삼성전자㈜ 마스터)
	송용호 (삼성전자㈜ 부사장)	엄낙웅 (한국전자통신연구원 연구위원)	오정우 (연세대학교 교수)
	이강윤 (성균관대학교 교수)	조태제 (삼성전자㈜ 고문)	최규명 (서울대학교 교수)
	최병호 (한국전자기술연구원 본부장)	최윤경 (고려대학교 교수)	최준민 (경북대학교 교수)
재무이사	권기원 (성균관대학교 교수)	이성수 (송실대학교 교수)	
산학이사	김경수 (넥스트칩 대표이사)	김동현 (CTK 사장)	김보은 (라온텍 사장)
	김준석 (ADT 사장)	나준호 (㈜LX세미콘 전무)	손재철 (어보브반도체 부사장)
	송태훈 (휴인스 사장)	신용석 (케이던스코리아 사장)	이도영 (옵토레인 사장)
	이윤중 (동부하이텍 부사장)	이장규 (텔레칩스 대표이사)	
회원이사	노원우 (연세대학교 교수)	문웅 (송실대학교 교수)	
연구위원	최우영 (서울대학교 교수) - 반도체소재	김상인 (아주대학교 교수) - 광파밍양자전자공학	
	김중선 (홍익대학교 교수) - SoC설계	김영진 (한국항공대학교 교수) - RF집적회로	
	정원영 (강원공업㈜ 본부장) - PCB&Package	김익균 (한국전자통신연구원 본부장) - 정보보안시스템	
	정이준 (경희대학교 교수) - 내방사선 반도체 설계 및 소자	김한규 (삼성전자공과대학교 교수) - ESD/EOS & Latchup	
	노원우 (연세대학교 교수) - 인 메모리 컴퓨팅		
협동위원	강명근 (한국교통대학교 교수)	강석형 (포항공과대학교 교수)	권구덕 (강원대학교 교수)
	권영수 (한국전자통신연구원 책임연구원)	김수연 (동국대학교 교수)	김영민 (홍익대학교 교수)
	김재욱 (KIST 그룹장)	김현 (서울과학기술대학교 교수)	류성주 (송실대학교 교수)
	박성정 (건국대학교 교수)	송준영 (인천대학교 교수)	양준성 (연세대학교 교수)
	오윤호 (성균관대학교 교수)	윤명국 (이화여자대학교 교수)	이영주 (포항공과대학교 교수)
	이우주 (중앙대학교 교수)	이윤명 (성균관대학교 교수)	이영민 (고려대학교 교수)
	전동석 (서울대학교 교수)	전성훈 (삼성전자㈜ 상무)	정무경 (㈜사피온코리아 CTO)
	정윤호 (한국항공대학교 교수)	제민규 (KAIST 교수)	채형일 (건국대학교 교수)
	최재혁 (KAIST 교수)	최재혁 (성균관대학교 교수)	추민성 (한양대학교 교수)
	한정환 (충남대학교 교수)	황태호 (한국전자기술연구원 센터장)	김태환 (한국항공대학교 교수)

컴퓨터소사이터티

회 명	장 회 장	최용수 (신한대학교 교수)	김승천 (한성대학교 교수)	김형중 (고려대학교 교수)
자 문 위 원	사 회 장	김민우 (강릉원주대학교 교수)	박주영 (한국교통대학교 교수)	신인철 (단국대학교 명예교수)
부 회 장	부 회 장	박영민 (홍익대학교 교수)	안현식 (동명대학교 교수)	이규대 (공주대학교 교수)
협 동 부 회 장	부 회 장	정영준 (가천대학교 교수)	허영 (스마트의료기기산업진흥재단 부이사장)	홍유식 (상지대학교 교수)
총 무 이 사	부 회 장	김민우 (강릉원주대학교 교수)	남상엽 (국제대학교 교수)	정교일 (한국전자통신연구원 책임연구원)
재 무 이 사	부 회 장	김민우 (강릉원주대학교 교수)	박수현 (국립대학교 교수)	윤은준 (경일대학교 교수)
보 호 이 사	부 회 장	김민우 (강릉원주대학교 교수)	심정연 (강남대학교 교수)	김영학 (산업기술평가관리원 본부장)
집 이 사	부 회 장	김민우 (강릉원주대학교 교수)	김병서 (홍익대학교 교수)	유성철 (LG히다씨 본부장)
학 술 이 사	부 회 장	김민우 (강릉원주대학교 교수)	우운택 (한국과학기술원 교수)	정민성 (홍익대학교 교수)
사 업 이 사	부 회 장	김민우 (강릉원주대학교 교수)	이후진 (한성대학교 교수)	김선욱 (고려대학교 교수)
산 학 이 사	부 회 장	김민우 (강릉원주대학교 교수)	기장근 (공주대학교 교수)	김한울 (서울과학기술대학교 교수)
연구회위원장	부 회 장	김민우 (강릉원주대학교 교수)	김천식 (세종대학교 교수)	윤상훈 (한국전자기술연구원 책임연구원)
	부 회 장	김민우 (강릉원주대학교 교수)	김규성 (한경대학교 교수)	이세호 (전북대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	이민호 (경북대학교 교수)	정해명 (김포대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	이찬수 (영남대학교 교수)	김중윤 (경동대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	김명선 (한성대학교 교수)	이정선 (울지대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	서민석 (고려대학교 교수)	임재균 (명지병원 소장)
	부 회 장	김민우 (강릉원주대학교 교수)	임경원 (대림대학교 교수)	한태화 (연세대학교 의료원 팀장)
	부 회 장	김민우 (강릉원주대학교 교수)	한영선 (부경대학교 교수)	황진영 (한국항공대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	황재정 (군산대학교 교수)	김흥선 (연세세브란스병원 연구원)
	부 회 장	김민우 (강릉원주대학교 교수)	차시호 (청운대학교 교수)	신동희 (대보정보통신 부장)
	부 회 장	김민우 (강릉원주대학교 교수)	지은영 (투와이스스템즈 이사)	이학준 (이노지이스코리아 연구소장)
	부 회 장	김민우 (강릉원주대학교 교수)	송치봉 (웨이버스 이사)	진훈 (경기대학교 교수)
	부 회 장	김민우 (강릉원주대학교 교수)	이재홍 (유비벨록스모바일 대표이사)	진훈 (경기대학교 교수) - 휴먼ICT
	부 회 장	김민우 (강릉원주대학교 교수)	조병영 (㈜태진인포텍 전무)	이민호 (경북대학교 교수) - 인공지능/신경망/퍼지
	부 회 장	김민우 (강릉원주대학교 교수)	김민우 (강릉원주대학교 교수)	우운택 (한국과학기술원 교수) - 중강휴먼
	부 회 장	김민우 (강릉원주대학교 교수)	김민우 (강릉원주대학교 교수)	김명선 (한성대학교 교수) - AI응용

인공지능 신호처리소사이터티

회 명	장 회 장	송병철 (인하대학교 교수)	김중욱 (고려대학교 교수)	김창익 (한국과학기술원 교수)
자 문 위 원	부 회 장	김정태 (이화여자대학교 교수)	박종민 (한양대학교 교수)	심동규 (광운대학교 교수)
부 회 장	부 회 장	김영민 (광주과학기술원 교수)	전병우 (성균관대학교 교수)	조담익 (서울대학교 교수)
협 동 부 회 장	부 회 장	홍민철 (숭실대학교 교수)	전세영 (서울대학교 교수)	정승원 (고려대학교 교수)
이 사	부 회 장	김진진 (계명대학교 교수)	권기홍 (부경대학교 교수)	김담수 (서울대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김창수 (고려대학교 교수)	백준기 (중앙대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김재우 (LG전자 연구위원)	이병욱 (이화여자대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	지인호 (홍익대학교 교수)	최강선 (한국기술교육대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	하정우 (네이버 AI연구소장)	한재준 (삼성전자(주) 마스터)
	부 회 장	김진진 (계명대학교 교수)	강제원 (이화여자대학교 교수)	고영준 (충남대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	고현석 (한양대학교 교수)	곽노준 (서울대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	곽수하 (포스텍 교수)	구본학 (LG전자 연구원)
	부 회 장	김진진 (계명대학교 교수)	김준모 (한국과학기술원 교수)	김진규 (고려대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김학구 (중앙대학교 교수)	김한울 (서울과학기술대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	박상현 (DGIST 교수)	박영명 (이화여자대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	박철수 (광운대학교 교수)	백승덕 (연세대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	심재영 (UNIST 교수)	심현정 (연세대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	우상민 (한국기술교육대학교 교수)	우종택 (한국과학기술원 교수)
	부 회 장	김진진 (계명대학교 교수)	유재준 (UNIST 교수)	유준빈 (아주대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	유정인 (한화디펜스 연구원)	이금하 (KT 연구원)
	부 회 장	김진진 (계명대학교 교수)	이범식 (조선대학교 교수)	이상철 (인하대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	이철 (동국대학교 교수)	임성훈 (DGIST 교수)
	부 회 장	김진진 (계명대학교 교수)	전세영 (서울대학교 교수)	정근일 (한국전자통신연구원 연구전문위원)
	부 회 장	김진진 (계명대학교 교수)	정희철 (경북대학교 교수)	조동현 (충남대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	조성현 (포스텍 교수)	차영우 (고려대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	최상준 (고려대학교 교수)	최욱 (인천대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	최종현 (연세대학교 교수)	최해철 (한밭대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	한재호 (고려대학교 교수)	함범섭 (연세대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	황효석 (경희대학교 교수)	곽노준 (서울대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김현수 (충북대학교 교수)	권구락 (조선대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	구형일 (아주대학교 교수)	김동현 (연세대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김기백 (숭실대학교 교수)	김용환 (한국전자기술연구원 수석연구원)
	부 회 장	김진진 (계명대학교 교수)	김진주 (연세대학교 교수)	김성진 (계명대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김재민 (한국항공대학교 교수)	박진우 (경성대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	김중우 (한국항공우주연구원 연구원)	박구만 (서울과학기술대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	남승우 (한국전자통신연구원 연구원)	박성훈 (한국과학기술원 교수)
	부 회 장	김진진 (계명대학교 교수)	박상훈 (명지대학교 교수)	배성호 (경희대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	박홍준 (광운대학교 교수)	서진근 (연세대학교 교수)
	부 회 장	김진진 (계명대학교 교수)	서정일 (한국전자통신연구원 박사)	신재섭 (㈜픽스트리 대표이사)
	부 회 장	김진진 (계명대학교 교수)	송진호 (연세대학교 교수)	

간 사
중 무
회 위
원 장

신 지 태 (성균관대학교 교수)	신 학 준 (캐논메디칼시스템즈코리아 박사)	양 현 중 (UNIST 교수)
어 영 정 (연세대학교 교수)	임 일 규 (부산대학교 교수)	오 세 홍 (한국외국어대학교 교수)
우 태 진 (포스텍 교수)	유 일 동 (인텔리빅스 대표이사)	유 양 모 (서강대학교 교수)
윤 국 근 (한국과학기술원 교수)	유 일 동 (한국외국어대학교 교수)	이 기 승 (건국대학교 교수)
이 상 근 (중앙대학교 교수)	이 상 용 (연세대학교 교수)	이 상 현 (DGIST 교수)
이 상 훈 (연세대학교 교수)	이 승 용 (포스텍 교수)	이 연 정 (경북대학교 교수)
이 의 진 (서울과학기술대학교 교수)	이 장 원 (한국항공대학교 교수)	이 재 성 (서울대학교 교수)
이 중 석 (연세대학교 교수)	이 중 실 (한국전자기술연구원 책임연구원)	이 중 하 (계명대학교 교수)
이 준 재 (계명대학교 교수)	이 주 호 (성균관대학교 교수)	이 창 우 (가톨릭대학교 교수)
임 재 열 (한국기술교육대학교 교수)	임 재 윤 (제주대학교 교수)	임 종 우 (한양대학교 교수)
장 세 진 (한국전자기술연구원 센터장)	장 용 준 (제주한라대학교 교수)	장 인 수 (한국전자통신연구원 연구원)
전 진 원 (국가수리과학연구소 박사)	전 해 근 (광주과학기술원 교수)	정 미 라 (계명대학교 교수)
정 원 기 (고려대학교 교수)	정 찬 호 (연벌대학교 교수)	정 호 기 (한국교통대학교 교수)
정 원 열 (영남대학교 교수)	정 훈 (HDXMILL 박사)	조 승 롱 (한국과학기술원 교수)
조 유 민 (네이버 연구원)	진 현 성 (제주대학교 교수)	최 승 호 (서울과학기술대학교 교수)
지 시 식 (연세대학교 교수)	최 장 환 (이화여자대학교 교수)	최 준 원 (한양대학교 교수)
최 현 철 (연세대학교 교수)	한 중 기 (세종대학교 교수)	최 용 석 (이주대학교 교수)
홍 성 훈 (전남대학교 교수)	황 도 식 (연세대학교 교수)	황 성 주 (한국과학기술원 교수)
홍 배 (충북대학교 교수)		
김 회 용 (경희대학교 교수)	한 재 호 (고려대학교 교수)	
민 동 보 (이화여자대학교 교수)		
강 수 주 (서강대학교 교수) - 영상처리	김 원 준 (건국대학교 교수) - 영상이해	
이 조 호 (서울대학교 교수) - 바이오영상신호처리	황 원 준 (아주대학교 교수) - 딥러닝	
장 진 (경북대학교 교수) - 음향 및 음성신호처리	김 성 우 (서울대학교 교수) - 로봇 지능	

시스템 및 제어사이어티

회 장	김 영 진 (한국생산기술연구원 수석연구원)	주 영 복 (한국기술교육대학교 교수)	김 기 연 (한국산업기술시험원 선임연구원)
감 회 장	이 덕 진 (전북대학교 교수)	정 길 도 (전북대학교 교수)	
사	김 영 철 (군산대학교 교수)	강 성 복 (한국생산기술연구원 수석연구원)	
편집/학술이사	권 중 원 (한국산업기술시험원 책임연구원)		
홍 보 이사	최 현 태 (한국해양과학기술원 책임연구원)	문 태 주 (주부일하우징 대표이사)	신 대 현 (주대연씨앤아이 대표이사)
산 학 연 이사	김 광 식 (㈜소울아이티 전무)		
	유 별 기 (현대전기주 대표이사)		
회 원 이사	권 오 민 (충북대학교 교수)	김 수 찬 (한경대학교 교수)	김 종 만 (전남도립대학교 교수)
	김 지 훈 (전주비전대학교 교수)	김 호 철 (울지대학교 교수)	남 기 향 (동국대학교 교수)
	류 지 형 (한국전자통신연구원 박사)	문 정 호 (강릉원주대학교 교수)	박 명 진 (경희대학교 교수)
	박 재 영 (전북대학교 교수)	변 영 재 (UNIST 교수)	서 영 석 (영남대학교 교수)
	송 철 구 (전북대학교 교수)	여 희 주 (대전대학교 교수)	유 재 현 (한경대학교 교수)
	이 상 준 (선문대학교 교수)	이 수 열 (경희대학교 교수)	이 용 귀 (한국전자통신연구원 선임연구원)
	이 태 희 (전북대학교 교수)	이 학 상 (세종대학교 교수)	정 재 훈 (동국대학교 교수)
	최 수 변 (KISTI 연구원)	최 우 영 (전북대학교 교수)	박 이 (한국산업기술시험원 선임연구원)
자 문 위 원	김 덕 열 (연세대학교 교수)	최 희 식 (서울시립대학교 교수)	박 종 국 (경희대학교 교수)
	서 일 훈 (한양대학교 교수)	오 상 록 (KIST 부원장)	오 승 롱 (단국대학교 교수)
	오 창 현 (고려대학교 교수)	유 정 봉 (공주대학교 교수)	허 경 무 (단국대학교 교수)
연구회위원장	김 규 식 (서울시립대학교 교수) - 전력전자	한 수 희 (포항공과대학교 교수) - 제어계측	
	남 기 창 (동국대학교 교수) - 의용전자 및 생체공학	정 재 훈 (동국대학교 교수) - 지능로봇	
	이 성 준 (한양대학교 교수) - 회로 및 시스템	이 석 재 (대구보건대학교 교수) - 국방정보 및 제어	
	연 규 봉 (한국자동차연구원 센터장) - 자동차전자	오 창 현 (고려대학교 교수) - 의료영상시스템	
	권 원 철 (한국산업기술시험원 책임연구원) - 스마트팩토리	정 범 진 (서울과학기술대학교 교수) - 스마트미터링	

산업전자사이어티

회 장	김 은 원 (대림대학교 교수)	김 동 식 (인하공업전문대학 교수)	남 상 연 (국제대학교 교수)
명 예 회 장	김 창 수 (유한대학교 교수)	윤 기 방 (인천대학교 교수)	이 병 선 (김포대학교 교수)
	원 영 진 (부천대학교 교수)	이 원 석 (동양미래대학교 교수)	
자 문 위 원	이 상 회 (동서울대학교 교수)	김 용 민 (충청대학교 교수)	이 상 준 (수원과학대학교 교수)
	김 대 휘 (한국정보기술 대표이사)	최 영 일 (조선이공대학교 교수)	한 성 준 (아이티센 고문)
	조 규 남 (로봇신문사 대표이사)		
수 석 부 회 장	고 정 환 (인하공업전문대학 교수)	김 상 범 (한국폴리텍대학 교수)	김 영 로 (명지전문대학교 교수)
상 임 이사	김 남 섭 (서일대학교 교수)	김 현 (부천대학교 교수)	서 병 석 (상지대학교 교수)
	김 연 선 (대림대학교 교수)	임 우 용 (인하공업전문대학 교수)	원 우 현 (한국폴리텍대학 교수)
	서 춘 원 (K-My지능정보기술 대표이사)	이 시 현 (동서울대학교 교수)	조 도 현 (인하공업전문대학 교수)
협 동 상 임 이사	윤 중 현 (조선이공대학교 교수)	권 오 병 (㈜넷케이아이 전무이사)	김 세 중 (㈜SJ정보통신 이사)
	김 현 석 (로보웰코리아 대표이사)	김 정 석 (㈜오디에이테크놀로지 대표이사)	박 현 영 (㈜보교택 대표이사)
	김 용 철 (㈜울포랜드 이사)	서 승 현 (㈜글로발텔레콤 대표이사)	성 재 용 (㈜오픈링크시스템 대표이사)
	송 관 식 (㈜아이티웨이 상무이사)	송 광 현 (복두출판사 대표이사)	송 치 봉 (웨이버스 대표이사)
	유 성 철 (㈜클로센 상무이사)	이 영 준 (㈜투비콤 대표이사)	장 대 현 (㈜LGCS 상무이사)
	전 한 수 (㈜세립TSG 전무이사)	조 병 영 (㈜태진인포텍 대표이사)	조 한 일 (㈜하이제이컨설팅 대표이사)
	최 서 우 (㈜한국정보기술 상무이사)	Gemma 김 (엔티데이타 이사)	
이 사	김 희 훈 (여주대학교 교수)	구 자 일 (인하공업전문대학 교수)	권 오 상 (경기과학기술대학교 교수)
	권 유 중 (세명대학교 교수)	김 대 수 (전주비전대학교 교수)	김 덕 수 (동양미래대학교 교수)
	김 백 기 (강릉원주대학교 교수)	김 종 오 (동양미래대학교 교수)	김 대 원 (상지대학교 교수)
	김 필 중 (조선이공대학교 교수)	성 훈 석 (부천대학교 교수)	안 태 원 (동양미래대학교 교수)
	우 찬 일 (서일대학교 교수)	이 문 구 (김포대학교 교수)	이 정 재 (대림대학교 교수)
	이 용 구 (한림성심대학교 교수)	이 주 연 (전주비전대학교 교수)	장 기 동 (동양미래대학교 교수)
	정 성 수 (동양미래대학교 교수)		
협 동 이사	변 용 필 (㈜넷케이아이 전무이사)	고 강 일 (이지테크 대표이사)	곽 정 희 (㈜한국정보기술 부장)
	김 응 연 (㈜인터그래픽 대표이사)	김 철 동 (㈜클로센 책임연구원)	신 동 희 (주대보정보통신 부장)
	신 우 현 (시티랩스 이사)	오 재 곤 (㈜한국정보기술 상무이사)	이 경 일 (주동해종합기술공사 상무이사)
	이 병 건 (주오픈링크시스템 이사)	이 성 대 (시티랩스 이사)	이 승 민 (주튜시스템 대표이사)
	이 승 태 (주하나텍시스템 이사)	이 용 우 (주쌍용정보통신 상무이사)	이 지 학 (주송암시스템 대표이사)
	임 준 섭 (주대신정보통신 차장)	장 기 용 (주나날에스엠아이 부장)	장 철 (주우송정보기술 대표이사)
	정 민 우 (주대보정보통신 상무이사)	한 찬 석 (주동해종합기술공사 부사장)	
감 사	홍 성 수 (용인예술과학대학교 교수)	이 태 동 (국제대학교 교수)	

제24대 평의원 명단

- 강명곤 (한국교통대학교 교수)
 강석주 (서강대학교 교수)
 강성원 (한국전자통신연구원 부원장)
 강진구 (인하대학교 교수)
 고성제 (고려대학교 교수)
 고희호 (충남대학교 교수)
 공준진 (삼성전자공과대학교 주임교수)
 구분태 (한국전자통신연구원 본부장)
 권구덕 (강원대학교 교수)
 권오규 (인하대학교 교수)
 권태수 (서울과학기술대학교 교수)
 김경기 (대구대학교 교수)
 김규식 (서울시립대학교 교수)
 김대순 (전주비전대학교 교수)
 김도현 (제주대학교 교수)
 김동현 (CTK 대표이사)
 김병서 (홍익대학교 교수)
 김상원 (서강대학교 교수)
 김선욱 (고려대학교 교수)
 김성진 (울산과학기술원 교수)
 김수중 (경북대학교 명예교수)
 김승천 (한성대학교 교수)
 김영로 (명지전문대학교 교수)
 김영선 (대림대학교 교수)
 김영진 (한국생산기술연구원 수석연구원)
 김용규 (한국철도기술연구원 수석연구원)
 김원중 (한국전자통신연구원 책임연구원)
 김유철 (LG AI연구원 부부장)
 김익국 (한국전자통신연구원 본부장)
 김정범 (강원대학교 교수)
 김종욱 (고려대학교 교수)
 김지훈 (이화여자대학교 교수)
 김진태 (건국대학교 교수)
 김철영 (충남대학교 교수)
 김태진 (덕택 대표이사)
 김현 (부천대학교 교수)
 김형준 (한국과학기술연구원 소장)
 김홍국 (광주과학기술원 교수)
 남궁선 (유니트론텍 부회장)
 남일구 (부산대학교 교수)
 노태문 (한국전자통신연구원 센터장/책임연구원)
 류승택 (한국과학기술원 교수)
 문용 (송실대학교 교수)
 박규태 (연세대학교 명예교수)
 박성욱 (SK하이닉스 부회장)
 박수현 (국민대학교 교수)
 박종선 (고려대학교 교수)
 박항구 (소암시스템 회장)
 배현철 (한국전자통신연구원 책임연구원)
 백준기 (중앙대학교 교수)
 변영재 (울산과학기술원 교수)
 서창호 (한국과학기술원 교수)
 성광모 (서울대학교 명예교수)
 손보익 (LX세미콘 대표이사)
 송민규 (동국대학교 교수)
 송상현 (중앙대학교 교수)
 신오안 (송실대학교 교수)
 심동규 (광운대학교 교수)
 안승권 (연암공과대학교 총장)
 양준성 (연세대학교 교수)
 연규봉 (한국자동차연구원 팀장/수석연구원)
 오성근 (이주대학교 교수)
 오정훈 (삼성전자^주 마스터)
 우정호 (비전넥스트 대표이사)
 유명식 (송실대학교 교수)
 유찬세 (한국전자기술연구원 센터장)
- 강문식 (강릉원주대학교 교수)
 강석판 (LG전자 상무)
 강윤희 (백석대학교 교수)
 고병철 (계명대학교 교수)
 고정환 (인하공업전문대학 교수)
 공배선 (성균관대학교 교수)
 광진태 (고려대학교 교수)
 구용서 (단국대학교 교수)
 권기룡 (부경대학교 교수)
 권종기 (한국전자통신연구원 연구전문위원)
 권혁인 (중앙대학교 교수)
 김광수 (서강대학교 교수)
 김남 (충북대학교 교수)
 김덕진 (고려대학교 명예교수)
 김동규 (한양대학교 교수)
 김명선 (한성대학교 교수)
 김봉태 (한국전자통신연구원 소장)
 김상태 (한국산업기술평가관리원 연구위원)
 김성대 (한국과학기술원 명예교수)
 김소영 (성균관대학교 교수)
 김수찬 (한경대학교 교수)
 김시호 (연세대학교 교수)
 김영민 (서울대학교 교수)
 김영재 (해동과학문화재단 이사장)
 김영철 (군산대학교 교수)
 김용석 (성균관대학교 교수)
 김원준 (건국대학교 교수)
 김윤 (서울시립대학교 교수)
 김재현 (이주대학교 교수)
 김정호 (이화여자대학교 교수)
 김주성 (한밭대학교 교수)
 김진상 (경희대학교 교수)
 김창수 (고려대학교 교수)
 김철우 (고려대학교 교수)
 김한구 (삼성전자공과대학교 교수)
 김현수 (삼성전자^주 상무)
 김형진 (인하대학교 교수)
 김홍 (인천대학교 교수)
 남기창 (동국대학교 교수)
 노원우 (연세대학교 교수)
 동성수 (웅인송담대학교 교수)
 문병인 (경북대학교 교수)
 민경식 (국민대학교 교수)
 박부건 (포항공과대학교 교수)
 박성욱 (강릉원주대학교 교수)
 박영훈 (숙명여자대학교 교수)
 박종일 (한양대학교 교수)
 배준성 (강원대학교 조교수)
 백광현 (중앙대학교 교수)
 범진욱 (서강대학교 교수)
 서승우 (서울대학교 교수)
 선우경 (서울대학교 교수)
 성원진 (서강대학교 교수)
 손일수 (서울과학기술대학교 교수)
 송민협 (한국전자통신연구원 선임연구원)
 송준영 (인천대학교 교수)
 신창환 (고려대학교 교수)
 심정연 (강남대학교 교수)
 안현식 (동명대학교 교수)
 엄낙웅 (한국전자통신연구원 책임연구원)
 예종철 (한국과학기술원 교수)
 오윤호 (고려대학교 교수)
 오성민 (한국기술교육대학교 교수)
 원재형 (도쿄일렉트론코리아 대표이사)
 유윤섭 (한경대학교 교수)
 유창동 (한국과학기술원 교수)
- 강민석 (LGO노텍 부사장)
 강석형 (포항공과대학교 교수)
 강제원 (이화여자대학교 교수)
 고석준 (제주대학교 교수)
 고진환 (경상대학교 교수)
 공성호 (경북대학교 교수)
 구민석 (인천대학교 교수)
 권건우 (홍익대학교 교수)
 권오경 (한양대학교 석좌교수)
 권종원 (한국산업기술시험원 책임연구원)
 권호열 (정보통신정책연구원 원장)
 김광순 (연세대학교 교수)
 김달수 (티엘이이 대표이사)
 김도현 (국민대학교 명예교수)
 김동식 (인하공업전문대학 교수)
 김명준 (한국전자통신연구원 원장)
 김부균 (송실대학교 교수)
 김선용 (건국대학교 교수)
 김성우 (서울대학교 교수)
 김수연 (동국대학교 교수)
 김수환 (서울대학교 교수)
 김영권 (후레대학교 명예총장)
 김영민 (홍익대학교 교수)
 김영진 (한국항공대학교 교수)
 김영한 (UC San Diego/기우스랩스 교수/대표이사)
 김용신 (고려대학교 교수)
 김유선 (LGO노텍 TASK LEADER)
 김은원 (대림대학교 교수)
 김재희 (연세대학교 명예교수)
 김종선 (홍익대학교 교수)
 김준모 (한국과학기술원 교수)
 김진영 (광운대학교 교수)
 김창익 (한국과학기술원 교수)
 김태욱 (연세대학교 교수)
 김현 (서울과학기술대학교 교수)
 김현철 (울산대학교 교수)
 김형택 (홍익대학교 교수)
 나정웅 (한국과학기술원 명예교수)
 남상욱 (서울대학교 교수)
 노정진 (한양대학교 교수)
 류수정 (사피온코리아 대표이사)
 문영식 (한양대학교 교수)
 민동보 (이화여자대학교 교수)
 박성민 (이화여자대학교 교수)
 박성한 (한양대학교 명예교수)
 박인규 (인하대학교 교수)
 박진욱 (육군사관학교 명예교수)
 배준호 (가천대학교 교수)
 백상현 (고려대학교 교수)
 변대석 (삼성전자^주 마스터)
 서정욱 (전 과학기술부 장관)
 선우영훈 (이주대학교 교수)
 손교민 (삼성전자^주 마스터)
 송문섭 (삼텍 회장)
 송병철 (인하대학교 교수)
 신오순 (송실대학교 교수)
 신현철 (광운대학교 교수)
 안상철 (한국과학기술연구원 책임연구원)
 안호균 (한국전자통신연구원 실장)
 여희주 (대진대학교 교수)
 오상록 (한국과학기술연구원 강릉분원장)
 오의열 (LG디스플레이 연구위원)
 우윤택 (한국과학기술원 교수)
 유동훈 (삼성전자^주 전문)
 유정봉 (공주대학교 교수)
 유창식 (삼성전자^주 전문)

- | | | |
|-----------------------|------------------------|---------------------------|
| 윤광섭 (인하대학교 교수) | 윤명국 (이화여자대학교 교수) | 윤상훈 (한국전자기술연구원 책임연구원) |
| 윤석진 (한국과학기술연구원 원장) | 윤석현 (단국대학교 교수) | 윤성로 (서울대학교 교수) |
| 윤영걸 (삼성전자㈜ 마스터) | 윤일구 (연세대학교 교수) | 윤종용 (한국공학교육인증원 이사장) |
| 이강윤 (성균관대학교 교수) | 이경중 (연세대학교 교수) | 이광엽 (서경대학교 교수) |
| 이규대 (공주대학교 교수) | 이규복 (한국전자기술연구원 부원장) | 이규필 (삼성전자㈜ 부사장) |
| 이남윤 (고려대학교 교수) | 이덕기 (연암공과대학교 교수) | 이덕진 (전북대학교 교수) |
| 이동규 (카카오모빌리티 부사장) | 이문기 (연세대학교 명예교수) | 이병선 (김포대학교 교수) |
| 이상만 (시스메이트 대표이사) | 이상설 (한양대학교 명예교수) | 이상윤 (연세대학교 교수) |
| 이상훈 (웨이브피아 대표이사) | 이석호 (한국전자통신연구원 책임연구원) | 이석희 (SK하이닉스 대표이사) |
| 이성수 (숭실대학교 교수) | 이성준 (한양대학교 교수) | 이수민 (한국센서연구소 대표이사) |
| 이승용 (포항공과대학교 교수) | 이승은 (서울과학기술대학교 교수) | 이성호 (한밭대학교 교수) |
| 이영렬 (세종대학교 교수) | 이윤식 (울산과학기술원 교수) | 이인규 (고려대학교 교수) |
| 이재관 (한국자동차연구원 본부장) | 이재성 (고려대학교 교수) | 이재진 (숭실대학교 교수) |
| 이재홍 (서울대학교 명예교수) | 이재훈 (유정시스템 대표이사) | 이정우 (중앙대학교 교수) |
| 이정원 (서울대학교 선임연구원) | 이종호 (숭실대학교 교수) | 이종호 (서울대학교 교수) |
| 이중호 (서울대학교 교수) | 이주연 (전주비전대학교 교수) | 이진구 (동국대학교 석좌교수) |
| 이창한 (한국반도체산업협회 상근부회장) | 이재은 (인하대학교 교수) | 이천희 (전) 청주대학교 교수) |
| 이철 (동국대학교 교수) | 이철 (동국대학교 교수) | 이충용 (연세대학교 교수) |
| 이충용 (서울대학교 명예교수) | 이태동 (국제대학교 교수) | 이태원 (고려대학교 명예교수) |
| 이한림 (중앙대학교 교수) | 이한호 (인하대학교 교수) | 이혁재 (서울대학교 교수) |
| 이형민 (고려대학교 교수) | 이흥노 (광주과학기술원 교수) | 이희덕 (충남대학교 교수) |
| 인치호 (세명대학교 교수) | 임매순 (한국과학기술연구원 책임연구원) | 임신일 (서경대학교 교수) |
| 임재택 (한양대학교 명예교수) | 임혜숙 (이화여자대학교 교수) | 장길진 (경북대학교 교수) |
| 장석호 (건국대학교 교수) | 장성진 (삼성전자㈜ 부사장) | 장익준 (경희대학교 교수) |
| 전국진 (서울대학교 명예교수) | 전동석 (서울대학교 교수) | 전병우 (성균관대학교 교수) |
| 전선익 (파이낸셜뉴스 사장) | 전세영 (서울대학교 교수) | 전영현 (삼성SDI 부회장) |
| 전정훈 (성균관대학교 교수) | 전홍태 (중앙대학교 명예교수) | 정교일 (한국전자통신연구원 연구전문위원) |
| 정길도 (전북대학교 교수) | 정민수 (라운텍 부사장) | 정방철 (충남대학교 교수) |
| 정범진 (한국외국어대학교 교수) | 정성엽 (차세대융합기술연구원 선임연구원) | 정승원 (고려대학교 교수) |
| 정용규 (울지대학교 교수) | 정원영 (강원공업 본부장) | 정윤호 (한국항공대학교 교수) |
| 정은승 (삼성전자㈜ 사장) | 정일권 (한국전자통신연구원 본부장) | 정정화 (한양대학교 석좌교수) |
| 정준문 (연세대학교 교수) | 정준 (솔리드 대표이사) | 정진곤 (중앙대학교 교수) |
| 정진균 (전북대학교 교수) | 정해준 (경희대학교 교수) | 제민규 (한국과학기술원 교수) |
| 조경순 (한국외국어대학교 교수) | 조남익 (서울대학교 교수) | 조도현 (인하공업전문대학 교수) |
| 조민호 (고려대학교 교수) | 조성인 (동국대학교 교수) | 조성현 (한양대학교 교수) |
| 조영민 (SkyMirr CEO) | 조중휘 (인천대학교 교수) | 조진웅 (한국전자기술연구원 센터장/수석연구원) |
| 조현중 (강원대학교 교수) | 조해정 (삼성물산 그룹장) | 진훈 (경기대학교 교수) |
| 차철용 (한국전자기술연구원 센터장) | 차혁규 (서울과학기술대학교 교수) | 채영철 (연세대학교 교수) |
| 채주형 (광운대학교 교수) | 천경준 (씨젠 회장) | 최강선 (한국기술교육대학교 교수) |
| 최광성 (한국전자통신연구원 실장) | 최광표 (삼성전자㈜ 마스터) | 최병수 (부경대학교 교수) |
| 최병호 (한국전자기술연구원 본부장) | 최성민 (해치텍 대표이사) | 최수일 (전남대학교 교수) |
| 최승범 (삼성전자㈜ 부사장) | 최승종 (LG전자 부사장) | 최영규 (한국교통대학교 교수) |
| 최용수 (신한대학교 교수) | 최우영 (연세대학교 교수) | 최웅 (숙명여자대학교 교수) |
| 최윤석 (한밭대학교 교수) | 최재혁 (한국과학기술원 교수) | 최준림 (경북대학교 교수) |
| 최중호 (서울시립대학교 교수) | 최진성 (도이치텔레콤 부사장) | 최창범 (한밭대학교 교수) |
| 최창식 (DB하이텍 부회장) | 최천원 (단국대학교 교수) | 최현택 (한국해양과학기술원 책임연구원) |
| 한동석 (경북대학교 교수) | 한영선 (부경대학교 교수) | 한은혜 (에스에스앤씨 대표이사) |
| 한재호 (고려대학교 교수) | 한정환 (충남대학교 교수) | 한태희 (성균관대학교 교수) |
| 함범섭 (연세대학교 교수) | 함철희 (삼성전자㈜ 마스터) | 허재두 (한국전자통신연구원 책임연구원) |
| 허준 (고려대학교 교수) | 현유진 (대구경북과학기술원 책임연구원) | 홍국태 (LX세미콘 연구위원) |
| 홍대식 (연세대학교 교수) | 홍민철 (숭실대학교 교수) | 홍승홍 (인하대학교 명예교수) |
| 홍유식 (상지대학교 교수) | 홍인기 (경희대학교 교수) | 홍제형 (한양대학교 교수) |
| 홍철호 (중앙대학교 교수) | 황성운 (가천대학교 교수) | 황승훈 (동국대학교 교수) |
| 황원준 (아주대학교 교수) | 황인정 (명지병원 수석연구원) | 황인철 (경원대학교 교수) |
| 황인태 (전남대학교 교수) | 황진영 (한국항공대학교 교수) | |

사무국 직원 명단

- 송기원 국장 - 기획, 산학연, 신규 사업, 자문/서울IT포럼, 지부, 인사, 규정, 회장단 관련, 대외협력 및 업무총괄
- 이안수 부장 - 허계학술대회, 주요 운영회의(이사회, 평의원회 및 총회), 총무업무(선거, 공문수발, 임원관련, 송년회, 포상 및 Wiset 등), 산업전자소사이터
- 배지영 부장 - 국제학술대회 총괄(ITC-CSCC, ICEIC, ICCE-Asia), 사업(기술워크샵 등), 추계학술대회, 컴퓨터소사이터, 시스템 및 제어소사이터
- 배기동 부장 - AI 관련, 국문지, 학회지, 표준화, 운영, 기타 지원업무, 인공지능 신호처리소사이터
- 변은정 부장 - 본회/소사이터/연구회 재무, 회원관리(개인회원 및 특별회원), 홍보, 통신소사이터
- 김천일 차장 - 정보화 관련(웹사이트 관리 및 온라인/디지털 업무지원 등), 교육, 전산장비 관리, 반도체소사이터
- 성다희 서기 - 국제학술대회 담당(ITC-CSCC, ICEIC, ICCE-Asia), 외국 기관과 국제협력(Joint Award 등), JSTS 및 SPC 발간

학회소식

제1차 상임이사회 회의록

제 1차 상임이사회가 1월 6일(금) 오후 5시 AC호텔 서울 강남(2층)에서 개최되었다. 이번 회의 결과는 다음과 같다.

- 다 음 -

1. 성원 보고

- 40명의 상임이사 중 31명이 참석하여 성원되었음.

2. 위원회 보고

- 각 위원회 별로 위원장들이 서면 및 발표보고사항을 진행하고 위원회별 주요사항을 논의함.

3. 심의사항

- 신규회원 가입을 승인함.
- 주요 승인위원회 (재정위원회 / 선거관리위원회) 위원 선정에 대해 원안대로 승인함.

4. 기타

특별회원 및 유관기관 방문

학회 회장단은 1월~2월 중 한국전자기술연구원, 삼성전자 및 정보통신산업진흥원을 방문하여 현안과제를 논의하고 상호 협력방안을 모색하였다.



한국전자기술연구원 - 신희동 원장



삼성전자/메모리사업부 - 이정배 사장



정보통신산업진흥원 - 허성욱 원장

2023년 제1회 로봇지능 콜로키엄 - 사람과 로봇의 공진화

인공지능 신호처리소사이머티 로봇지능연구회(위원장 : 김성우 교수(서울대)) 개최로 “2023년 제1회 로봇지능 콜로키엄 - 사람과 로봇의 공진화”를 1월 17일(화) 웰리힐리 리조트 루비홀(본관5층)에서 개최되었다.

본 콜로키엄에서는 사람과 로봇의 공진화를 통해 다가올 미래에 대해 미리 전망하고 관련 기술의 연구개발 최전선에 있는 국내 대표 연구실과 연구 기관의 현황에 대해 듣고 토론하는 시간이 되었다. 참석은 약 50여명 이었다.

2022년 연구회 신설 후, 개최하는 첫 행사로, 이번 콜로키엄을 시작으로 연구회 활성화를 추진코자 준비중이다.

2023년 대한전자공학회 영상이해연구회 겨울학교

인공지능 신호처리소사이머티 영상이해연구회(위원장 : 김원준 교수(건국대)) 개최로 “2023년 대한전자공학회 영상이해연구회 겨울학교”를 1월 17일(화)~19일(목)까지 웰리힐리 리조트 루비홀(본관5층)에서 개최되었다.

이번 영상이해연구회 겨울학교에서는 컴퓨터 비전 분야에서 최근 널리 연구되고 있는 주제의 기초 이론부터 최신 연구 동향까지 폭넓은 주제로 진행되었다. 특히, 다양한 신경망 구조, 멀티모달 데이터 기반 응용, NeRF/Diffusion 모델, 3차원 모델링 등 많은 관심을 받고 있는 연구 주제로 심도있는 강연이 진행되었다. 참석은 약 160여명 이었다.



2023년 영상이해연구회 겨울학교 운영위원회 기념촬영



2023년 영상이해연구회 겨울학교 강연모습

ICEIC 2023 국제학술대회

ICEIC는 우리나라 정보통신기술 학문과 산업을 아시아 및 유라시아 지역에 전파할 목적으로 대한전자공학회에서 주관하는 학술행사다. 이전에는 ICEIC를 2년에 한 번씩 개최하였으나 2012년부터 IEEE CE Society의 Co-sponsorship을 받아 국제 학술대회로 육성하기 위해 매년 개최하고 있다.

금년에는 2월 5일(일)~2월 8일(수)에 싱가포르 상그릴라에서 성황리에 개최되었고, 전 세계 12개국에서 약 309편의 논문이 발표되었다. 이번 행사는 이혁재 학회장(서울대학교)이 General Chair, 송병철 교수(인하대학교)가 Organization Chair, 강석주 교수(서강대학교)가 Technical Program Chair로 참여하였으며, 많은 교수들이 조직위원회를 구성하여 행사 준비에 만전을 기하였다. 이번 학술행사는 코로나 19 이후 3년 만에 오프라인과 형태로 진행되었다.

첫째 날 2월 5일(일)에는 Welcome Reception에는 이혁재 학회장의 환영 인사와 함께 여러 국가로부터 온 약 50명의 참가자들이 환영 인사와 담소를 나누었다.

둘째 날 2월 6일(월) 개최식에는 이혁재 학회장의 환영사와 IEEE CE Society의 Wen-Chung Kao 회장의 환영사로 시작을 알리며, National University of Singapore의 학장 Aaron Thean 교수와 싱가포르 한국과학기술자협회 KSEASG의 회장 조남준 교수의 환영 인사와 더불어 KSEASG와 대한전자공학회의 MOU 체결식이 있었다. 뒤이어 National University of Singapore Aaron Thean 교수와 삼성전자 손영수 부사장의 Plenary Talk가 진행되었다. 같은 날, Tokyo University of Agriculture and Technology와 Sigron Inc. Toshihisa Tanaka 교수, MOLOCO 권재명 박사의 Keynote, A*STAR Institute for Infocomm Research Yan Wu 박사의 Tutorial과 함께 12개의 구두 발표 세션과 4개의 포스터 발표 세션이 진행되었다.

셋째 날 2월 7일(화)에는 KAIST 권인소 교수와 GLOBALFOUNDRIES 구정모 이사, IEEE Fellow Tony Q.S. Quek 교수의 Plenary Talk가 있었으며, National University of Singapore 유담 교수의 Keynote 강연과 경희대학교 김정욱 교수의 Tutorial 강연이 진행되었다. 그리고 9

개의 구두 발표 세션과 4개의 포스터 발표 세션이 진행되었다. 저녁에는 이혁재 학회장의 환영사로 만찬의 시작을 알렸으며, Best Paper 시상이 있었다. 이어서 대한전자공학회 명예 회장 이재홍 교수의 건배사로 본격적인 Conference Banquet이 진행되었다. 마지막 날 2월 8일(수)에는 폐회식을 끝으로 행사가 마무리되었다. 이번 ICEIC 2023은 세계적인 연구자들이 서로의 연구 성과를 공유할 수 있는 학술교류의 장이 되었을 뿐만 아니라, 향후 정보통신기술 분야의 연구와 새로운 기술 발전을 도모할 수 있는 기회를 제공하였다.



Welcome Reception



IEIE & KSEASG MOU 체결식



Plenary Talk 1 – University of Singapore Aaron Thean 교수



Opening Ceremony – 이혁재 회장 환영사



Plenary Talk 2 – 삼성전자 손영수 부사장



Opening Ceremony – IEEE CTSoc Wen-Chung Kao 회장 환영사



Plenary Talk



주요참가자 기념촬영 - 초청인사, 조직위원 및 명예회장



Best Paper Award



Conference Banquet

신규회원 가입현황

기간 : 2023년 1월 1일 - 1월31일

〈정회원〉

김진규(고려대학교), 이완호, 전기완(국가수리과학연구소), 송철(대구 경북과학기술원), 유현우(울산과학기술원), 임승찬(한경대학교),

이상 6명

〈평생회원〉

김민준(한국과학기술원)

이상 1명

〈학생회원〉

박지수, 유진우, 이도영, 장호민(대구경북과학기술원), Nghia Nguyen, Truong Hien, 정이식(서울대학교), 김기현, 손주현(이화여자대학교), 신승석(전남대학교), 김동근, 박채호(한국기술교육대학교), 김경준(한국외국어대학교)

이상 12명

학회일지

www.theieie.org

THE INSTITUTE OF
ELECTRONICS AND INFORMATION
ENGINEERS

(2023년 1월 16일 ~ 2월 15일)

1. 행사 개최

구분	행사명	기간	장소
로봇지능연구회	2023년 제1회 로봇지능 콜로키엄 - 사람과 로봇의 공진화 -	1.17	웰리힐리조트
영상이해연구회	2023년 대한전자공학회 영상이해연구회 겨울학교	1.17 - 19	웰리힐리조트
본회(국제협력)	ICEIC 2023 국제학술대회	2.5 - 8	싱가포르

인공지능 컴퓨팅 최적화 기술



김명선 편집위원
(한성대학교)

인공지능 응용의 성능 향상을 위해 딥러닝 모델 자체에 관련된 연구가 활발히 진행되고 있으나 딥러닝 모델을 효율적으로 실행할 수 있는 컴퓨팅 시스템 측면에서의 연구는 상대적으로 미흡하다. 갈수록 거대해지는 모델의 크기와 메모리 요구로 인하여 리소스가 제약된 시스템 환경에서 최신 딥러닝 모델을 적용하는 것은 매우 어려운 일이다. 본 특집호는 딥러닝 모델의 컴퓨팅에 있어서 최근 개발

되고 있는 최적화 기술들을 살펴본다. 다양한 분야의 전문가들이 총 5편의 논문을 기고하였으며, 이를 통해 인공지능을 최적으로 컴퓨팅할 수 있는 하드웨어부터 소프트웨어까지의 요소 기술들을 들여다 볼 수 있도록 특집호가 구성되었다.

첫째, “딥러닝 모델의 초매개변수 최적화(김신규)”에서는 초매개변수 최적화 기법(hyperparameter optimization, HPO)들을 소개한다. 딥러닝 모델의 학습 과정에서 학습을 수행하는 사람들의 전문적인 사전 지식에 의존하여 여러 초매개변수들을 설정한다. 이후 학습된 모델의 성능 결과치를 비교하여 더 좋은 성능을 달성할 수 있을 때까지 새로운 설정값을 경험치에 의존하면서 계속해서 바꿔가며 학습한다. 본 특집호에서는 이러한 비효율성을 극복할 수 있는 HPO 기술을 소개한다.

둘째, “트랜스포머 모델의 경량화(오지훈)”에서는 트랜스포머 기반 대규모 자연어 모델을 메모리와 연산량 측면에서 경량화하는 방법들을 소개한다. 최근 자연어 처리(NLP) 분야에서 트랜스포머(Transformer)를 활용한 모델들이 늘어나고 있으며 그 성능 또한 기존 모델들과 비교 시 매우 탁월하다. 하지만 수천억 개에 달하는 파라미터 수를 가지는 대규모 언어 모델들은 제한된 메모리 크기를 가지는 연산 유닛(GPU 등)에 로딩하기가 어렵고 대규모 컴퓨팅 인프라가 지원되지 않는 경우 학습에만 몇 달의 시간이 소요된다. 본 특집호에서는 이러한 이슈를 해결하기 방법들을 소개한다.

셋째, “AI Inference를 위한 하드웨어 플랫폼과 소프트웨어 프레임워크의 동향(이연복)”에서는 AI Inference 과정에 이용되고 있는 대표적인 하드웨어 플랫폼과 소프트웨어 프레임워크들을 소개한다. Inference는 데이터센터 내 서버나 IoT/엣지 디바이스 등에서 실행되고 각각의 실행 환경에서 처리 속도 혹은 저지연/저전력 등 가장 중요하게 추구하는 요소들이 조금씩 다를 수 있다. 따라서 NVIDIA와 Intel과 같은 선진사를 중심으로 최적 딥러닝 실행에 부합하는 다양한 하드웨어 플랫폼을 출시하고, 또한 각 하드웨어 제조사들은 해당 제품을 효율적으로 활용하기 위한 소프트웨어 프레임워크들을 제공한다. 본 특집호에서는 이러한 하드웨어/소프트웨어 솔루션들을 살펴본다.

넷째, “전자설계자동화 기술과 차세대 지능형반도체 소자 및 회로 설계 응용(정성엽)”에서는 AI 반도체 내부의 저지연/저전력 이슈를 해결하기 위한 차세대 지능형 반도체인 저항변화 현상 기반의 메모리 소자를 소개한다. 더불어 이를 활용하는 집적 회로에 대한 구체적인 전자설계자동화(electronic design automation) 사례에 대하여 살펴본다.

마지막으로 “Kubernetes 기반 오픈소스 AI/ML 플랫폼(박용석)”에서는 실제 상용 환경에서 AI 응용들의 효율적인 운용을 지원하기 위한 체계적인 기능을 갖춘 시스템에 대하여 살펴본다. AI 응용의 생명주기(Lifecycle)는 데이터 확보, ML (machine learning) 모델 설계 및 훈련, 모델 배포, 성능 관찰(monitors)로 구성된다. 하지만 이러한 생명주기를 진행할 때 전문인력의 확보 및 유지, AI 도구 및 인프라 환경, 프로젝트 운용에 있어서 복잡성 등은 AI 응용의 생명주기를 매우 느리게 만든다. 본 특집호에서는 AI 응용의 생명주기를 효율적으로 운용할 수 있는 Kubernetes 기반 오픈소스 AI/ML 플랫폼 기술들을 소개한다.

본 특집호를 위하여 바쁜 일정 중에도 기고문을 준비해주신 집필진 여러분께 감사의 말씀을 드린다. 본 특집호의 인공지능 컴퓨팅 최적화 기술들에 대한 기고문이 우리나라 인공지능 산업의 발전과 경쟁력 강화에 기여할 수 있기를 기원한다.

딥러닝 모델의 초매개변수 최적화

I. 서론

딥러닝 기술이 기존에 해결하기 어려웠던 문제들을 해결할 수 있음을 보이면서 점점 더 다양한 분야에서 딥러닝 기술을 활용하려는 시도가 늘어가고 있다. 이에 따라 다양한 모델과 학습 기법이 개발되었고 점점 더 높은 성능을 달성하게 되었다. 하지만, 딥러닝 모델이 좋은 성능을 낼 수 있도록 잘 학습시키는 것은 점점 더 어려운 문제가 되었다.

딥러닝 모델은 두 가지 유형의 매개변수 (parameter) 를 갖고 있다. 첫번째는 초매개변수 (hyperparameter) 인데, 이것은 학습을 시작하기 전에 미리 설정되어야 하는 값이다. 대표적인 hyperparameter에는 학습률 (learning rate) 과 배치 크기 (batch size) 등이 있다. 두 번째는 모델 매개변수 (model parameter) 로서 학습을 통해서 설정되는 값이다. 대표적인 model parameter에는 convolution의 weight 를 들 수 있다. 동일한 모델을 서로 다른 데이터셋으로 학습시킬 때 일반적으로 기존에 사용했던 hyperparameter를 그대로 재사용하는 경우가 많은데, 이는 최적의 hyperparameter를 선택하는 것이 어려운 문제이기 때문이다. 모델의 학습을 시작하기 전에 미리 설정해야하는 초매개변수(hyperparameter)의 갯수가 점점 더 늘어났을 뿐만 아니라, 좋은 설정값을 찾는 것은 온전히 모델을 학습시키는 사람의 경험과 능력에만 의지할 수 밖에 없었기 때문이다. 하지만, 이는 최고의 성능을 달성하기 위해서는 피해야할 습관이다.

최고의 성능을 달성할 수 있는 최적의 hyperparameter 설정값을 찾기 위해서는 많은 시간과 노력이 필요하다. 먼저, 모델을 학습시키는 사람의 사전 지식을 활용하여 몇 가지 설정값으로 학습을 수행한다. 그리고 학습된 모델의 결과를 비교하여 더 좋은 성능을 달성할 수 있을 것으로 예상되는 새로운 설정값을 선택하여 모델을 다시 학습시킨다.



김신규
Intel

새롭게 얻은 결과가 원하는 성능에 도달하지 못했다면, 지금까지의 결과들을 바탕으로 새로운 설정값을 선택하여 다시 학습을 시키는 작업을 반복하게 된다. 이와 같은 특성으로 인하여 최적의 hyperparameter를 선택하는 것은 실질적으로 매우 어려운 문제인 것이다.

이 문제를 해결하기 위해서 등장한 것이 초매개변수 최적화 기법 (hyperparameter optimization, 이하 HPO) 이다. HPO는 위와 같은 반복적인 작업을 사람을 대신하여 수행할 뿐만 아니라, 주어진 시간과 자원을 최대한 효율적으로 활용하여 최적의 hyperparameter 설정값을 찾는 것을 목표로 한다. 본고에서는 지금까지 소개된 HPO 기법들과 이를 구현한 응용을 소개하고, 앞으로 HPO가 풀어야 할 문제들을 제시한다.

II. HPO 알고리즘

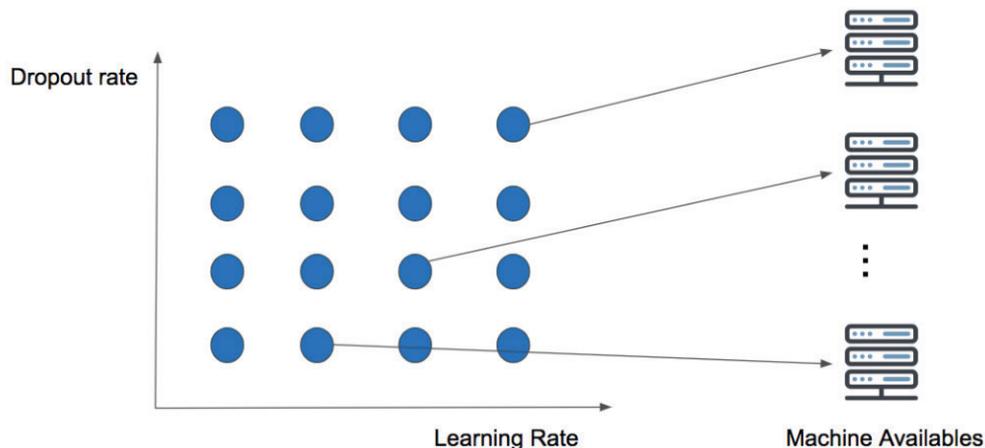
HPO 알고리즘은 크게 두 가지 부류로 나눌 수 있다. 하나는 black-box optimization이고, 다른 하나는 multi-fidelity optimization이다^[1]. 대부분의 HPO 알고리즘은 black-box optimization으로서 대상 모델의 특성을 전혀 모르는 상태에서 최적의 hyperparameter를 찾아내게 된다. 이 중에서 처음에는 신뢰성이 낮은 결과를 이용해서 hyperparameter를 걸러내다가 점점 더 신뢰성이 높은 결과를 이용해서 최적의 hyperparameter를 선택하는 방법을 별도로 multi-fidelity optimization

이라고 분류하였다. 본고에서는 Grid Search, Random Search, Bayesian Optimization, Population-based Training의 네 가지 black-box optimization과 Successive Halving Algorithm, HyperBand Algorithm, Asynchronous Successive Halving Algorithm, Bayesian Optimization HyperBand Algorithm의 네 가지 multi-fidelity optimization에 대해서 설명한다.

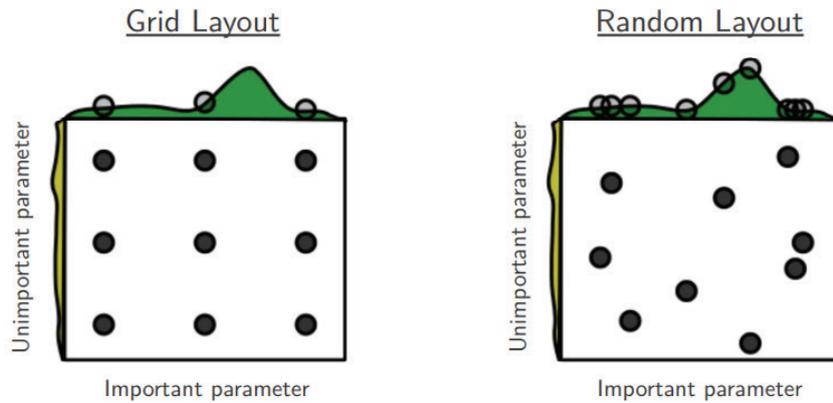
1. Grid Search

Grid Search는 가능한 모든 hyperparameter의 조합을 확인하는 방법이다. <그림 1>은 "Dropout rate"와 "Learning rate" 두 개의 hyperparameter에 대한 최적값을 찾는 Grid Search를 도식화한 것이다^[2]. 파란 점이 하나의 hyperparameter 설정값을 나타낸다. 각각 네 가지의 값을 가질 수 있다고 가정했을 때 총 16개의 설정이 가능하므로, 이 16개의 설정에 대해서 모두 학습을 해서 결과를 비교하면 최적의 hyperparameter를 찾을 수 있게 된다.

이 방법은 각각의 hyperparameter 설정값이 독립적이므로, <그림 1>과 같이 모델을 학습시킬 수 있는 계산 자원이 충분하다면 병렬적으로 실행할 수 있다는 특성이 있다. 하지만, 최적화해야 할 hyperparameter의 종류가 늘어나면 탐색할 설정값이 급격하게 늘어나기 때문에 확장성 (scalability) 가 매우 떨어진다는 단점이 있다.



<그림 1> 두 개의 hyperparameter에 대한 Grid Search의 병렬 처리



〈그림 2〉 두 개의 hyperparameter에 대한 Grid Search와 Random Search의 비교

2. Random Search

Random Search는 Grid Search와 달리 모든 가능한 설정을 확인하지 않고, 임의로 선택된 몇 가지 설정에 대해서만 확인하여 그 중에 가장 좋은 성능을 보인 설정값을 선택하는 기법이다^[3]. Random Search가 임의의 설정값을 선택하기 때문에 최적값을 항상 찾지 못할 것이라고 생각할 수도 있지만, deep learning에 적용했을 경우 일반적으로 Grid Search보다 더 좋은 hyperparameter 설정값을 찾는 경우가 많다.

〈그림 2〉는 두 개의 hyperparameter를 최적화하는 Grid Search와 Random Search를 비교한다. 이 때 가로축의 hyperparameter를 결과에 미치는 영향이 크고, 세로축의 hyperparameter는 결과에 미치는 영향이 작다고 하자. 이 때 각각 9개의 설정값을 확인한다면, Grid Search는 왼쪽 그림과 같이 세 개의 중요한 hyperparameter만 확인하게 된다. 하지만, 오른쪽 그림과 같이 Random Search는 중요한 hyperparameter의 가능한 값 중에서 9개를 확인할 수 있기 때문에 결과적으로 Grid Search보다 더 좋은 설정값을 찾을 수 있게 된다. 그리고, Random Search도 Grid Search와 같이 병렬적으로 수행이 가능하다는 장점이 있다.

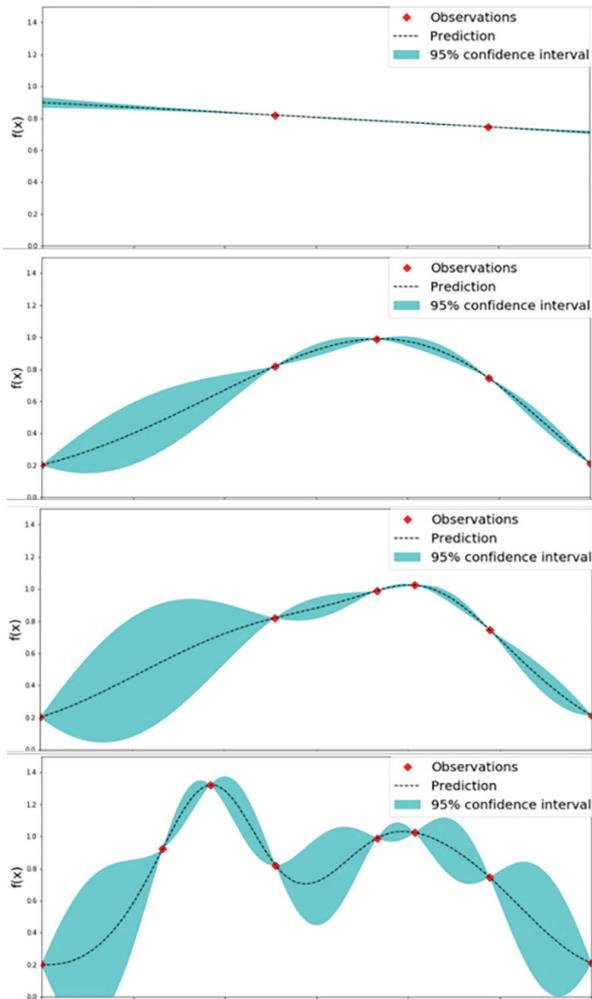
이와 같은 상황은 딥러닝 학습에서 예외적인 경우가 아니라 거의 항상 발생한다. 이런 이유로 일반적으로 딥러닝에서는 Random Search가 Grid Search보다 더 좋은 결과를 낼 수 있게 되며, 대부분의 HPO 알고리즘 연구에서 비교군으로 Random Search가 사용되고 있다.

3. Bayesian Optimization

앞서 소개한 Grid Search와 Random Search는 각각의 설정값을 확인할 때 앞선 결과가 다음 선택에 영향을 미치지 않아서 병렬적으로 수행할 수 있다는 장점이 있지만, 이미 최적의 설정값을 찾았음에도 불필요한 확인작업을 수행하게 되어 시간과 자원의 낭비가 일어날 수도 있다는 단점이 있다. 이와 같은 문제를 해결하기 위하여 개발된 방법이 Bayesian Optimization (이하 BO) 이다. BO는 Sequential Model Based Optimization이라고도 불리는 방법으로서 앞선 테스트 결과를 바탕으로 다음 Sampling Point를 선택하는 것을 특징으로 하고 있다.

BO는 두 가지 핵심 요소를 갖고 있는데 하나는 Surrogate Model이고, 다른 하나는 Acquisition Function 이다. BO는 지금까지 얻은 테스트 결과를 이용하여 대상 모델을 잘 근사하는 Surrogate Model을 만들게 되는데 이 Surrogate Model이 실제 모델과 점점 더 유사해질 수록 최적의 hyperparameter를 더 빨리 찾을 수 있게 된다. Surrogate Model을 만들 때 많이 사용하는 방법중의 하나가 Gaussian Process (이하 GP) 이다.

Acquisition Function 은 다음 Sampling Point를 결정하는 함수인데, 여기서 고려해야할 특성이 바로 Exploration과 Exploitation이다. Exploration은 기존 결과와 상관없는 지점을 찾을 수 있도록 하는 성질이고 Exploitation은 기존 결과를 기반으로 더 좋은 결과가 예측되는 지점으로 수렴하게 하는 성질이다. Acquisition Function은 Exploration 과 Exploitation 을 잘 조정



(그림 3) Gaussian Process를 이용한 Bayesian Optimization

하여 Local Optimum에 빠지지 않고 더 좋은 설정값을 찾을 수 있도록 한다. 일반적으로 가장 많이 사용되는 Acquisition Function은 Expected Improvement Algorithm이다.

〈그림 3〉은 GP를 이용한 Bayesian Optimization의 과정을 보여준다^[4]. 첫 번째 그림에서 두 개의 빨간 점은 임의로 선택된 두 개의 hyperparameter 설정값에 대한 결과이다. 그리고 점선은 이 결과를 기반으로 얻어진 Surrogate Model이고 파란 영역은 95% 신뢰 구간을 나타낸다. 처음에는 두 개의 결과밖에 없으므로 원래 모델이 선형적인 모습을 갖고 있다고 예측하게 된다. 여기서 Acquisition Function이 두 점의 중간정도에 해당하는 위치를 다음 Sampling Point로 선택하여 학습을 진

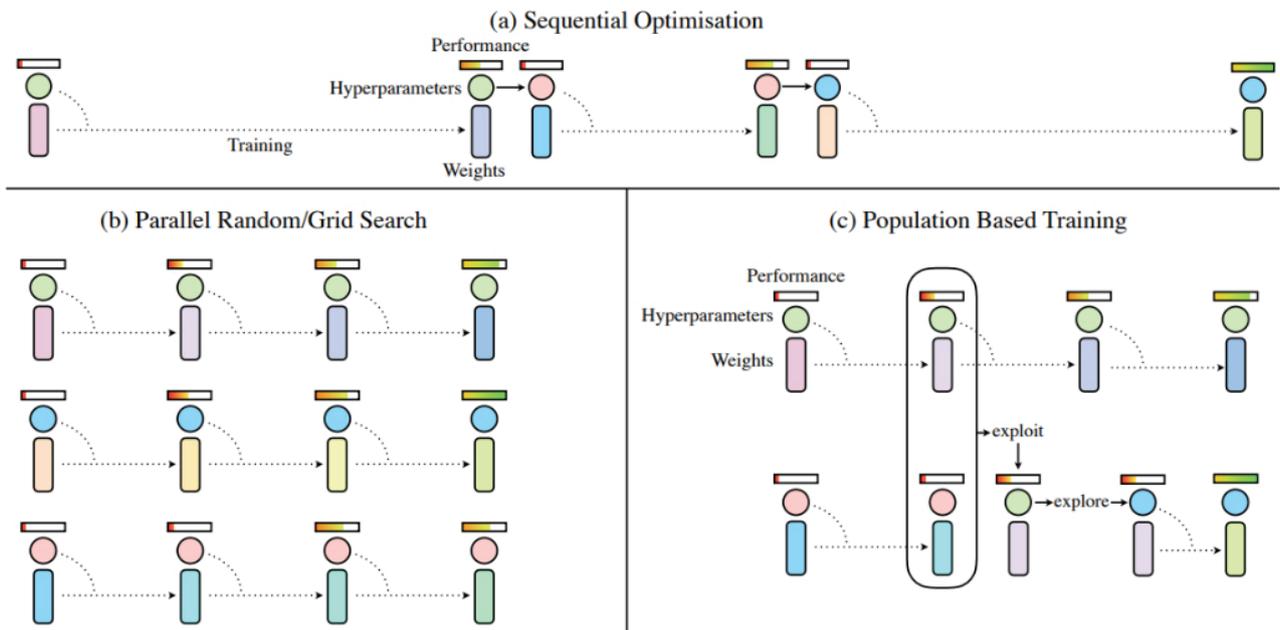
행한다. 이를 통해서 얻어진 값이 표시된 것이 두 번째 그림이다. 이때까지 얻어진 결과를 이용해서 다시 한 번 Surrogate Model을 개선한다. 첫번째 그림보다는 좀 더 복잡한 모습을 갖게 된 것을 확인할 수 있다. 그리고 다시 새로운 Sampling Point를 선택해서 학습을 진행하여 결과를 얻고 다시 이를 활용하여 Surrogate Model을 개선해나가다보면 마지막 그림과 같은 상태가 된다.

BO는 지금까지 얻은 결과를 바탕으로 다음 Sampling Point를 결정하기 때문에 불필요한 학습을 줄일 수 있어서 시간과 계산 자원을 효율적으로 활용할 수 있다는 큰 장점이 있다. 하지만, 이런 특징으로 인하여 Grid Search나 Random Search와는 달리 병렬화가 어렵다는 단점이 있다. 또한 Surrogate Model을 만들고 Acquisition Function을 이용해서 다음 Sampling Point를 선택하는 과정이 좀 더 복잡하여 시간이 좀 걸린다는 단점도 있다. 하지만 딥러닝 모델을 학습시키는 시간에 비하면 이는 큰 문제는 아니다.

앞서 그림을 통해서 설명한 GP는 연속적인 값을 갖는 hyperparameter에 대해서는 잘 동작하지만, 여러 Optimizer중에 하나는 선택하는 것과 같이 비연속적인 값을 갖는 hyperparameter에 대해서는 바로 적용하기 어렵다는 한계를 갖고 있다. 각각의 비연속적인 값을 숫자에 대응하여 GP를 이용할 수도 있지만 정확도가 좀 떨어진다는 문제가 있다. 이를 극복하기 위하여 소개된 방법이 Tree-structured Parzen Estimators (이하 TPE)이다. 저자의 경험으로는 GP가 일반적으로 TPE보다 최적값을 더 빨리 찾는 경향이 있음을 확인하였다. 그러므로, TPE가 GP보다 항상 더 좋다고 할 수는 없으므로, 최적화할 hyperparameter에 비연속적인 값이 있는 경우에만 TPE를 선택하는 것을 추천한다.

4. Population-based Training

Population-based Training (이하 PBT)는 유전알고리즘 (Genetic Algorithm)에 기반을 둔 HPO 기법이다^[5]. 〈그림 4〉는 PBT와 다른 HPO 기법을 비교하여 보여주고 있다. 〈그림 4〉의 (a)는 BO를 도식화하고 있다. 선택된 hyperparameter를 이용하여 학습을 진행한 후 기



〈그림 4〉 Population-based Training과 다른 HPO 기법과의 비교

존의 결과에 기반하여 새로운 hyperparameter를 선택한다. 그림에서 연두색 hyperparameter가 분홍색으로 변하는 것이 이런 과정을 나타내는 것이다. 〈그림 4〉의 (b)는 Grid Search와 Random Search를 도식화하고 있다. (a)와 달리 병렬적으로 실행되고 있음을 도식화하고 있다. (c)는 PBT의 진행과정을 보여주고 있다. 처음에는 Random Search처럼 여러개의 hyperparameter를 병렬적으로 학습을 진행하다가 미리 정해진 진행률에 도달하면 학습을 멈추고, 그 때의 결과를 비교한다. 이 때 BO처럼 Exploitation 과정을 진행하는데, PBT에서는 상대적으로 성능이 낮은 hyperparameter를 성능이 더 좋은 설정값으로 대체한다. 그리고, 학습이 시작되기 전에 유전 알고리즘의 Mutation을 적용하여 hyperparameter 설정값을 수정하는데 이는 Exploration이라고 한다. 이런 과정을 계속 거치면서 최적의 hyper-parameter를 찾게 된다.

PBT가 다른 HPO기법과 다른 점은 hyperparameter의 설정값만 결정하는 것이 아니라, 어느 시점에 변경이 되었는지도 결정한다는 것이다. 이런 특성으로 인하여 다른 기법에 비해서 실제 적용할 때 코드상의 수정이 많이 필요하다는 단점이 있다.

5. Successive Halving Algorithm

Successive Halving Algorithm (이하 SHA)는 Multi-fidelity Optimization에 속하는 HPO 기법이다 [6, 7]. SHA는 임의로 선택된 여러개의 hyperparameter 설정값으로 동시에 학습한다는 점에서 Random Search와 유사한 점이 있다. 하지만 모든 설정값에 대한 학습을 끝까지 진행하지 않고, 미리 정해진 진행률에 도달할 때마다 성능이 좋지 않은 절반의 학습을 중단한다는 점에서 차이가 있다.

〈그림 5〉는 SHA가 어떻게 진행되는지를 보여준다. 가로축 budget은 학습횟수 (iteration) 을 의미하고, 세로축 loss는 성능을 비교하기 위한 수치로서 낮을 수록 높은 성능을 의미한다. 전체 학습횟수 8분의 1 지점까지는 모든 설정값을 다 학습을 진행한다. 그리고 이 시점에서 낮은 성능을 보이는 절반의 설정값은 학습을 중단하고 나머지 절반만 계속 학습을 진행한다. 그리고 전체 학습횟수의 4분의 1 지점에서 다시 한 번 성능을 비교해서 성능이 낮은 절반은 학습을 중단한다. 이런 과정을 계속 반복하면서 마지막 하나의 설정값이 선택될때까지 반복하며 최적의 hyperparameter를 찾게 된다.

이 방식은 안좋은 결과를 낼 것으로 예상되는 설정값

에 대한 학습을 이른 시점에 중단시키기 때문에 Random Search나 BO보다 자원의 활용 측면에서 보다 효율적인 알고리즘이다. 하지만, 설정값의 갯수와 자원 할당을 어떻게 조절하느냐가 결정하기 어려운 문제이다. 그 이유는 설정값의 갯수를 늘리면 각 설정값당 자원이 적게 할당이 되고 이로 인해 너무 빨리 종료가 되어서 제대로 된 비교를 할 수 없게 되기 때문이다. 이를 극복하기 위해 자원을 많이 할당하면 자원의 낭비가 발생하게 되고 자원의 할당량을 줄이면 최적의 hyperparameter를 찾기 어렵게 된다.

6. 이외의 Multi-fidelity Optimization 기법

SHA의 단점을 극복한 것이 HyperBand (이하 HB) 알고리즘이다^[8]. SHA가 고정된 설정값의 갯수와 고정된 자원할당량을 갖고 있는 것에 비해 HB는 다양한 설정값의 갯수와 자원할당량의 조합에 대해서 학습을 한다. 특히, SHA는 무조건 절반씩 중단시키지만 HB는 이 비율도 조절이 가능하다. 현재까지 알려진 바에 따르면 HB가 앞서 설명한 Random Search나 BO보다 더 적은 자원을 사용해도 딥러닝 모델에서 더 좋은 결과를 낸다고 알려져 있다.

SHA와 HB는 한꺼번에 여러 hyperparameter 설정값을 학습시킬 수 있지만 성능이 안좋은 것으로 예상되는 절반을 결정하기 위해서 일정 시간마다 모든 학습이 종료되기를 기다린다는 문제가 있다. 이를 극복한 것이 Asynchronous Successive Halving Algorithm (이하 ASHA)이다^[9]. ASHA는 일정 시간마다 모든 학습이 종료되기를 기다리지 않기 때문에 자원의 활용률 측면에서 SHA나 HB보다 효율적이다.

마지막으로, Bayesian Optimization HyperBand (이하 BOHB)라는 알고리즘도 있다. SHA, HB, ASHA는 모두 Random Search를 사용한다는 공통점이 있다. Random Search는 알고리즘 초반에는 효과적이지만, 최종 성능이 더 좋을 것으로 예상되는 hyperparameter 설정값을 선택하는 후반부에 갈수록 효과적이지 못하다는 단점이 있다. 이것을 Bayesian Optimization을 활용하여 극복한 것이 BOHB이다.

아래의 <표 1>에 지금까지 소개한 Multi-fidelity

Optimization Algorithm을 비교하여 정리하였다. SHA는 항상 절반의 hyperparameter 설정값을 중단하도록 고정되어 있지만, 나머지는 조절이 가능하다. 그리고 iteration의 횟수와 자원할당량의 관계 역시 SHA만 고정되어 있고, 나머지는 여러가지 조합을 활용한다. Hyperparameter의 설정값은 BOHB를 제외한 나머지는 모두 Random Search를 사용한다. 여기까지만 고려한다면 BOHB가 가장 좋아보이지만, 확장성까지 고려한다면 ASHA가 BOHB보다 나은 점도 있다.

III. HPO 응용

앞서 소개한 다양한 HPO 알고리즘은 이미 다양한 형태로 구현하여 공개된 것이 많이 있으며, 용도에 따라 선택하여 사용하는 것이 가능하다. UC-Berkeley에서 공개한 Ray-Tune^[10], Microsoft의 Neural Network Intelligence^[11], Google의 Vizier^[12], Amazon의 SageMaker^[13], Intel의 Sigopt^[14] 등이 대표적인 HPO를 구현해놓은 응용이다. 이 중에서 본고에서는 Ray-Tune과 Sigopt를 소개한다.

1. Ray-Tune

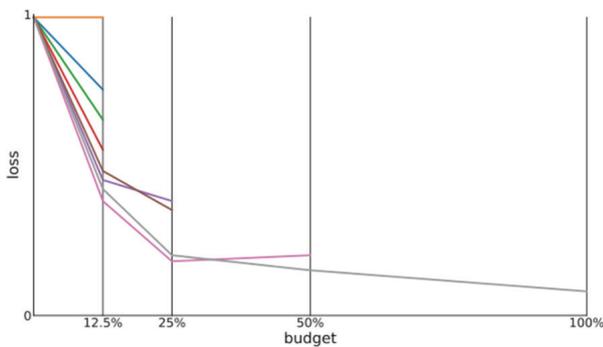
Ray-Tune은 UC-Berkeley의 RISE lab에서 개발한 HPO 프레임워크이다. 앞서 소개한 모든 HPO 알고리즘을 모두 구현하여 제공하고 있기 때문에 다양한 HPO 알고리즘의 성능 및 특성을 비교해보기에 매우 편리하다. 특히 오픈소스 소프트웨어이기 때문에 내부 구현도 확인해볼 수 있다는 점에서 큰 장점을 갖고 있다. 상업용 소프트웨어가 아니기 때문에 클라우드를 통해서 서비스되는 형태는 아니지만, 숙련된 사용자라면 클라우드 환경에서 설치하여 사용하는 것도 가능하다.

2. Sigopt

Sigopt는 Ray-Tune과 달리 오픈소스 소프트웨어가 아니라 클라우드 기반 상업용 소프트웨어이다. Ray-Tune처럼 다양한 HPO 알고리즘을 제공하지는 않지만, 훌륭한 시각화 도구와 과거 이력을 검색해볼 수 있는 기

〈표 1〉 Multi-fidelity Optimization Algorithm의 비교

	SHA	HB	BOHB	ASHA
Discard ratio	half	adjusted by user (1/3)	adjusted by user	adjusted by user
B vs n	one choice	several combinations	several combinations	several combinations
Sampling	random	random	TPE	random
Parallelization	synchronous	synchronous	synchronous	asynchronous
Scalability	small scale	small scale	small scale	large scale
Time	2015	2018	2018	2018



〈그림 5〉 Successive Halving Algorithm의 진행

능을 제공하기 때문에 편리한 사용법이 Ray-Tune 대비 장점이다. 다만, 클라우드 기반 서비스이기 때문에 무조건 인터넷에 연결되어야 한다는 제약이 있지만, 인터넷에 연결이 가능하다면 훌륭한 시각화 도구를 제공하기 때문에 각각의 hyperparameter가 최종 성능에 어떤 영향을 미치는지 분석하는데 적합한 HPO 응용이다.

IV. HPO의 한계

HPO를 사용하면 최적의 hyperparameter를 찾아서 최고의 성능을 얻을 수 있는 모델을 학습시킬 수 있는 것은 사실이지만, 이를 위해서는 많은 시간과 자원이 필요하기 때문에 쉽게 사용하기에는 현실적인 어려움이 있다. 딥러닝 모델이 점점 더 거대해지면서 학습을 여러번 수행해본다는 것은 점점 더 어려워지고 있기 때문이다. 이런 문제를 해결하기 위하여 HPO를 좀 더 빠르게 수행하기 위한 노력들이 끊임없이 이어지고 있다.

Median Stopping은 BO에 적용할 수 있는 기법으로서, BO의 테스트 설정값을 끝까지 학습시키지 않고 특

정 조건에 부합하지 않으면 학습을 중단시키는 방법이다^[15]. Iteration S에서 현재의 테스트 설정값 X로부터 얻어진 중간결과가 완료된 hyperparameter의 설정값들의 Iteration S에서의 모든 중간결과들의 median 값보다 작으면 중단시키는 기법이다. 이 기법은 Google의 Vizier, Ray-Tune, Microsoft의 NNI에 구현되어 적용되었다.

HPO의 또다른 문제점은 HPO를 사용하기 위해서 설정해야 하는 매개변수가 최적화하려는 hyperparameter의 갯수보다 많은 경우가 발생할 수 있다는 것이다. 예를 들어 어떤 사용자가 Learning Rate와 Momentum만을 최적화하고 싶다고 했을 때 이를 위한 HPO를 설정하려면 이보다 훨씬 많은 HPO 알고리즘의 매개변수를 설정해야 할 수도 있다. 〈그림 6〉은 Ray-Tune으로 두 가지 hyperparameter를 최적화하기 위한 예제 코드를 보여주고 있다. 예제 코드에서 볼 수 있듯이, 각각의 hyperparameter에 대해서 최대값과 최소값을 정해주어야 하고, 각 hyperparameter가 어떤 분포를 갖고 있는지도 정해주어야 한다. 이 예제코드에서는 두 개의 hyperparameter를 최적화하기 위해서 네 개의 HPO 매개변수를 설정한 것이다. 이런 설정들은 자동으로 결정하기 어렵기 때문에 어쩔 수 없이 사용자의 노력이 필요한 부분이고, 이런 것이 HPO를 쉽게 사용하기 어렵게 만드는 요인 중의 하나이다.

V. 전망과 결론

본고에서는 지금까지 소개된 다양한 HPO 기법과 이를 구현한 응용들을 살펴보고 현재의 HPO가 가진 한계에 대해서 논의해보았다. HPO를 활용하면 최고의 성능을



```

search_space = {
    "lr": tune.sample_from(lambda spec: 10 ** (-10 * np.random.rand())),
    "momentum": tune.uniform(0.1, 0.9),
}

# Uncomment this to enable distributed execution
# `ray.init(address="auto")`

# Download the dataset first
datasets.MNIST("~/data", train=True, download=True)

tuner = tune.Tuner(
    train_mnist,
    param_space=search_space,
)
results = tuner.fit()

```

〈그림 6〉 Learning Rate와 Momentum을 최적화하기 위한 Ray-Tune의 예제 코드

달성할 수 있도록 딥러닝 모델을 학습시킬 수 있음은 자명하다. 하지만, 많은 시간과 자원이 필요하고, 이를 적용시키는데에도 많은 사전지식이 필요하기 때문에 HPO를 쉽게 사용하기 어려운 것이 사실이다. 앞으로 HPO 기술이 현재보다 훨씬 더 적은 시간과 자원을 사용하여 실행이 가능해지고, 사용법이 단순해진다면, HPO의 사용처가 그에 비례하여 늘어나게 될 것이고, 이를 통해 더 좋은 모델을 만들어서 인류의 삶에 도움이 되기를 바란다.

참고 문헌

- [1] Radwa Elshawi, et al. "Automated Machine Learning: State-of-The-Art and Open Challenges", <https://arxiv.org/abs/1906.02287>, 2019
- [2] Alessio Gozzoli, "Practical Guide to Hyperparameters Optimization for Deep Learning Models", <https://blog.floydhub.com/guide-to-hyperparameters-search-for-deep-learning-models/>, 2018
- [3] James Bergstra et al. "Random Search for Hyper-Parameter Optimization", Journal of Machine Learning Research, 2012
- [4] Tong Yu, et al. "Hyper-Parameter Optimization: A Review of Algorithms and Applications", <https://arxiv.org/abs/2003.05689>, 2020
- [5] Max Jaderberg, et al. "Population Based Training of Neural Networks", <https://arxiv.org/abs/1711.09846>, 2017
- [6] Kevin Jamieson, et al. "Non-stochastic Best Arm Identification and Hyperparameter Optimization", AISTATS, 2016
- [7] André Biedenkapp, et al. "BOHB: Robust and Efficient Hyperparameter Optimization at Scale" https://www.automl.org/blog_bohb/, 2018
- [8] Lisha Li, et al. "Hyperband: A novel bandit-based approach to hyperparameter optimization", The Journal of Machine Learning Research, 2017
- [9] Liam Li, et al. "Massively parallel hyperparameter tuning", <https://arxiv.org/abs/1810.05934>, 2018.
- [10] Ray-Tune, <https://docs.ray.io/en/latest/tune/>, UC-Berkeley
- [11] Neural Network Intelligence, <https://nni.readthedocs.io/>, Microsoft
- [12] Vizier, <https://github.com/google/vizier>, Google
- [13] SageMaker, <https://aws.amazon.com/sagemaker/>, Amazon
- [14] Sigopt, <https://sigopt.com/>, Intel
- [15] Daniel Golovin, et al. "Google Vizier: A Service for Black-Box Optimization", ACM SIGKDD, 2017



김신규

- 2006년 2월 서울대학교 컴퓨터공학부 학사
- 2008년 2월 서울대학교 컴퓨터공학부 석사
- 2013년 8월 서울대학교 컴퓨터공학부 박사
- 2013년 9월 ~ 2020년 10월 삼성전자
- 2020년 11월 ~ 현재 Intel

〈관심 분야〉

Distributed Computing, Parallel Processing, Compiler,
Machine Learning

트랜스포머 모델의 경량화

I. 서론

최근 자연어 처리 (NLP) 분야에서 GPT 모델처럼 트랜스포머 기반 대규모 언어 모델의 우월한 성능이 지속적으로 입증되고 있다. 그 모델 사이즈는 상상을 초월할 정도로 해를 거듭할 수록 커지고 있다. 파라미터와 데이터, 학습 시간을 폭발적으로 늘리면서 언어 모델은 언어의 미묘한 뉘앙스까지 이해할 수 있게 되었다. NVIDIA Megatron과 DeepSpeed 기반 Megatron-Turing Natural Language Generation (MT-NLG)은 무려 파라미터의 수가 5300억 개에 달한다. 2022년 12월 기준으로 현존하는 가장 대규모 자연어 모델이다.

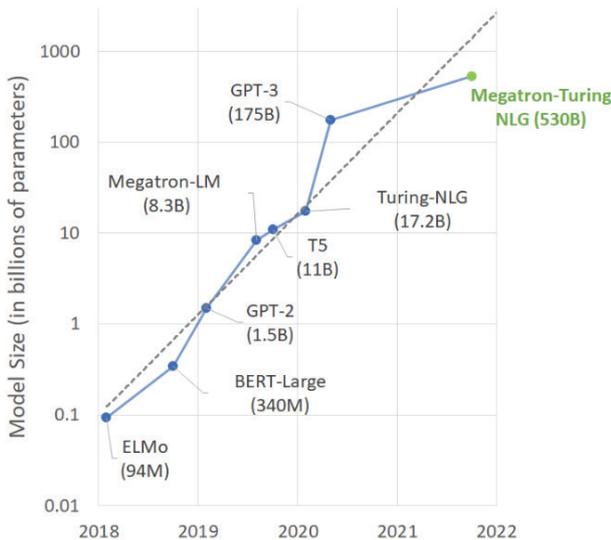
그러나 이러한 대규모 언어 모델에는 여러 가지 어려움이 따른다. 현존하는 GPU 메모리로는 이 모델들의 파라미터를 로드하기가 어렵고 대규모 컴퓨팅 인프라가 부족한 열악한 상황에서는 그 모델을 활용하기가 어렵다. 그리고 모델을 학습하는 동안 GPT-3 모델의 경우에 수백 페타플롭(PF)의 과도한 컴퓨팅을 몇 달에 걸쳐 요구하기에 심각한 탄소배출을 초래하는 환경피해를 야기하고 있다. 이 문제를 해결하기 위해서 트랜스포머 기반의 대규모 자연어 모델을 메모리와 연산량 측면에서 경량화하는 노력은 매우 필연적이다.

II. 트랜스포머 모델의 응용

대표적인 트랜스포머의 응용 분야로 자연어 처리 (NLP)를 들 수 있다. BERT (Pre-training of Deep Bidirectional Transformers for Language Understanding) 모델은 bi-directional 트랜스포머 모델을 사용하는 모델이다. BERT는 대형 코퍼스에서 Unsupervised Learning으로 광범위한 목적으로 언어를 이해하도록 사전 모델을



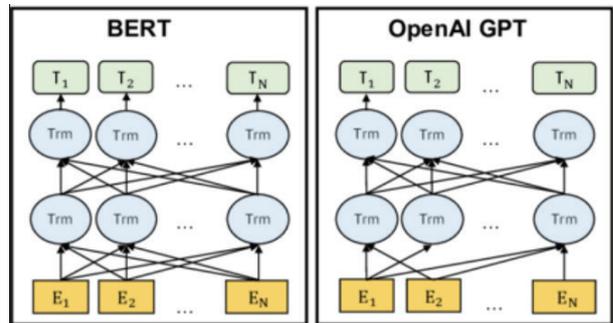
오지훈
Neubla



〈그림 1〉 자연어 모델 사이즈의 변화^[1].

구축하는 것이다. 많은 Downstream MLP Task에서 Language Model Pre-training을 활용하는 것은 효과적이라고 연구되었기 때문이다. 사전 모델 학습시에는 Masked Language Model (MLM)과 Next Sentence Prediction (NSP) Task를 가지고 Unsupervised Learning을 한다. 이렇게 pretrained 모델을 이용하여 문장 레벨로 추론하는 Natural Language Inference 같은 Sentence-level Task와 Question-Answering, Name Entity Recognition와 같이 토큰 레벨로 추론하는 Token-level Task를 Supervised Learning으로 Fine-tuning하여 추론에 이용한다.

트랜스포머의 인코더를 이용하는 BERT와 달리, GPT (Generative Pre-trained Transformer)는 left-to-right 트랜스포머를 사용하는데, 한 단어가 입력되면, 다음에 올 적절한 토큰을 생성하는 Autoregressive 방식으로 학습된 언어 모델이다. 2022년 OpenAI에서 GPT-3 버전까지 공개를 하였는데, 1,750억 개의 학습 파라미터를 가지고 있어 그 이전 버전인 GPT-2보다 2배 이상 큰 모델이다, 그 이전에도 마이크로소프트에서 공개한 튜링 NLG가 있는데, GPT-3 보다 파라미터가 10배 적다. 활용 분야는 시나 소설을 쓰는 등의 텍스트를 완성해준다거나, 상식적인 질문에 대한 응답, 자유 대화, 기본적인 수준의 컴퓨터 프로그래밍까지, 그리고 GPT-3를 발전시킨



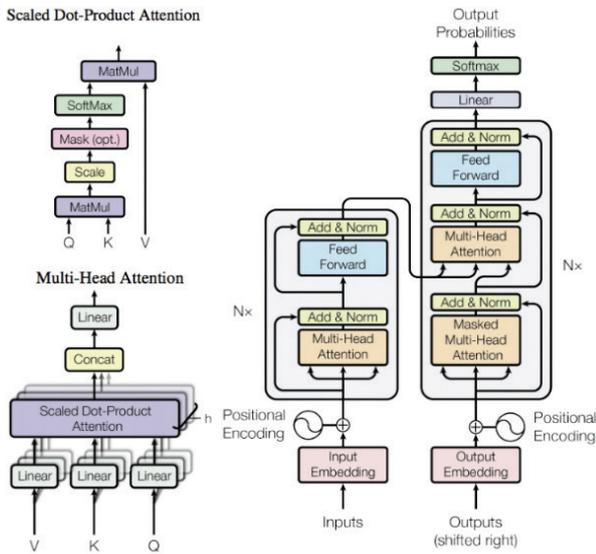
〈그림 2〉 BERT와 GPT 모델의 차이점. BERT는 bi-directional Transformer를 사용하고 GPT는 left-to-right Transformer를 사용한다^[2].

GPT-3.5를 기반으로 학습을 거친 ChatGPT처럼 인간과 구분하기 어려울 정도로 자연스런 대화 전문의 인공지능 챗봇 서비스를 구현 할 수 있다.

또 다른 예로, 대용량 학습 능력이 가능하고 여러 형태의 데이터를 의미 있는 형태로 변환하는 representation learning 목적을 위해서도 사용할 수 있다. 2021년에 발표된 OpenAI의 CLIP^[3]은 텍스트와 이미지 간의 representational learning의 성능을 향상시키기 위해서 기존 ImageNet보다 방대한 4억 개의 이미지 데이터를 사용하였는데, 수작업 레이블링 대신에 웹 크롤링을 통해서 자동으로 이미지와 캡션을 추출하여 거대한 데이터셋을 구축하였다. 관련된 이미지들과 텍스트들 사이에 유사성을 찾는 문제로 신경망을 학습하는데, 여기서 이미지와 텍스트를 각각 인코딩할 때 트랜스포머 인코더의 출력값을 사용한다. 이 인코더의 출력값 (임베딩 벡터)을 다양한 downstream task에 사용할 수 있는데, 그 범위가 NLP, computer vision, generative model까지 매우 무궁무진하다고 할 수 있다.

III. 트랜스포머 모델의 구조

다양한 NLP task에 대해서 단일 architecture로 수행하는 것이 트랜스포머 모델의 핵심이다. 구체적인 예로 BERT 모델 구조는 다중 레이어 양방향 트랜스포머 인코더 구조다. Pre-training부터 Fine-tuning까지 Downstream Task를 수행하기 위해서 공통된 인코더 레이어를 N개 쌓아 올려 사용하는 모델 구조의 단순함이 큰

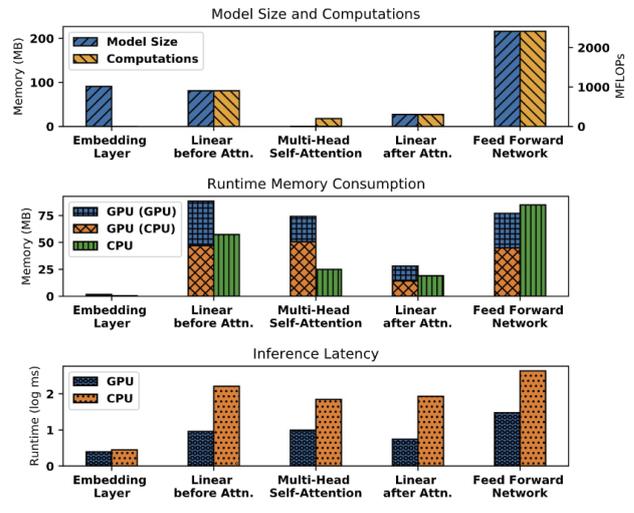


〈그림 3〉 트랜스포머 모델 - 인코더-디코더 모델 구조와 주요 구성 요소^[4].

장점이다.

입력과 출력은 다양한 Downstream Task를 위해서 임베딩 벡터 형태로 변형된다. BERT에서 문장이란 언어적 문장이 아닌, 단순히 토큰의 연속적인 시퀀스라고 간주한다. 복수의 문장을 분리하기 위해서 [SEP] 토큰으로 구별한다. BERT의 입력 Representation은 세 가지 Embedding을 합성하게 되는데, 자연어를 토큰화하여 Vocab ID로 맵핑하는 Token Embedding, 다음 문장을 예측하는 태스크를 위해 문장간 구분을 위해서 도입한 Segment Embedding, 그리고 토큰의 순서 정보를 주는 Positional Embedding이다.

BERT-base 모델에서 학습된 파라미터 크기와 Floating point Operations (FLOPs) 측면에서 가장 많은 메모리와 연산량을 소비하는 레이어는 피드포워드 (Feed Forward Network: FFN)이다. 구체적으로 BERT-base 모델에서 하나의 피드포워드는 $(768 \times 3072) + (3072 \times 768) = 4829K$ 만큼의 학습파라미터를 요구한다. 그 다음으로 멀티헤드 어텐션 블록에서 어텐션 (Scaled dot-product attention) 이전의 선형 레이어는 각 어텐션이 3개의 입력(key, value, query)으로 단 하나의 출력을 만들기 때문에 어텐션 이후의 선형 레이어 크기의 약 3배다. 실제로 BERT-base 모델



〈그림 4〉 BERT base 모델의 메모리/연산 프로파일^[5].

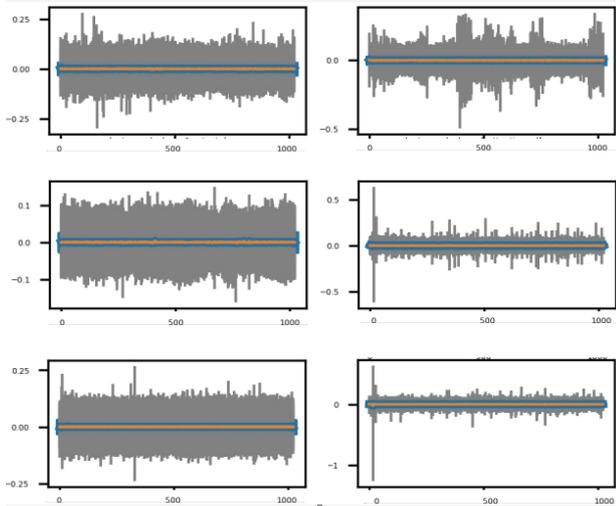
의 멀티헤드 어텐션 메커니즘에서는 768차원 = 64차원 x 12헤드로 구성되어 있고 어텐션 이전의 선형 레이어에서는 $(768 \times 64) \times 12$ 헤드 x 3 (key, value, query) = 1769K 그리고 어텐션 이후의 선형 레이어에서는 $(768 \times 768) = 589K$ 파라미터를 필요로 한다. Multi-head Self-Attention 모듈은 파라미터가 존재하지 않고 key, query embedding vector 간에 MATMUL, Scale 그리고 Softmax 연산을 수행한다. 이 때, 시퀀스 길이의 제곱에 비례하여 이 모듈의 연산량이 결정된다.

그래픽 처리 장치 (GPU)에서 모델을 실행할 때 총 런타임 메모리에는 GPU 측과 중앙 처리 장치 (CPU) 측 모두의 메모리가 포함되며, 이는 두 장치에 중복 텐서가 존재하기 때문에 CPU에서만 실행되는 모델보다 크다. 그래서 멀티헤드 어텐션 블록은 이론보다 실제로 훨씬 더 많은 비용이 든다. GPU는 더 복잡한 어텐션 레이어보다 선형 레이어를 더 빠르고 효율적으로 가속화하도록 설계되었다. 그러므로 FFN 레이어는 전체 모델에서 가장 큰 병목현상을 만든다고 할 수 있다.

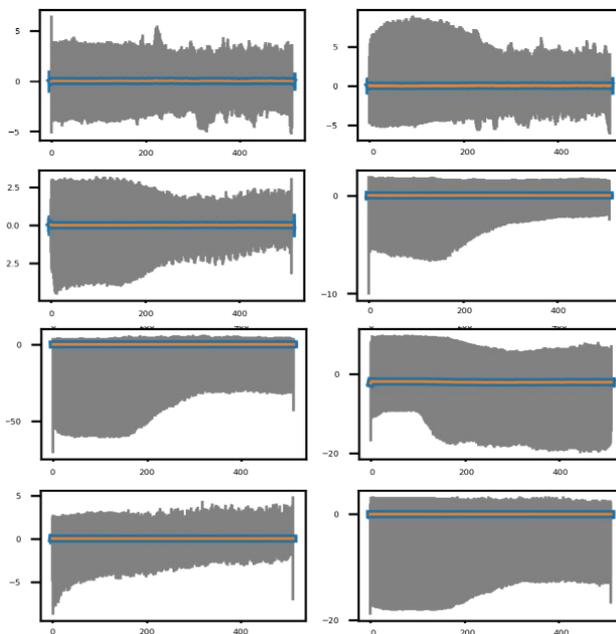
IV. 트랜스포머 모델의 압축기법들

1. 양자화 (Quantization)

양자화는 신경망 모델의 학습된 실수형 (FP32) 파라미터를 정수형 파라미터 (INT, Fixed point)로 변환하



〈그림 5〉 BERT large 모델의 첫번째 인코더의 여섯 레이어별 각 채널 (가로축) 파라미터의 min-max 범위 (세로축). 좌측에서 우측으로, 위에서 아래 순으로 (1) self_attention query layer, (2) self_attention key layer, (3) self_attention value layer, (4) self_attention output dense layer, (5) intermediate dense layer, (6) output dense layer.



〈그림 6〉 BERT large 모델의 첫 번째 인코더의 여섯 레이어별 각 토큰 (가로축)의 moving average min-max 범위 (세로축). 좌측에서 우측으로, 위에서 아래 순으로 (1) self_attention query layer, (2) self_attention key layer, (3) self_attention value layer, (4) self_attention output dense layer, (5) self_attention output layernorm, (6) intermediate dense layer (7) output dense layer (8) output layernorm.

는 과정을 의미한다. 보통 학습은 Weight, Activation의 Gradient의 표현 스케일이 다양해야 하기 때문에, 부동 소수점 방식인 FP32 또는 FP16 둘 중에 선택하거나 레이어들 간에 혼용해서 사용하고 있다. 그러나 학습을 마친 모델을 추론 (Inference)에 이용할 때는 INT8과 같이 Low integer bit으로 변환함으로써 모델의 사이즈를 축소할 수 있어서 저장 용량을 효율화할 수 있다. 또한, 계산과 메모리 대역폭을 절약할 수 있어서 추론 속도를 높이고 전력 소비를 효율화할 수 있다.

양자화 시점에 따라 런타임 사전에 하는 정적 양자화 (Static quantization)와 런타임 과정에 하는 동적 양자화 (Dynamic quantization)로 크게 나눌 수 있다. 정적 양자화를 하기 위해서 타겟 테스트 도메인을 충분히 대표할만한 적당한 수의 샘플 데이터들로 추론하였을 때, 레이어별로 weight와 activation의 표현이 일정 분포를 가져야 한다. 그 범위가 매우 크거나 샘플 데이터마다 매우 다른 분포도를 가진다고 한다면, 정적 양자화 이후의 정확도는 원래 모델이 가지고 있던 정확도를 유지하기 어렵다. 위 그림에서 BERT의 학습된 모델을 이용하여 레이어별로 파라미터의 min-max 범위를 도식화하였을 때, intermediate dense layer와 output dense layer에서 특정 채널에서 다른 채널보다 범위가 매우 큰 것을 관찰할 수 있다. 이 문제를 해결하기 위해서 채널 단위로 양자화하는 channel-wise quantization으로 해결할 수 있다. 하지만, 레이어의 출력값인 activation에 대해서 moving average min-max 범위를 구했을 때, self_attention output layernorm, intermediate dense layer, 그리고 output layernorm의 범위가 다른 레이어보다 상대적으로 크다는 것을 확인할 수 있다. 이는 INT8 데이터 타입에서 양자화 에러가 커질 수 있다. 그리고 중간 레이어들의 표현 분포가 토큰들 간에 매우 상이한 특성을 가진다. 이를 극복하기 위해서 양자화 연산 커널을 지원하는 연산자에서 정적 양자화된 파라미터를 효율적으로 로드하고 activation에 대해서 동적 양자화를 적용한다. 즉, 연산 커널에서 프로세싱하기 전에 float으로 저장되어 있는 activation을 동적으로 integer로 양자화하고 프로세싱 후에는 다시 역양자화를 하여 float으로 저

장한다.

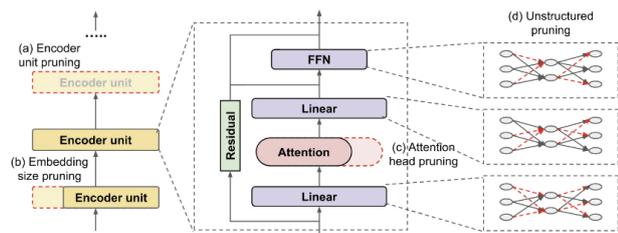
동적 양자화를 지원하는 하드웨어의 선택지는 CPU 또는 제한된 GPU 환경에서만 가능하다. 딥러닝 모델의 효율적인 추론을 위해서 만들어진 신경망 가속기(NPU)의 경우에는 동적 양자화를 지원하지 않는 경우가 많기 때문에, 양자화 학습(Quantization-Aware Training)을 활용해야 한다. 이는 학습 진행 시점에 양자화 적용에 의한 영향을 Fake Quantization 방식으로 시뮬레이션함으로써 양자화를 고려한 학습을 가능하게 하여, 정적 양자화 모델로 최종 변환 시 정확도 감소 폭을 최소화할 수 있다. 다만, 양자화 학습이 고려되지 않은 보통의 모델이라면, 양자화 학습을 추가적으로 진행해야 한다는 단점이 존재한다.

추가적으로 트랜스포머 모델에는 선형 연산자 MATMUL 외에 Softmax, Layer Normalization, GELU같은 비선형 연산자들이 있다. 이 연산자들의 출력값을 작은 오차로 근사화하면서 Integer기반으로 연산 속도를 빠르게 하기 위한 연구들(Look-up table, polynomial 기반^{[6])이 활발히 진행되고 있다.}

2. 전정화 (Pruning)

전정화는 신경망에서 밀도 높은 Synapse (Edge) 또는 Neuron (Node)의 연결 중에서 덜 중요한 부분을 없애는 것이다. 파라미터의 절대값 크기가 작거나 학습 과정 중에 업데이트가 상대적으로 덜 된다는 지표를 가지고 중요도가 낮다고 판단할 수 있다.

전정화할 때 구조를 고려하느냐에 따라서 크게 Unstructured Pruning과 Structured Pruning으로 나눌 수 있다. Unstructured Pruning은 파라미터의 구조



〈그림 7〉 트랜스포머 모델의 Structure Pruning 방법들^[6].

에 상관 없이 0 근처의 파라미터를 가지치기 하듯이 0으로 만드는 것이다. 95%의 높은 비율로 전정화를 하더라도 좋은 정확도 성능을 낸다고 하는 논문 발표들이 있다. 하지만, 높은 비율로 전정화를 하였다고 해도 0을 스킵하면서 연산할 수 있는 Sparse Matrix Kernel (e.g., cuSPARSE)이 없다면 실질적으로 추론 속도를 개선하지 못한다.

대신, Structured Pruning은, 트랜스포머 모델에서 설명하자면, Attention Head, Encoder Unit, Embedding Size 등 Unit 단위로 전정화하는 것을 의미한다. Attention Head의 중요성에 대한 의문은 계속 제기되어 왔는데, BERT-base 인코더당 전체 12개의 Attention Head 중에 1-2개 만으로도 높은 정확도가 가능하다고 주장하는 논문들이 있다^[7,8]. 훈련 단계 동안 Attention Head를 임의로 가지 치기 하는데, 이것은 다양한 수의 Attention Head에 모델을 견고하게 만들고 이후 추론 시에는 더 작은 모델을 추출한다. 다음으로 Encoder Unit을 임의로 가지치기하거나 이미 정의된 전략으로 인코더를 dropout하면서 학습하는 것이다. 마지막으로, 모델의 embedding size를 줄이는 것인데, 이것도 반복적으로 학습하는 중에 제일 중요하지 않은 feature dimension을 제거하는 방식으로 동작한다.

3. 행렬 분해 (Matrix Decomposition)

BERT에서의 컴퓨팅 오버헤드는 주로 Linear Layer와 Self-Attention Layer에 있는 매우 큰 MATMUL 연산에서 발생한다. 행렬 분해는 큰 매트릭스 연산을 작은 복수의 매트릭스 연산으로 분해하는 것을 의미한다. Linear Layer와 Embedding Layer에서는 Weight Matrix와 Input Matrix의 곱 ($A \times B$)을 복수의 작은 Weight Matrix와의 연속적인 곱 ($A \times C$ and $C \times B$)으로 분해한다.

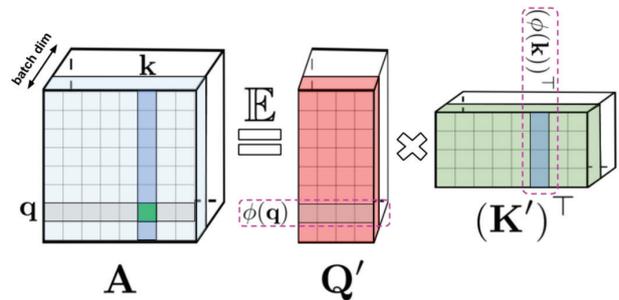
Attention 분해를 살펴보면, 다수의 토큰으로 이루어진 입력 문장에 대해서 모든 토큰 조합 간에 Attention 연산을 하는데, 토큰의 개수 따라 이차적인 (quadratically) 메모리와 연산량 증가를 야기한다. 이는 불필요한 연산의 증가를 발생시킬 수 있다. 왜냐하면, 문장이 길어질수록 (토큰

크기가 증가할수록 모든 토큰이 글로벌 컨텍스트 연관성을 가지고 있지 않을 수 있기 때문이다.

이 문제를 해소하기 위해서, 토큰들을 공간 지역성, 크기에 기반한 지역성, 또는 Adaptive Attention Span을 기준으로 그룹핑할 수 있고 동일 그룹 내부에 있는 토큰끼리의 Attention 조합으로 축소할 수 있다. 게다가, Attention 출력값들은 독립적으로 연산이 가능하기 때문에, 병렬 프로세싱을 활용하여 가속이 가능하다. 추가로, key-query matrix를 낮은 차원으로 Projection하여 연산량을 감소시키거나 Top-k key-query product 값에 대해서만 Softmax 연산을 하는 방법도 제안되었다^[5]. 또, Softmax-attention Matrix를 분해하는 방법을 제안하는 연구가 소개되었다^[9]. Query와 Key를 Normalize dot product 한 후 Softmax를 취해서 만들어진 Attention Matrix를 다시 원래의 Query와 Key의 임의의 비선형 함수의 곱으로 분해할 수 있다. 이를 통해 유사성 정보를 보다 효율적인 방식으로 인코딩할 수 있다.

4. 커널 합성 (Kernel Fusion)

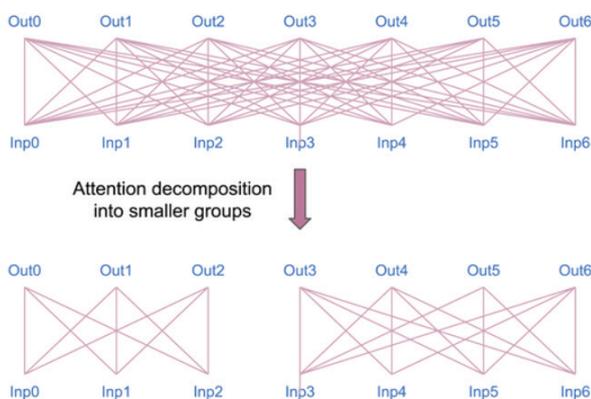
Kernel Fusion은 신경망 모델의 가속을 위해서 인접한 복수의 Layer 또는 Operator들을 하나의 Operator로 병합하여 처리하는 방법이다. 다시 말해, Fusion은 계산 단계에서의 공통화 작업을 의미하는데, 기본적으로 단일 하드웨어 (GPU, CPU, NPU) 작업에서 연속된 작업을 결합하여 코드를 보다 효율적으로 실행하는 구현 트릭이다. 여러 레이어를 처리하면서 데이터 이동이 발생하



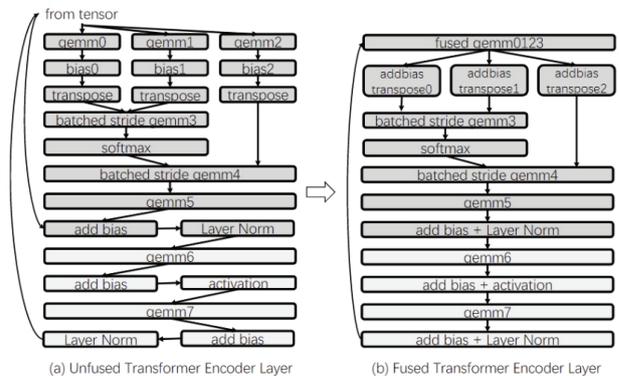
〈그림 9〉 좌변은 모든 토큰 페어간의 query, key의 scaled dot product 연산과 softmax를 통해 나온 유사성을 포함하는 표준적인 attention matrix를 보여주고 있고 우변은 lower-rank randomized matrix 통해서 근사화된 attention matrix를 보여줌^[9].

는 Context Switching에 대한 오버헤드, 메모리 액세스 감소와 캐시 지역성 증가 등의 긍정적인 효과를 가져올 수 있다. CNN 모델에서는 흔히 등장하는 Convolution과 BatchNorm 그리고 Linear Activation 함수들의 패턴을 하나의 커널로 동작하도록 합성한다. 트랜스포머 모델에서는 Batch Size와 Sequence 길이가 상대적으로 큰 BERT 케이스에서 non-GEMM 커널들 (LayerNorm, Softmax, Add Bias, Transpose)의 비효율적인 구현 때문에 전체적으로 낮은 효율의 성능을 발생시킨다. Latency 측면에서 GEMM 커널에 61.8%, 그리고 Non-GEMM 커널에 38.2%의 Latency를 초래하고 있다.

단일 구조가 반복되는 트랜스포머 모델에서도 Kernel Fusion을 십분 활용할 수 있다. 예를 들면, GEMM 커널의 경우 Scaled dot product attention의 입력 단에서 query, key, value에 대해서 독립적으로 Linear Layer



〈그림 8〉 Attention 그룹핑^[5].



〈그림 10〉 오리지널 트랜스포머 인코더 구조와 커널 합성 이후의 인코더 구조^[5].

를 처리하고 있는데, 이를 하나의 query, key, value 매트릭스로 병합하여 단일 Linear Layer로 구현할 수 있다. 그리고 연속된 non-GEMM 커널들 (Add bias와 Transpose, Add bias와 Layer Norm, Add bias와 Activation)들을 합성된 커널로 커스터마이징할 수 있다. 실제로 PyTorch용 오픈 소스 딥러닝 최적화 라이브러리인 DeepSpeed는 고효율의 커스터마이징된 트랜스포머 커널을 개발하여 싱글 혹은 멀티 GPU에서의 학습 Throughput을 획기적으로 개선하였다.

5. 숫자 정밀도 감소 (Numeric Precision Reduction)

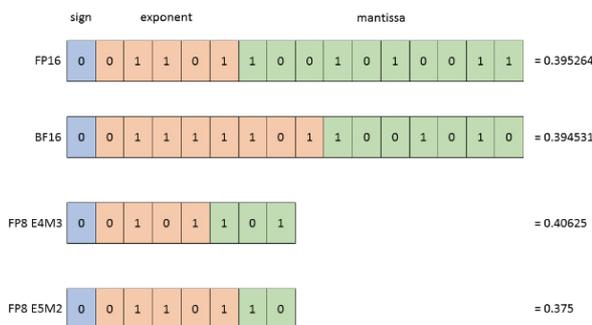
IEEE754 표준을 따르는 Floating Point 표기법에 따르면, 주로 32-bit Floating Point을 Single Precision Floating Point (Sign=1, Exponent=8, Mantissa=23)라고 부른다. 반대로 더 적은 bit로 수를 표현할 때 16 bit를 사용하며 이를 Half Precision (Sign=1, Exponent=5, Mantissa=10)이라고 부른다. 최근 신경망 모델의 사이즈가 점차 커지면서 계산량과 메모리 크기 등이 기하급수적으로 증가하였고 학습을 위해 많은 리소스가 필요해지게 되었다. 이러한 문제점을 감쇄하기 위해서 딥러닝 모델에서 주로 사용하였던 Single Precision 대신 Half Precision을 사용하는 대안이 떠올랐다. Bit 수가 반으로 줄어서 표현할 수 있는 범위가 크게 줄었으나, 계산량과 메모리에서는 크게 개선이 되었다. 그러나 Half Precision으로 학습을 하면 Training Loss가 중간부터 증가하는 문제가 발생하는데, Half Precision으로 표현할 수 있는 수의 범위가 협소하다 보니 중간 이후부

터 작아지는 Gradient의 미세한 오차가 누적되면서 학습이 잘 되지 않게 되는 것이다.

이 문제를 해소하기 위해서 2018년 NVIDIA에서 Mixed-Precision Training을 소개하였는데, 파라미터를 FP32에서 FP16으로 변환한 뒤, Forward Pass와 Backward Pass 모두 FP16으로 연산을 수행한다. 파라미터를 업데이트할 때, 다시 FP32로 변환이 되어 weight가 업데이트가 되는데, 표현 범위보다 작은 Activation Gradients를 Loss Scaling이라고 하는 기법을 통해서 스케일업하는 방법을 적용하였다. NVIDIA Volta와 Turing 아키텍처에서 Tensor Core를 소개하기 시작하면서 현대 GPU는 기존 Single Precision 위주의 학습에서 탈피하여 Single Precision와 Half Precision을 혼합하여 학습 속도를 2~3배 빠르게 할 수 있게 되었다.

2019년에 구글에서 Bfloat16을 소개되는데, Float32와 지수부가 동일한 8bit를 사용하는 특징이 있다. 그래서 FP32보다는 정확도는 떨어지지만, FP16이 가졌던 협소했던 표현 범위의 제약은 해소될 수 있었다. NVIDIA Ampere 아키텍처인 A100, RTX 3000, TPU v2, v3에서 Bfloat16을 하드웨어적으로 지원하기 시작하였다.

2022년에는 트랜스포머 모델에 대해서 Intel, NVIDIA, Arm에서 FP8에 대한 공동 연구를 진행하였고, 이 새로운 데이터포맷은 FP16 대비 Precision Loss가 조금 늘어난 대신 급격한 학습 속도 향상을 가져왔다. 하지만 온전히 FP8로 학습을 하면 트랜스포머 모델을 forward, backward pass 때 요구되는 numeric precision의 차이 때문에 학습이 수렴에 이르기 어렵다. 이를 해결하기 위해서 FP8에 binary interchange format을 도입하는데, 이는 Weight와 Activation Tensor에는 E4M3 포맷을 사용하고 Gradient Tensor에는 E5M2 포맷을 사용하도록 하는 것이다. E4M3의 dynamic range를 확장하기 위해서 특수값 표현의 패턴을 재정의하였고 FP8에 per-tensor scaling을 도입하여 Exponent bias를 소프트웨어적으로 조정할 수 있게 하였다. 연산 커널의 입력은 FP8이지만, 누적 연산이 필요한 경우 FP16 또는 BF16으로 더 넓은 bit precision을 사용한 후에 출력시에 다시 FP8로 변환하도록 함으로써 누적 연산의 정확도 하



〈그림 11〉 딥러닝에서 사용되는 여러 Floating Point 데이터타입의 구조 (FP16, BF16, FP8-E4M3, FP8-E5M2).

락을 최소화한다. 추가로, GEMM 연산의 결과가 non-GEMM (softmax, layernorm, gelu)의 입력이 되는 경우 FP8로 변환하지 않고 높은 precision을 유지한다.

FP8로 학습된 모델을 FP8 추론으로 심플하게 사용할 수 있으나, FP16이나 FP32로 학습된 모델을 FP8로 추론할 경우, Post-Train Quantization과 유사한 Post-Processing 과정이 요구된다. 새로 도입한 Scaling factor를 이용하여 Exponent bias를 시프트 할 수 있게 하여 Weight와 Activation의 다양한 범위에 대응할 수 있다. 특히, GEMM and Add operation은 다이내믹 Scaling factor의 필요성이 GEMM only보다 크다고 알려져 있다.

6. 모델 구조 변형을 통한 경량화

트랜스포머 구조 자체를 하드웨어 친화적으로 변형하는 연구들이 많이 진행되고 있다. Albert 모델^[11]은 트랜스포머 인코더 레이어간의 파라미터를 공유하도록 하고 임베딩 레이어의 Projection 분할을 통해서 모델 파라미터의 개수를 줄였다. MobileBERT 모델^[12]은 기존 MobileNet의 Inverted-Bottleneck을 트랜스포머 모델에 적용하여 BERT-large만큼 레이어는 깊지만, 너비는 얇아진 버전이다. 결과적으로 BERT-base보다 4.3배 작으면서 5.5배 빨라졌고 GLUE 태스크에서 BERT-base 대비 성능 하락은 0.6에 불과하였으나 SQuAD에 있어서는 BERT-base보다 오히려 높은 정확도를 달성하였다. Pixel 4 스마트폰에서 63ms의 latency로 추론할 수 있을 정도의 light-weight BERT 모델이다. 이 가벼운 모델로 무거운 모델의 정확도에 근접하기 위해서 학습의 전략이 중요한데, Progressive knowledge transfer을 통해서 단계적으로 지식을 이식하는 것이 성능 유지에 도움이 되

었다고 한다. 다른 연구^[13]는 BERT 아키텍처에서 NAS를 통해 최적의 파라미터를 추출하여 더 작고 효율적인 아키텍처를 만들었다. BERT-large의 16% 크기, 추론 속도는 ROBERTa의 1.21%, CPU에서 8배 빠른 결과를 가져왔다.

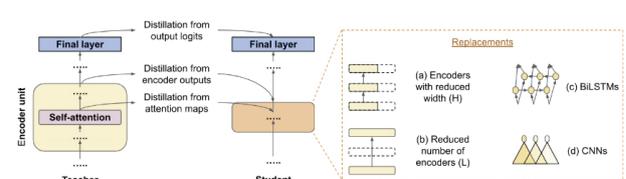
7. 지식 증류 (Knowledge Distillation)

지식 증류는 성능이 좋은 큰 Teacher 모델로부터 작은 Student 모델로 지식을 전파하는 방법이다. 트랜스포머 모델의 일부를 BiLSTM이나 CNN 같은 병렬화가 가능한 효율적인 구조로 대체를 할 수 있다. 또, Encoder output로부터 Semantic, Contextual한 Feature-based Knowledge를 더 작은 embedding size나 encoder unit을 가진 모델로 지식을 증류할 수 있다. 또는 Attention Maps (Relation-based Knowledge)은 입력 토큰간에 contextual 의존성을 가리키는 softmax 분포를 작은 Student 모델이 모사하도록 할 수 있다. Transformer Distillation 방법을 사용함으로써, TinyBERT^[14]는 BERT-base 모델보다 7.5배 작고 추론 속도가 9.4배 빠르면서 정확도 성능을 유지하였고, 6-layer TinyBERT는 Teacher 모델인 BERT-base와 유사한 정확도 성능을 확보하였다.

지식 증류 전략에는 Offline, Online, 그리고 Self 방식이 있다. Offline 방식은 가장 대중적인 방법으로 Pretrained Teacher 모델로부터 Student 모델로 학습 중에 지식을 전파하는 방법이다. Online 방식은 Pretrained Teacher 모델이 제공되지 않을 때, Teacher, Student 모델을 동시에 학습을 하면서 지식을 전파하는 방식이다. 보통 Teacher와 Student 모델의 구조가 다르게 일반적인데, Self 방식은 동일한 구조를 사용한다. 가

		BERT _{LARGE}	BERT _{BASE}	IB-BERT _{LARGE}	MobileBERT	MobileBERT _{TINY}	
embedding	Embedding	1024	768	1024	128	128	
	Dropout	no-op	no-op	no-op	3-convolution	3-convolution	
body	Linear	Dropout	768	512	512	512	
		Dropout	1024	768	1024	128	128
	MHA	#Head	$\begin{pmatrix} 1024 \\ 16 \\ 1024 \end{pmatrix}$	$\begin{pmatrix} 768 \\ 12 \\ 768 \end{pmatrix}$	$\begin{pmatrix} 512 \\ 4 \\ 1024 \end{pmatrix}$	$\begin{pmatrix} 512 \\ 4 \\ 128 \end{pmatrix}$	$\begin{pmatrix} 512 \\ 4 \\ 128 \end{pmatrix}$
		Dropout	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 768 \\ 3072 \\ 768 \end{pmatrix} \times 12$	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 4$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 2$
	FFN	Dropout	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 768 \\ 3072 \\ 768 \end{pmatrix} \times 12$	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 4$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 2$
		Dropout	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 768 \\ 3072 \\ 768 \end{pmatrix} \times 12$	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 4$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 2$
	Linear	Dropout	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 768 \\ 3072 \\ 768 \end{pmatrix} \times 12$	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 4$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 2$
		Dropout	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 768 \\ 3072 \\ 768 \end{pmatrix} \times 12$	$\begin{pmatrix} 1024 \\ 4096 \\ 1024 \end{pmatrix} \times 24$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 4$	$\begin{pmatrix} 128 \\ 512 \\ 128 \end{pmatrix} \times 2$
	#Params		334M	109M	293M	25.3M	15.1M

〈그림 8〉 Attention 그룹핑 [5].



〈그림 13〉 지식 증류. Encoder 채널, Encoder 유닛의 개수, 다른 op으로 대체하는 등의 방법이 있다 [5].



령, DNN의 깊은 레이어에 있는 지식을 얇은 레이어로 전파할 수도 있다.

V. 전망과 결론

2017년 트랜스포머 모델이 소개되었고 2018년 구글의 BERT 등장을 계기로 대규모 언어 모델(LLM)은 여러 NLP 태스크에서 그 우수성을 입증하고 있다. 그 후로 보다 더 크고 강력한 LLM을 구축하기 위한 경쟁에 많은 AI 기업들이 뛰어들었고 2020년 OpenAI의 GPT-3 발표로 175B개의 파라미터를 가진 모델이 등장하였다. 사람처럼 흉내내기 어려울 것이라고 생각했던 논리적 사고와 창작 능력이 필요한 태스크들에서도 놀라운 결과를 만들었다. 그리고 NVIDIA는 Megatron 530B LLM을 소개하였는데, 전문가 훈련 또는 지도학습이 필요 없이도 많은 영역에서 깊이 있는 질문에 대답할 수 있다. 텍스트 프롬프트를 바탕으로 사실적이거나 창의적으로 이미지를 생성할 수 있도록 설계된 Generative Model (DALL-E2, Stable Diffusion)들도 세상을 놀라게 만들었다. 이 모든 LLM의 결과물들이 최근 5년 안에 일어난 일이고 앞으로 인간이 상상하기 어려운 크기의 모델로 계속 발전할 것인지 궁금하다. 이제 언어 모델은 규모, 기술 과시의 영역을 넘어서 서비스화되는 단계에 들어섰다. 구글의 문서 요약 기능 또는 하이퍼클로바의 네이버 노트, 텍스트를 주면 그림을 생성하는 DALL-E에 이르기까지 서비스 영역은 무궁무진하다. 하지만, 컴퓨팅 빈부 격차로 인해서 기업간 인공지능 서비스의 양극화가 발생하고 있으며, 오픈 소스를 지향하는 빅테크 기업들이 폐쇄형으로 전환하면서 보편적인 접근이 불가능해졌다. 대규모 언어 모델의 민주화를 위해서는 모델 경량화는 필수적인 요소다. 한 가지 분명한 것은 앞서 소개한 LLM 모델들이 모두 트랜스포머를 기반으로 만들어진 모델들이라는 것이다. 다행인 것은 대규모 LLM으로 진화할 수록 과잉의 파라미터와 연산, 메모리도 늘어나며, 압축할 수 있는 여지도 커지게 된다. 우리는 여기서 소개된 많은 압축 기술을 이용하여 모델을 경량화할 수 있다. 클라우드 영역에서 제한적으로 가능했던 트랜스포머 모델을 효율적으로 엣지 디바

이스에서 대중적으로 활용할 수 있는 날이 곧 올 것이라고 기대한다.

참고 문헌

- [1] <https://developer.nvidia.com>
- [2] Z. Huang et al., Recent Trends in Deep Learning Based Open-Domain Textual Question Answering Systems, IEEE Access 2020
- [3] R. Alec and Kim, Learning Transferable Visual Models from Natural Language Supervision, pp 8748–8763, PMLR 2021
- [4] A. Vaswani et al., Attention is All You Need, NeurIPS 2017
- [5] P. Ganesh et al., Compressing Large-Scale Transformer-Based Models: A Case Study on BERT, Transactions of the Association for Computational Linguistics, vol. 9 2021
- [6] S. Kim et al., I-BERT: Integer-only BERT Quantization, ICML 2021
- [7] O. Kovaleva et al., Revealing the dark secrets of BERT, EMNLP-IJCNLP 2019, pp 4356–4374
- [8] Y. Tay et al., Synthesizer: Rethinking self-attention in transformer models, arXiv:2005.00743, 2020
- [9] K. Choromanski et al., Rethinking Attention with Performers, ICLR 2021
- [10] <https://hoya012.github.io/blog/Mixed-Precision-Training/>
- [11] Lan et al., Albert: A lite bert for self-supervised learning of language representations, arXiv:1909.11942, 2019
- [12] Sun et al., MobileBERT: A compact task-agnostic BERT for resource-limited devices, arXiv:2004.02984
- [13] Wynter and Perry, Optimal Subarchitecture Extraction for BERT, arXiv:2010.10499
- [14] X. Jiao et al., TinyBERT: Distilling BERT for Natural Language Understanding, EMNLP2020



오지훈

- 2007년 2월 한양대학교 전자전기컴퓨터공학 학사
- 2009년 8월 Georgia Tech 전자전기컴퓨터공학 석사
- 2012년 12월 Georgia Tech 전자전기컴퓨터공학 박사
- 2013년 1월 ~ 2013년 12월 Childrens' National Medical Center 포닥 연구원
- 2014년 1월 ~ 2021년 2월 삼성리서치 책임 연구원
- 2021년 3월 ~ 2021년 12월 삼성리서치 수석 연구원
- 2022년 2월 ~ 현재 Neubla ML Lead Engineer

〈관심 분야〉

Computer Vision, Deep Learning Model Compression,
Medical Imaging

AI Inference를 위한 하드웨어 플랫폼과 소프트웨어 프레임워크의 동향

I. 서론

인공지능 모델의 규모, 즉 필요한 연산 수와 파라미터 수는 지속적으로 증가 해 왔다. 트랜스포머(transformer) 모델이 등장한 이후부터 그 증가 추세가 급격히 심화되어 수백, 수천억 개의 파라미터를 가진 모델들이 등장하고, 연산량뿐 아니라 메모리 footprint도 inference 성능 향상에 있어서 중요한 요소가 되고 있다. 인공지능 모델의 training은 일반적으로 방대한 학습용 데이터를 동시에 대량으로 처리해야 하므로, 많은 병렬처리용 연산기와 큰 메모리 bandwidth가 필요하며, 임의의 연산을 실행해야 하기에 범용성도 확보되어야 한다. 따라서 training에는 고성능 GPU (Graphics Processing Unit)을 이용하는 것이 일반적이다.

반면, 완성된 모델을 이용하여 실제 inference를 실행하기 위한 하드웨어 플랫폼과 소프트웨어 프레임워크는 서비스나 실행 시나리오에 따라 다각화되고 있다. Inference를 수행하는 환경은 크게 datacenter와 edge device로 나뉠 수 있는데, 전자는 사용자의 데이터를 통신망을 경유해 datacenter의 서버로 보낸 다음 서버에서 inference를 실행하고 그 결과를 이용자의 device로 되보내는 방식이다. 후자는 스마트폰이나 TV, 차량, CCTV등 사용자의 디바이스에서 직접 inference를 실행하는 방식이다. 대체로 datacenter에서는 동시 처리량이 많으므로 throughput을 중시하고, edge device에서는 지연(latency)이나 소비 전력을 중시하는 경향이 있다.

이에 따라, AI inference를 효율적으로 실행하기 위한 다양한 하드웨어들이 개발되고 있다. NVIDIA, Intel, AMD등 기존의 프로세서 제조사들은 물론이고, Google, Amazon, Meta 등의 소프트웨어 기업들도 자사 알고리즘에 특화된 AI 가속기를 개발하고 있다. 또한 각 하드웨어



이연복
Intel



의 제조사들은 해당 제품을 효율적으로 활용하기 위한 AI inference 프레임워크들을 제공한다. 물론 소프트웨어 기업들도 여러 하드웨어 타겟의 backend를 지원하고 있다. 본 고에서는 AI Inference에 이용되고 있는 대표적인 하드웨어 플랫폼과 AI inference 프레임워크들을 소개한다.

II. AI Inference 실행에 사용되는 대표적인 하드웨어 플랫폼

표 1은 AI Inference에 사용되고 있는 다양한 하드웨어 플랫폼과 각 제조사에 의해 제공되는 AI 프레임워크 혹은 소프트웨어 개발 도구(SDK: Software Development Kit)를 보여준다. AI Inference에 사용되는 하드웨어는 크게 CPU, GPU, custom AI Chip으로 나눌 수 있으며 본 장에서 각각의 특징을 설명한다.

2.1 CPU (Central Processing Unit)

AI inference에 사용되는 CPU에는 크게 Intel Xeon, AMD Epyc 제품군으로 대표되는 서버 CPU, Intel Core i, AMD Ryzen 제품군으로 대표되는 desktop CPU, 그리

고 ARM기반의 mobile CPU 등이 있다. 각 제품군에 따라 코어나 메모리 구성, clock 주파수 등에 차이가 있다.

CPU는 특정 응용을 빨리 실행하는 것보다 범용성, 호환성을 중시하므로 다양한 명령을 지원하고, 연산의 단위가 작으며, 여러 명령의 조합으로 임의의 연산을 수행 가능하도록 만들어져 있다. 일반적으로 CPU는 GPU나 custom AI 가속기에 비해 data level 병렬성보다는 instruction 수준 병렬성을 활용하기 위한 구조를 가진다. 즉, 하나의 명령으로 여러 데이터를 동시에 처리하기보다 프로그램 진행 흐름을 예측하고 명령들의 의존 관계를 파악해서, 최대한 많은 명령을 끊임 없이 연산 유닛에 공급하는 데 중점을 둔 구조이다. 이를 위한 prediction 회로와 instruction scheduler 회로는 CPU 면적과 전력 소모의 상당 부분을 차지한다. 또한 복수 개의 코어를 활용하여 thread/process 수준 병렬성도 이용 가능 하지만, 일반적으로 수 십개 이하로, GPU에 비하면 훨씬 작다. 따라서, 진행이 규칙적이고 충분히 많은 data level 병렬성을 갖는 프로그램의 경우, GPU나 custom 가속기에 비해 전성비가 낮다. 하지만, 대용량 메모리와 캐시를 활용 가능 하여, LSTM (Long Short-

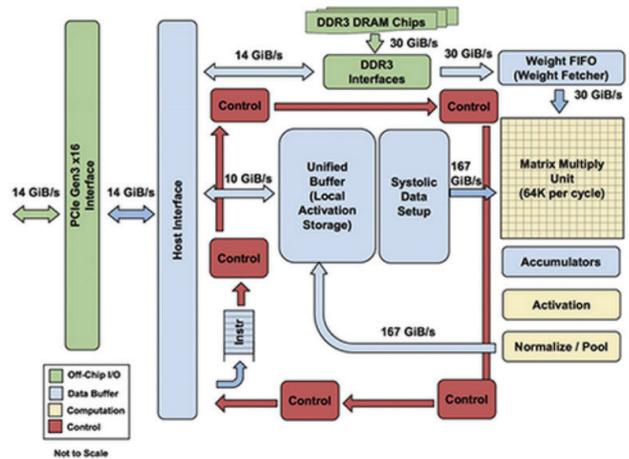
〈표 1〉 다양한 AI 응용 하드웨어 플랫폼 과 제조사별 AI 프레임워크

Company	Computing Power	Architecture	SDK/Framework	Target
Intel Greco	150 TOPS (INT8)	Tensor processor core	SynapseAI	Datacenter
Intel A770 16GB	275 TLOPS (INT8)	GPU	Vino™	Datacenter
Intel EyeQ Ultra	178 (INT8)	Mobile CPU + vision cores	EyeQ Kit	Edge
Intel MyriadX	4 (INT8)	VLIW cores	OpenVino™	Edge
NVIDIA H100 80GB (Tensor core)	624 TOPS (INT8)	GPU	TensorRT	Datacenter
Google TPU4	275 TOPS (INT8)	Systolic Array	TensorFlow	Datacenter
Google Edge TPU	4 TOPS (INT8)	Systolic Array	TensorFlow	Edge
AMD Radeon 7900 XT	64 TFLOPS (FP16)	GPU	ROCm	Datacenter
Qualcomm Cloud A100	400 TOPS (INT8)	AI Core	AIC SDK	Datacenter
Amazon Inferentia	32-512 TOPS (INT8)	Systolic Array	Neuron	Datacenter
Groq GroqChip	750 TOPS (INT8)	Systolic Array	GroqWare	Datacenter
Arm Ethos N77	4 TOPS (INT8)	Compute Engine (MAC units)	Vela compiler	Edge
Tesla FSD chip	73 TOPS (INT8)	NPU (MAC units)	NN Compiler	Edge



Term Memory) 모델과 같은 시계열 알고리즘이나, 트랜스포머 기반 모델과 같이 메모리 접근이 병목이 되는 AI 응용의 경우에는 전성비나 가성비를 고려해 CPU를 채택하는 경우도 많다.

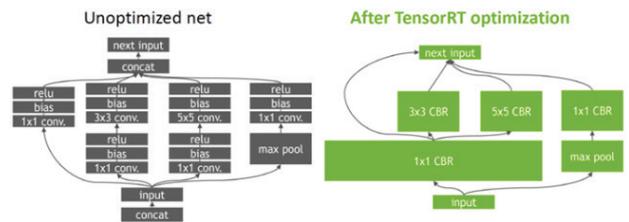
한 편, 스마트폰이나 웨어러블 디바이스에 주로 이용되는 모바일 CPU는 칩 규모와 전력 소모가 매우 작은 만큼 이용 가능한 메모리나 연산량이 매우 한정되어 있어, MobileNet과 같이 depthwise convolution 등을 활용하여 연산량을 줄인 네트워크가 이용되고, 양자화(quantization)에 의한 모델 경량화가 한층 더 중시되는 경향이 있다. 한 편, 내장 GPU를 포함 하고 있는 CPU도 있어, 이런 경우는 CPU와 GPU를 함께 Inference에 활용할 수도 있다.



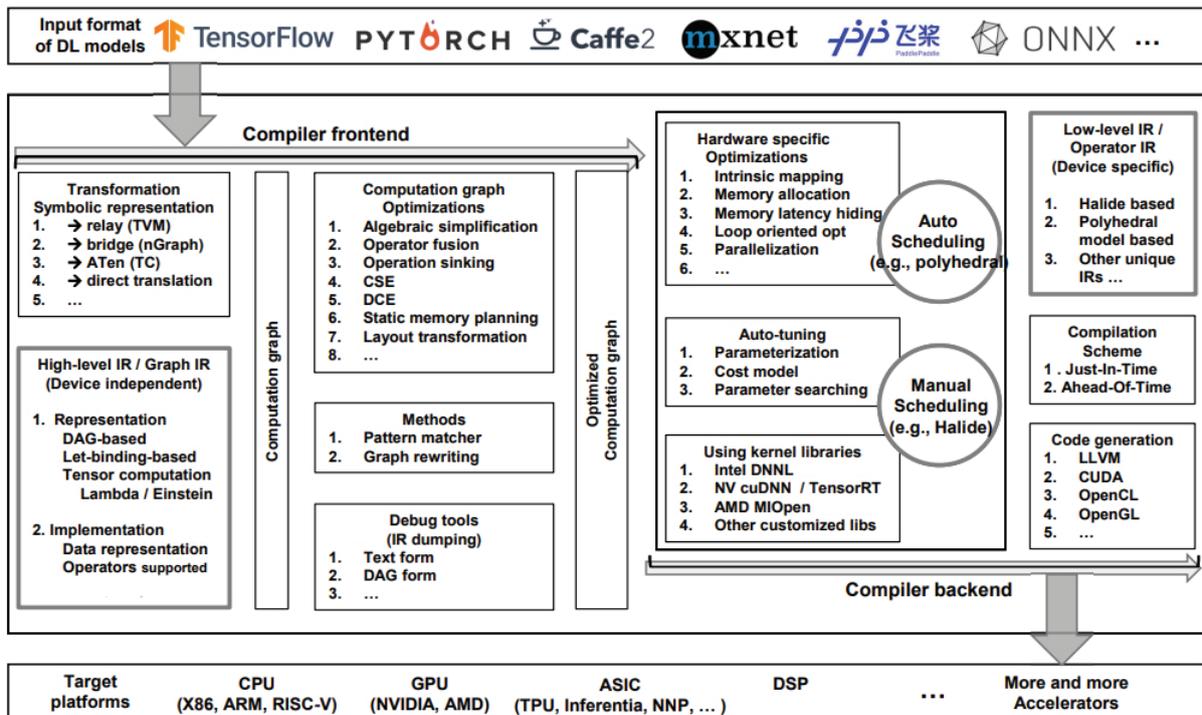
〈그림 1〉 Google TPU V2 Block Diagram [2]

2.2 GPU (Graphic Processing Unit)

GPU는 그래픽스 작업을 가속하기 위해 만들어진 아키텍처이다. 일반적으로 그래픽스 응용은 대량의 픽셀에 대해 동일한 연산을 반복적으로 수행하는데, 이를 효율적으로 처리하기 위해, GPU는 수백 또는 수천 개의 코어로



〈그림 3〉 Layer fusion in TensorRT (Vertial/Horizontal) [1]



〈그림 2〉 AI Framework 에서 적용 되는 최적화 [3]

구성하여 병렬성을 높였다. 이는 단순한 연산의 반복이 많은 신경망 기반의 AI 모델의 요구사항과 맞아, GPU의 발전은 신경망 기반 AI 모델의 비약적 발전에 큰 역할을 하였다. 특히, 신경망 모델의 training에는 대량의 학습 데이터에 대한 병렬 연산이 필요하여, 현재도 training에는 GPU를 이용하는 경우가 압도적으로 많다.

Inference에서도 트랜스포머 모델의 등장 전, AI 모델의 주류를 이루었던 CNN (Convolutional Neural Network) 기반 모델들은 대부분 weight의 재사용성이 높고 계산 집약적인 모델들이었기에, GPU를 사용할 경우 매우 효율이 높았다.

단, GPU는 코어 수가 많고, AI 응용에 쓰이지 않는 graphics 처리 전용 회로도 내포하고 있어 전력 소모가 매우 크며 별도의 비용 부담이 있다. 따라서, 병렬성을 충분히 확보 하지 못하는 AI 모델의 경우에는, 비용 대비 기대 성능을 얻을 수 있는지 고려 할 필요가 있다.

2.3 Custom AI Inference Chip

AI inference에서는 matrix multiplication, convolution, pooling, activation 등 자주 쓰이는 연산들이 있는데, 이러한 연산들에 특화된 AI Inference 전용 하드웨어를 만들 수 있다. AI 가속 하드웨어의 대표적인 예로, Google TPU [2] 등에 사용 된 systolic array를 들 수 있는데, 이는 matrix multiplication을 가속하기 위한 아키텍처로, CPU나 GPU처럼 연산 결과를 레지스터에 쓰는 것이 아니라, 하드웨어 상의 데이터 흐름을 최적

화하여 연산된 결과를 최대한 다른 많은 연산에서 곧바로 사용할 수 있도록 구현한 것이다.

AI inference 전용 custom AI 칩들은, 특정 범위의 AI 응용을 실행하여 전력소모/면적 대비 최대 성능을 얻기 위해 불필요한 다른 하드웨어 부분을 줄이고 AI 응용을 돌리는데 필요한 가속 하드웨어 위주로 구성된 경우가 많다. 최근에는 NVIDIA 의 V100, A100 의 Tensor Core, Intel Arc GPU의 XMX Core와 같이, 고성능 GPU, CPU 제품에도 이러한 AI 연산 전용 하드웨어를 경쟁적으로 탑재하는 추세이다. 또한, TPU V3처럼 inference뿐만 아니라 training도 함께 가속하기 위한 하드웨어도 개발되고 있다.

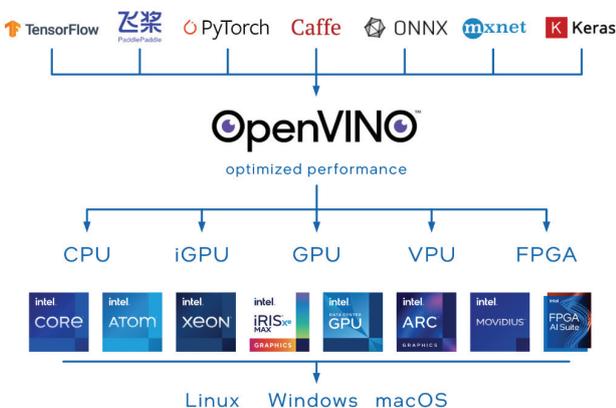
III. AI 소프트웨어 프레임워크

AI 소프트웨어 프레임워크는 AI 응용을 실행하는데 필요한 런타임과 커널 그리고, 유용한 라이브러리와 API 들을 가진 소프트웨어 도구를 말한다. AI Inference 성능을 높이기 위해서는 하드웨어 의존적인 최적화가 필요한 경우가 많다. <표 1>에서와 같이, 각 하드웨어의 제조사들은 이러한 하드웨어 특성을 고려한 최적화를 적용 가능하게 해주는 AI 소프트웨어 프레임워크를 제공하고 있다. 한편, Google의 TensorFlow, Meta 의 PyTorch 와 같은 범용 AI 프레임워크들도 이러한 제조사가 제공하는 커널 라이브러리카 각종 최적화 backend 모듈을 지원하고 있다.

본 장에서는 3.1절 에서 AI 프레임워크에서 적용되고 있는 공통적인 최적화 기술들을 소개하고, 3.2절 이후 잘 알려진 몇 가지 AI 프레임워크들을 소개한다.

3.1 AI 응용의 대표적인 최적화 기술

<그림 2>에서 볼 수 있듯이 AI 응용의 최적화는 Graph level 부터 저수준 명령 레벨까지 다양한 단계에서 여러가지 기술들이 적용된다. 즉, TensorFlow 나 PyTorch 등으로 training 된 모델이 주어지면, 일반적으로 최초에 각 프레임워크 의 IR (Intermediate Representation)로 변환하는 작업이 수행된다. 이러한



<그림 4> Intel OpenVino™ Framework [6]



IR은 대부분 모델을 computation graph 형식으로 나타낸 것이며, OpenVino™의 IR (구 NGraph IR)^[4]를 예로 들 수 있다.

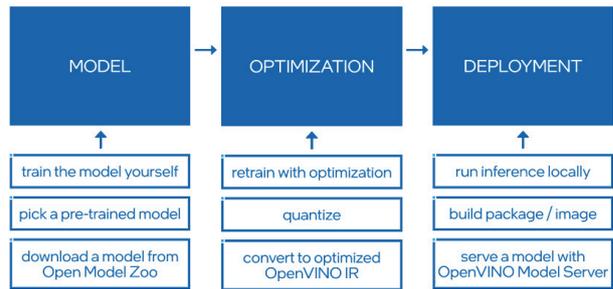
Graph로 표현된 IR, 즉 graph level IR 상에서 여러 가지 최적화를 거치는 작업을 graph compilation 또는 graph optimization이라고 한다. 또한, 많은 tool들이 여러 수준의 IR을 가진다. Graph IR의 각 node는 최종적으로 kernel source code로 변환되어 기계어로 컴파일되는데, 이 과정에서 Graph보다 저수준의 IR이 이용되는 경우가 많다. 이하, 대표적인 graph level 최적화 기술들을 소개한다.

- **Data Layout 최적화** : 각 레이어의 데이터 저장 순서를 최적화하는 것을 말한다. 데이터를 접근할 때, 최적화 커널이나 가속 하드웨어의 특성에 따라, 특정 메모리 layout이 유리한 경우가 많다. 예를 들어 convolution kernel의 경우, channel 방향의 데이터가 메모리상에 먼저 저장되고, 다음으로 width, height 방향으로 저장하는 것이 유리하도록 kernel 혹은 하드웨어가 설계되어 있는 경우가 많다. 이러한 특성을 고려한 data layout을 결정하고, 적절히 재구성해주는 최적화가 적용된다.

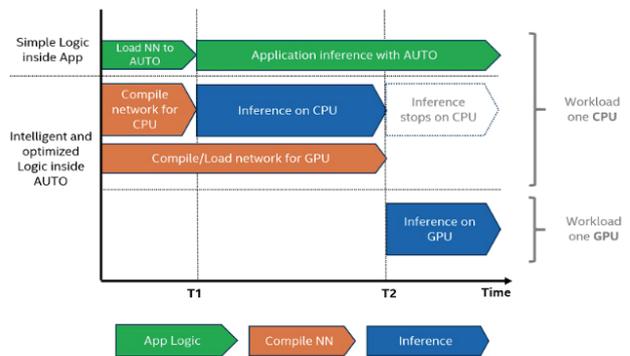
- **Layer Fusion** : <그림 2>에서와 같이 레이어들을 복수의 레이어를 각각 따로 실행하지 않고, 한 번에 처리하는 것을 말한다. 예를 들어 convolution, pooling, add 연산들이 연속적으로 실행될 경우, 각각의 연산에 대한 kernel을 따로 실행하는 것보다 한 번에 처리하는 것이, latency나 메모리 사용량 모든 측면에서 효율적이다. 일반적으로 이러한 fusing을 적용하기 위해서는 fused 커널이나 post processing을 한번에 처리해 주기 위한 하드웨어 기능이 존재해야 한다.

- **Layer Conversion** : 하드웨어 기능이나 이용 가능한 커널에 따라 특정 연산을 다른 연산으로 대체하는 것이 유리할 수 있다. 예를 들면 deconvolution을 transposed convolution으로 대체하거나 reduce-mean operation을 average-pool로 바꾸기도 한다.

- **하드웨어 가속 명령 활용** : Systolic array 등의 가속 하드웨어를 적절히 사용하기 위해서는 해당 최적화 kernel을 구현하고 런타임 parameter들을 설정해야 한다.



<그림 5> OpenVino™의 workflow^[6]



<그림 6> OpenVino™ Auto Plugin을 활용한 First inference latency 최적화^[7]

- **Kernel tuning** : data size나 shape, 사용되는 하드웨어 플랫폼 상의 연산 유닛, 메모리 사이즈 등 각종 조건에 따라, 커널의 실행 단위나 진행 방향, tiling size 등의 파라미터들이 잘 설정되어야 한다.

- **Memory optimization** : Cache, local 메모리 등을 잘 활용하기 위한 tiling, 메모리 bank 접근효율을 위한 접근 순서 조정, 레이어 간 메모리 reuse, device-host 간 data이동 최소화 등이 적용될 수 있다.

- **Batch size 결정** : 메모리 사이즈, 연산 유닛 개수 등 하드웨어 플랫폼의 특성을 고려하여 한번에 실행할 최적의 batch size 등을 결정해 주는 기능을 제공할 수 있다.

- **Multi-device optimization** : CPU의 bit/little core, 혹은 CPU와 GPU, AI custom chip 등, heterogeneous한 구성에서 multi device를 활용할 수 있을 때, 레이어 단위 혹은 모델, stream 단위로 각 디바이스에 적절히 분배, 할당해서 성능을 높일 수 있다.

3.2 Intel OpenVINO™ [6]

Intel은 Xeon, Core i 제품군 등 CPU뿐만 아니라 HD Graphics, Iris Graphics, XE Graphics 등 내장 GPU를 함께 개발해 왔다. 또한 MyriadX, VPU 등의 다양한 AI custom chip도 개발하고 있으며 최근에는 Arc 제품군으로 외장 GPU 도 출시하는 등 다양한 AI 응용 가속용 하드웨어를 생산하고 있다.

OpenVino™ 는 이러한 Intel 하드웨어에서 inference 를 최적화하기 위한 오픈소스 프레임워크다. <그림 4>에서 보이는 것과 같이 대부분의 AI 모델 포맷을 지원하며, Intel 의 CPU, GPU, VPU (MyriadX), FPGA 등의 하드웨어 Plugin을 지원하고 있다.

<그림 5>는 OpenVino™를 사용하기 위한 기본적인 workflow를 보여준다. 먼저, 주어진 trained model을 Model Optimizer라는 최적화 툴을 이용하여 OpenVino™ IR로 바꾸어 준다. 그런 다음 해당 IR을 Python/C++ API를 통해 컴파일하고 Inference를 실행하게 된다.

OpenVino™ 의 최적화는 모든 플러그인에 공통적으로 적용되는 IR 레벨의 common transformation, 그리고 각 플러그인 별 최적화로 나뉘어진다. 선술 한 layer fusion, kernel optimization, 메모리 reuse, layer conversion, 등의 모든 최적화가 적용되고 있다. 최적화 커널은 CPU는 OneDNN [5]을, GPU는 OpenCL 기반의 자체 커널 (a.k.a clDNN)과 OneDNN 을 둘 다 활용하고 있다.

또한 OpenVino™ 는 Auto Plugin과 Auto Device configuration 기능을 제공하여, 다양한 하드웨어로 구성된 시스템 레벨에서의 자원 분배를 자동으로 최적화해 준다. Auto Plugin을 사용하면, 이용 가능한 하드웨어를 자동으로 감지하고, 각 레이어들을 적절한 리소스에 할당한다. Auto Device Configuration 기능은 각 리소스에서 이용 가능한 최적의 batch size, stream 개수 등을 계산해 준다. 이를 이용하여 사용자는 다음 code에서 보여주는 것과 같은 최소한의 설정으로 최적의 시스템 성능을 얻을 수 있다.

```
from openvino.runtime import Core
ie = Core() # Initialize OpenVino
model = ie.read_model(model="model.xml")
compiled_model = ie.compile_model(model=model,
device_name = "AUTO")
```

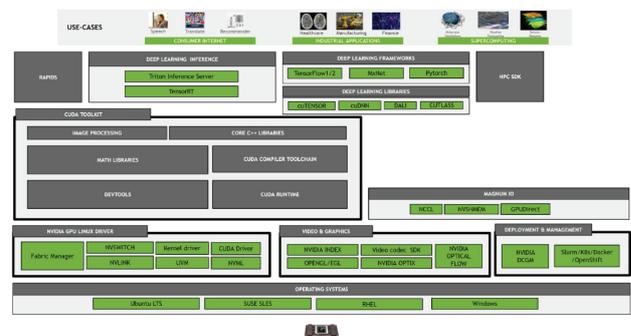
Auto plugin 은 <그림 5>에서의 예시처럼, First inference latency를 줄이는 목적으로도 활용 된다. GPU Plugin은 커널 컴파일을 런타임에 수행하는 JIT (Just In Time) 방식을 사용하고, 최적화 복잡도도 높으므로, 최초에 모델을 컴파일하는 시간이 CPU보다 많이 걸린다. Auto Plugin을 이용하면, GPU 타겟으로 컴파일하는 동안, CPU를 이용하여 최초의 inference를 실행하는 것이 가능해진다.

3.3 NVIDIA TensorRT [1]

TensorRT는 NVIDIA에서 제작한 자사의 GPU를 대상으로 한 AI Inference 최적화 도구이다. <그림 7>에서는 NVIDIA GPU의 소프트웨어 구성을 보여준다.

TensorRT는 CUDA, CuDNN (CUDA Deep Neural Network Library) 등의 커널 library를 활용하여, layer fusion 등의 graph optimization, kernel auto tuning, precision 최적화 (FP16 or INT8), memory reuse 등의 최적화를 적용하고 있다 [9]. [9]에 의하면 TensorRT로 실행할 경우, 동일한 GPU에서 native TensorFlow만으로 실행한경우에 비해 6배의 성능 향상을 얻을 수 있었다고 한다.

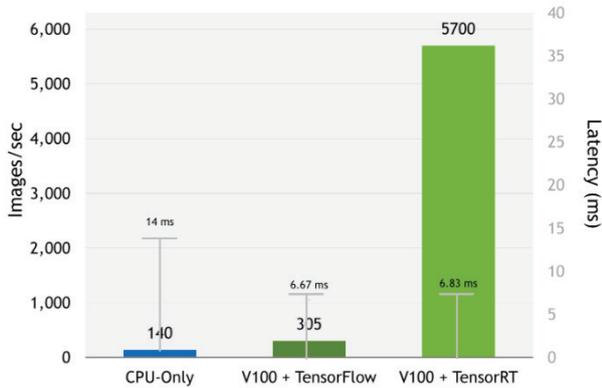
TensorRT를 사용하는 방법은 1) 학습 완료 된 모델



<그림 7> NVIDIA AI Framework Software Stack [1]

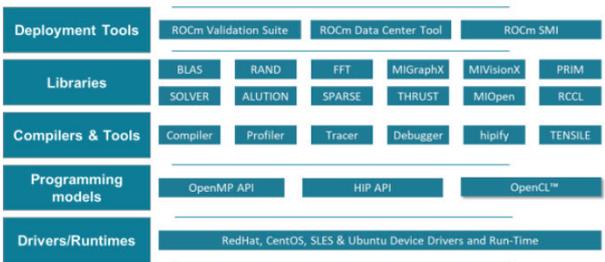


Up to 40x Faster CNNs on V100 vs. CPU-Only Under 7ms Latency (ResNet50)

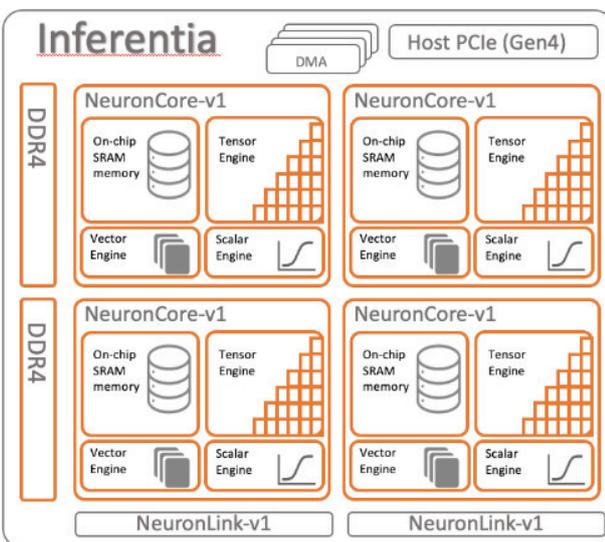


Inference throughput (images/sec) on ResNet50. **V100 + TensorRT**: NVIDIA TensorRT (FP16), batch size 39, Tesla V100-SXM2-16GB, E5-2690 v4@2.60GHz 3.5GHz Turbo (Broadwell) HT On **V100 + TensorFlow**: Preview of volta optimized TensorFlow (FP16), batch size 2, Tesla V100-PCI-E-16GB, E5-2690 v4@2.60GHz 3.5GHz Turbo (Broadwell) HT On. **CPU-Only**: Intel Xeon-D 1587 Broadwell-E CPU and Intel DL SDK. Score doubled to comprehend Intel's stated claim of 2x performance improvement on Skylake with AVX512.

〈그림 8〉 Performance improvement for NVIDIA GPU with TensorRT [8]



〈그림 9〉 AMD ROCm Software Stack [11]



〈그림 10〉 AWS Inferentia architecture [12]

을 TensorRT 엔진으로 직접 변환하여 이용하는 방법, 2) TensorFlow 와 PyTorch에 포함된 TensorFlow-TensorRT, Torch-TensorRT 기능을 이용하는 방법 두 가지가 있다.

한편, 최근 중요도가 높아진 트랜스포머 모델들을 가속하기 위한 boost 소프트웨어로 FasterTransformer [10]도 제공되고 있는데, 이를 이용하면 multi-head attention과 같이 트랜스포머에서 자주 사용되는 패턴들을 최적화 커널로 퓨징 함으로서 성능 향상을 얻을 수 있다.

3.4 AMD ROCm [11]

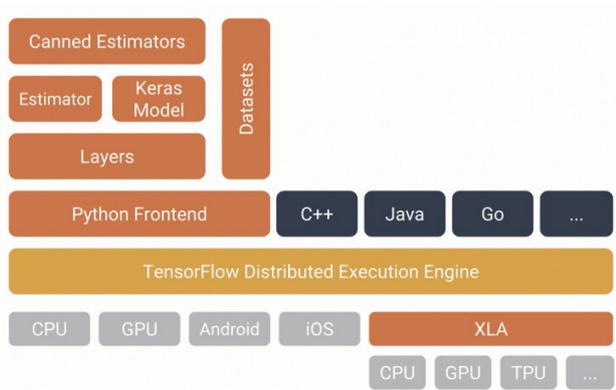
AMD CPU와 GPU에서 AI 응용을 실행하기 위한 프레임워크로는 ROCm (Radeon Open Compute Stack) [11]이 있다. ROCm 역시 오픈소스로 공개되어 있으며, 소프트웨어 stack 은 〈그림 9〉와 같이 구성되어 있다. TensorFlow와PyTorch를 지원하며, graph compiler인 MIOpen, 최적화kernel library인 MIOpen 을 포함한다. Native tool로서 혹은 ONNX Runtime의 backend 로 이용이 가능하다.

3.5 AWS Neuron [12]

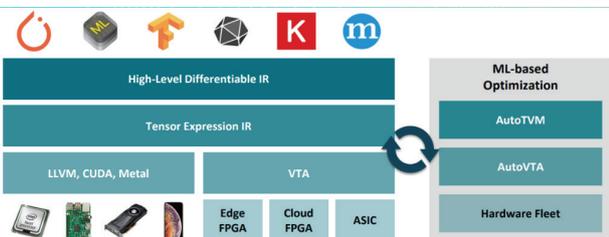
Amazon AWS에서 개발한 AI inference 가속용 칩인 inferentia를 위한 최적화 프레임워크로, AWS Neuron이 있다. 각 Inferentia Device는 〈그림 10〉과 같이 4개의 Neuron 코어로 이루어져 있다. Neuron은 TensorFlow와 PyTorch에 integration 되어 있으며, 기본적으로는 각 모델을 각 Neuron 코어에 대하여 컴파일한다. 한편, 최적화를 위해 여러 개의 Neuron 코어에 대해 모델을 한꺼번에 batch compilation 하는 기능도 제공하고 있고, 모델을 분할하여 자동으로 여러 개의 Neuron 코어에서 파이프라인 형태로 실행하는 최적화 기능도 제공한다.

3.6 Google TensorFlow/XLA [13]

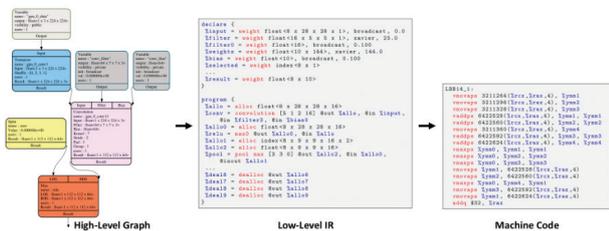
Google 의 TensorFlow는 Meta의 PyTorch와 더불어 AI training에 가장 많이 사용되는 프레임워크이다. Inference에 TensorFlow를 사용할 경우, backend를



〈그림 11〉 TensorFlow XLA [13]



〈그림 12〉 Apache TVM



〈그림 13〉 Glow Compiler Workflow

TensorFlow-TRT나 TensorFlow 에 integration 된 OpenVino™ 등 타겟 하드웨어 플랫폼 의 최적화 엔진 을 함께 사용 하는것이 효과적이다. 이러한 방식은, 모델의 graph를 partitioning하여 만들어진 subgraph 를 OpenVino™나 TensorRT를 통해 실행하게 된다.

또는, 〈그림 11〉에서 보여지는 TensorFlow XLA (Accelerated Linear Algebra) [13] 도 적용해볼 수 있다. TensorFlow XLA는 Tensorflow 프레임워크 의 AI compiler로, 그래프를 많은 작은 커널 단위로 실행하는 것이 아니라, 더 큰 커널로 묶어서 컴파일 함으로써 성능을 최적화하는 틀이다. 기계어로 런타임에 컴파일되는 JIT (Just In Time)과 실행 전에 바이너리를 미리 컴파

일 하는 AoT (Ahead of Time) 두가지 방식이 이용 가능하며, JIT 방식은 training에, AoT 방식은 Inference에 활용할 수 있다.

XLA는 TensorFlow모델의 레이어들을 HLO (HighLevel Operation) 라는 IR로 변환하고, affine transform, instruction fusion, common subexpression elimination, constant folding 등의 HLO 레벨 에서의 변환/최적화를 거친 후, LLVM [4] 컴파일러를 통해 타겟 플랫폼의 기계어로 변환된다. 또한, 이 과정에서 CuDNN [15] 과 CuBLAS [16] 등의 최적화 라이브러리를 사용한다. 현재 NVIDIA GPU, Intel CPU, TPU 등을 타겟으로 지원 중이며 custom backend를 직접 구현하여 등록할 수도 있다.

3.7 Apache TVM [17]

TVM 은 Apache 오픈소스 AI compiler 프레임워크이다. 〈그림 12〉 는 TVM 의 workflow를 보여준다. 즉, IR 레벨에서 constant folding, dead code elimination, layout transformation, scaling factor folding 등의 최적화가 이루어지고, tensor operator description이라는 것을 통해 각 operator의 실제 구현을 정의한다. 〈그림 12〉에서 보듯이, 대부분의 AI 프레임워크 frontend를 지원 하며, 기본적으로 지원하는 CPU, GPU 이외에 여러 가지 custom backend를 추가 할 수 있다. 또한 Auto TVM 기능을 제공하여 다양한 하드웨어 플랫폼의 최적화 커널 템플릿의 타일 사이즈 등 각 파라미터들을 자동으로 튜닝할 수 있다.

3.8 Meta PyTorch/Glow [18]/AITemplate

PyTorch 는 Tensorflow와 더불어 AI training에서 가장 많이 사용되는 프레임워크이며, Torch-TensorRT, Intel extension 등을 이용하여 GPU나 CPU backend를 가속할 수 있다.

Meta에서는 AI inference 가속용 컴파일러인 Glow [18]를 개발하였으며, JIT/AoT 두 가지 모드가 있다. Backend로는 CPU/GPU 를 지원하며, custom backend 를 등록하는 것도 용이하다. 〈그림 13〉은 Glow compiler



의 workflow를 나타내며, 입력 모델이 target 하드웨어 플랫폼의 기계어로 변환되기까지 크게 High level graph IR / Low level IR을 거치며, 각각의 단계에서 최적화를 하게 된다. Graph level 에서는 transpose sinking, dead code elimination, concat optimization, common sub-expression elimination 등의 최적화가 이루어지고, Low level IR에서는 operator stacking 등의 최적화가 적용되고 있다. 이 중에서도 operator stacking은 Glow의 중요한 최적화의 하나로, 그래프 레벨에서의 layer fusing과 비슷하게 몇 개의 operation을 별도의 커널이 아니라 연속된 명령으로 실행하도록 해 주는 것이다. layer fusing의 경우에는 fusing을 적용한 커널이 존재해야 fusing을 적용 가능하지만, operator stacking은 자동으로 그러한 kernel을 만들어 준다는 장점이 있다.

한편, Meta에서는 최근 NVIDIA와 AMD GPU를 대상으로 한 가속용 프레임워크인 AITemplate^[19]를 발표하였다. AITemplate은, AI모델을 GPU용의 최적화 된 C++ 코드로 변환해 주는 툴이다. AITemplate은 다양한 그래프 변환을 담당하는 frontend layer와, GPU용 c++ kernel template을 생성하는 backend layer의 두 부분으로 구성된다. CUDA와 ROCm의 런타임 라이브러리를 이용하며, layer fusion 등의 일반적인 최적화에 더하여 multi-head attention block 최적화 등 트랜스포머 block를 위한 최적화도 포함하고 있다.

IV. 전망과 결론

본 고에서는 AI inference를 최적화하기 위해 사용되는 대표적인 하드웨어와 소프트웨어 프레임워크들에 대하여 소개하였다. 다양한 AI 서비스가 고도화, 복잡화 되어 가는 추세에서, 모델의 정확도뿐만 아니라 inference 성능도 중요해지고 있다. 고도의 inference 최적화를 위해서 소프트웨어 특성과 하드웨어 특성이 함께 고려되어야 하기에, 전통적인 하드웨어 제조사와 소프트웨어 기업의 역할도 점점 경계가 없어지고 있다. 즉, 하드웨어의 제조사들은 자사 제품을 효율적으로 활용하기 위한 AI inference 프레임워크들을 개발하고, 소프트웨어 기업들

은 자사의 서비스에 특화된 하드웨어를 직접 개발하고 있다. 다양한 하드웨어와 소프트웨어 프레임워크들이 경쟁적으로 출시되고 있기에, AI 응용을 실제 서비스로 다루는 경우, 이들의 특성을 정확히 이해하는 것이 중요해졌다. 3장에서 설명한 것처럼, TensorFlow나 PyTorch와 같은 범용 AI 프레임워크에 타겟 하드웨어의 가속 엔진을 이용하는 것은 어느 정도 성능 향상에 도움이 될 수 있다. 하지만, 모델을 실행할 하드웨어 환경의 제조사에서 제공하는 프레임워크가 있는 경우에는, 이들을 이용하는 것이, 해당 하드웨어에서 가정하고 있는 최적의 결과물을 얻을 가능성이 크다. 한편, NVIDIA의 솔루션과 프레임워크는 NVIDIA GPU만을 가속할 수 있지만, CPU를 포함하는 multi device 환경에서는 OpenVino와 같이 다양한 하드웨어 Plugin을 지원하며 이들 간의 자원 할당을 자동으로 최적화해 주는 기능을 갖춘 프레임워크를 이용하는 것이 유리할 수 있다. 또한 본 고에서는 graph level의 최적화를 주로 소개하였는데, 양자화(quantization) 등의 경량화도 성능 향상에 큰 도움이 될 수 있으며, 이 또한 다양한 툴이 혼재한다.

빠르게 복잡하고 거대해져 가는 AI 모델의 발전은 하드웨어와 소프트웨어 codesign의 필요성을 극대화 하였고, 관련 분야의 발전에도 기여하였으며, 더불어 유래없이 많은 기술들이 빠르게 개발되고 소프트웨어 툴과 하드웨어 제품들이 생겨나고 있다. 모델과 플랫폼의 상호 발전을 가속하는 경향은 앞으로도 계속될 것으로 전망되며, On chip training 등 다양한 도전적인 과제들이 해결되어 갈 것이라 기대된다.

참고 문헌

- [1] Jay Rodge et al., "Accelerate Deep Learning Inference in Production with TensorRT", GTC Digital November, 2021.
- [2] Google Cloud, "An in-depth look at Google's first Tensor Processing Unit (TPU)", <https://cloud.google.com/blog/products/ai-machine-learning/an-in-depth-look-at-googles-first-tensor-processing-unit-tpu>
- [3] Mingzhen Li et al., "The deep learning ompiler: A comprehensive survey", IEE trans. on Parallel and Distributed



Systems, vol.32, issue 3, Mar. 2021.

[4] Deep Learning Network Intermediate Representation and Operation Sets in OpenVINO™, https://docs.openvino.ai/latest/openvino_docs_MO_DG_IR_and_opsets.html

[5] OneDNN : <https://github.com/oneapi-src/oneDNN>

[6] OpenVino™, <https://www.intel.com/content/www/us/en/developer/tools/openvino-toolkit/overview.html>

[7] Intel OpenVino™ Technical Blog, "Automatic Device Selection and Configuration with OpenVino," <https://blog.openvino.ai/blog-posts/automatic-device-selection-and-configuration>, November 2022.

[8] NVIDIA Technical Blog : "TensorRT3 : Faster TensorFlow Inference and Volta Support", "<https://developer.nvidia.com/blog/tensorrt-3-faster-tensorflow-inference/>

[9] Jay Rodge et al., "Accelerate Deep Learning Inference in Production with TensorRT", GTC Digital November, 2021

[10] FasterTransformer : <https://github.com/NVIDIA/FasterTransformer>

[11] AMD ROCm : <https://rocmdocs.amd.com/en/latest/>

[12] AWS Neuron Documentation: <https://awsdocs-neuron.readthedocs-hosted.com/en/latest/index.html>

[13] XLA: Optimizing Compiler for Machine Learning, <https://www.tensorflow.org/xla>

[14] The LLVM Compiler Infrastructure: <https://llvm.org/>

[15] NVIDIA CuDNN <https://developer.nvidia.com/cudnn>

[16] CuBLAS : Basic Linear Algebra on NVIDIA GPUs <https://developer.nvidia.com/cublas>

[17] Tianqi Chen et al., "TVM: an automated end-to-end optimizing compiler for deep learning," Proc of the 13th USENIX Conference on Operating Systems Design and Implementation, Oct. 2018.

[18] Nadav Rotem et al., "Glow: Graph Lowering Compiler Techniques for Neural Networks," arXiv, <http://arxiv.org/abs/1805.00907>, Aug. 2018.

[19] AITemplate : <https://github.com/facebookincubator/AITemplate>



이 연 복

- 2006년 3월 사이타마대학교 정보시스템공학과 학사
- 2008년 3월 동경대학교 전자공학과 석사
- 2011년 3월 동경대학교 전자공학과 박사
- 2011년 4월 ~ 2017년 12월 삼성전자 삼성리서치 책임연구원
- 2018년 1월 ~ 2018년 10월 SK Telecom 종합기술원 연구원
- 2018년 10월 ~ 2022년 8월 인텔코리아 IoTG 시니어 소프트웨어 엔지니어
- 2022년 9월 ~ 현재 Intel Coreporation NSWE 시니어 소프트웨어 엔지니어

〈관심 분야〉

Computer Architecture, SW Framework, Compiler, AI model optimization, Computer Aided Design, GPU kernel optimization, System Software

전자설계자동화 기술과 차세대 지능형반도체 소자 및 회로 설계 응용

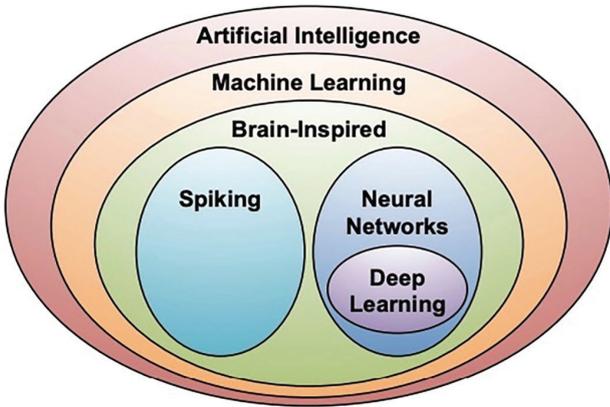
I. 서론

현재 우리는 디지털 대전환(digital transformation)을 맞고 있다. 그레고리 비알은 디지털 대전환을 어떤 조직이 정보, 컴퓨팅, 통신 및 연결 기술 등의 디지털 기술을 융합하여 해당 조직의 상품, 서비스 혹은 운영에 개선하기 위해 중대한 변화를 일으키는 과정으로 정의 하였다^[1]. 이는 기존의 데이터(data)를 유한한 자릿수의 숫자로 나타내는 방식으로 변환하는 디지털 데이터화(digitation)나 디지털화(digitalization)에 기반을 두지만 더한층 높은 수준의 총체적 개념이다. 무형의 비트로 된 디지털 데이터와 그 데이터를 1차 가공한 정보에 국한되지 않으며, 정보 수집, 처리, 저장, 그리고 공유하는 일련의 작업과 이에 필요한 하드웨어와 소프트웨어 상품, 서비스 그리고 운영을 모두 포함한다.

디지털 대전환의 시대에는 필연적으로 데이터의 양이 폭발적으로 증가하게 될 것이다. 데이터를 획득하는 센서(sensor)가 전통적인 컴퓨터와 통신기기뿐만 아니라, 사람, 자동차, 시계, 로봇 등 모든 사물(IoT, internet of things)에 탑재되어 상시 운용된다. 더 나아가, 이러한 사물이 인공지능(artificial intelligence, AI)를 통해 지능화(AIoT, artificial intelligence of things)되어 사람을 대신해 획득한 데이터를 기반으로 정보를 생산하고, 학습을 통해 지식을 축적하여, 지식을 기반으로 지적인 추론과 판단을 하게 된다. 그 대표적인 사례가 자동차이다. 자율주행 자동차(혹은 스마트 카)라고 불리는 현대의 자동차에는 기존에 사람이 수행하던 운전과 길 찾기 등 지능이 필요한 대부분의 활동을 대신해주는 제어 장치가 탑재되어 사람에게는 최소한의 관여만 요구하게 되었다. 앞으로는 이러한 변화가 자동차뿐만 아니라 우리 생활 전반으로 확산될 것이다.



정 성 엽
차세대융합기술연구원



〈그림 1〉 인공지능(artificial intelligence)의 맥락에서 DNN(deep neural network)과 SNN(spiking neural network)의 분류^[2]

사물의 지능화를 가져온 가장 대표적인 인공지능 알고리즘이 DNN(deep neural network)이었다 〈그림 1〉^[2]. 하지만, 그 연산이 복잡하고 그 양이 방대하여 기존의 CPU에서는 원하는 성능을 달성하기가 어려웠으며, 슈퍼컴퓨터 혹은 GPU와 같은 특수한 장치가 필요하였다. 그 중 가장 유명한 사례가 Google DeepMind의 AlphaGo로 1202개의 CPU와 176개의 GPU로 구성된 분산형 하드웨어를 사용하였다^[3]. 이를 1세대 지능형 반도체라고 부를 수 있다. 하지만, 최근 들어 CPU의 clock 주파수 등 성능이 더 이상 개선되지 않고 있으며, multi-core 형태로 한계에 다다른 CPU를 복수 개 집적하는 방향은 근본적인 von Neumann 구조가 가지고 있는 메모리 접근에 대한 전력 소모와 신호 지연이 극심해 저전력 고성능 인공지능 구현이 제한적이다.

이 같은 배경에서, SoC(system-on-chip) 형태로 컴퓨터 구조의 한계 극복을 시도하는 2세대 지능형 반도체가 연구 개발되었다. 테슬라의 자율주행용 칩이 그 대표적인 사례로 DNN의 한 계열로 영상 분석에 특화된 CNN(convolutional neural network) 알고리즘을 반도체 SoC로 구현하기 위해 3개의 Quad A72 CPU, 1개의 GPU와 더불어 DNN 기반 인공지능 알고리즘의 핵심인 합성곱 연산을 담당하는 9,216개의 MAC(multiply-and-accumulate) 회로와 가중치(weight)의 저장을 담당하는 32 MiB의 SRAM을 집적한 2개의 NPU(neural processing unit) 및 각 블록을 연결하는 NoC(network

on chip) 회로 등을 집적하였다. 한편, DNN을 구성하는 뉴런은 매우 단순화된 형식뉴런이라는 모델을 기반으로 하고 있으며, 뉴런이 발화 할지 여부(activation)를 뉴런 간의 경로이자 스위치인 시냅스의 가중치(w_i)와 신호(x_i)의 합성곱($\sum w_i \cdot x_i$)의 크기로 판단함에 있어 가중치를 저장할 메모리와 합성곱 연산을 수행할 CPU 간의 제한적인 대역폭 문제가 상존하여 저지연과 저전력이 요구되는 응용 분야에서 네트워크의 사이즈 증가를 기반으로 하는 인공지능 성능 개선이 자유롭지 못하다^[2].

연구/상업적으로 가장 성공적인 DNN 기반의 인공신경망과 가속기를 일상생활에서 접하고 있는 지금 이 시점에, SNN(spiking neural network)과 비휘발성 저항변화 메모리(nonvolatile resistive switching memory)로 대변되는 3세대 지능형 반도체에 대한 연구가 지속되고 있는 이유도 바로 이 때문이다 〈그림 1〉. SNN은 동물의 신경망을 모방하여 만든 인공신경망으로 기존의 DNN이 tensor 혹은 실수값(floating number)의 행렬곱 연산을 하는 반면, SNN은 특정 시점에 발생한 spike가 누적되어 활동 전위가 임계점을 넘으면 발화를 하는 단순한 원리 덕분에 낮은 전력 소모를 달성할 수 있다. 한편, 인텔의 CMOS기반 TrueNorth 계열 SNN 하드웨어는 소비 전력 측면에서 SNN의 장점을 온전히 활용하고 있지 못하고 있다^[4]. 그러나, ReRAM의 경우 비휘발성메모리 특성으로 인해 SRAM의 휘발성메모리에 비해 저전력 달성에 용이하다. 또한, ReRAM 기반 CBA(crossbar array)는 인텔의 CMOS기반 SNN 하드웨어가 여전히 덧셈기가 필요할 수 있는 반면 CBA만으로도 합성곱이 가능한 장점이 있다. 더불어, ReRAM은 SNN뿐만 아니라 DNN 기반 합성곱 연산기에도 사용될 수 있기 때문에 더 큰 주목을 받고 있다.

현재까지 DNN기반 지능형 반도체에 비해 상대적으로 열세인 SNN기반 지능형 반도체의 효과적인 학습 알고리즘 부재 및 실질적인 추론 응용처 검증 부족을 극복하기 위해서 전자설계자동화를 활용한 설계 및 최적화 관련 연구가 필수적이게 되었다.



II. 전자설계자동화

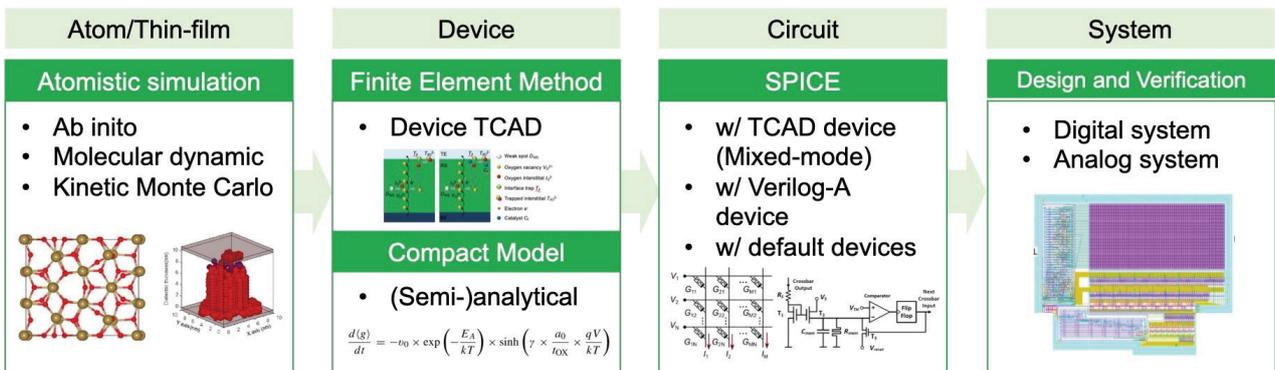
집적회로(integrated circuit)를 개발/생산하는 과정은 크게 설계(design)와 공정(technology)로 구분되며, 전자설계자동화(electronic design automation)는 컴퓨터 및 프레임워크를 활용한 집적회로 설계 과정의 자동화를 뜻한다. 집적회로가 발명되기 이전의 전자회로의 설계는 대략적인 손과 종이를 사용한 분석을 수행한 후 개별적인 전자소자들을 만능 보드(bread board)에 꽂아 특성을 측정하고 원하는 특성을 얻을 때까지 전자소자들을 바꿔가며 실험을 수행하는 과정을 반복하였다. 하지만, 집적회로의 개발에는 기관상에 회로가 제작된 이후에 그 특성이 요구조건을 충족하지 않더라도 수정할 수 없고 새롭게 제작하여야 하기 때문에 기존의 시행착오를 반복하는 방식은 개발 시간과 비용의 막대한 손실을 발생시키게 된다. 또한, 무어의 법칙을 따라 증가해온 집적회로의 소자의 개수도 미리 그 특성을 예측할 필요성을 낳았다.

전자설계자동화의 종류는 집적회로 이루는 시스템-모듈-회로-소자-원자의 구성 성분의 각 수준에 따라 분류하면 이해하기 쉽다 <그림 2>. 시스템-모듈 수준의 전자설계자동화의 발전은 1970년대 Applicon사, Calma사, Computervision사가 배치(layout) 설계를 자동화하는 기능을 대상으로 이니 컴퓨터와 전용 OS(operating system)를 활용하여 턴 키 시스템 설계 자동화를 최초로 제공한 컴퓨터 지원 설계(computer-aided design, CAD)가 1세대라고 본다. 이후, 1980년대, Mentor Graphics사, Daisy System사, Valid Logic System사

가 논리 설계(logic design) 기능을 추가하여 엔지니어링 워크스테이션(EWS)과 유닉스(Unix) OS를 활용하여 설계자 1인 1대를 가지고 자동화된 설계를 수행한 시대를 열었다.

회로 수준의 전자설계자동화는 SPICE(simulation program with integrated circuit emphasis)가 대표적이다. 캘리포니아 대학교 버클리의 Laurence Nagel이 1971년 개발한 SPICE1을 모태로, 1975년 SPICE2, 1989년 Thomas Quarles가 C언어로 개발한 SPICE3를 거쳐 상용 SPICE인 HSPICE(Synopsys), PSpice(Cadence Design Systems), SmartSPICE(Silvaco) 그리고 LTSpice(Analog Devices)와 오픈-소스 SPICE인 ngspice, XSpice(조지아텍) 등으로 이어져 왔으며 세대를 거쳐 분석의 종류(DC, AC, transient, noise, transfer function 등), 소자 압축 모델(compact model)의 종류, 입력 출력 기능이 강화되어 왔다. SPICE는 기본적으로 SPICE의 코어는 회로의 nodal analysis를 수행하며 비선형 미분방정식을 내재적분법(implicit integration method), 뉴턴법(Newton's method), 희소행렬법(sparse matrix technique)으로 푼다.

소자 수준의 전자설계자동화는 압축모델과 TCAD(technology computer-aided design)를 꼽을 수 있다. 회로 성분들의 정적-동적 전류-전압 특성의 수학적 표현을 모델이라고 하며, 압축 모델은 특히 물리적으로 엄밀성, 간결성, 정확성을 갖춘 수학적으로 표현된 모델 혹은 컴퓨터 코드화 된 모델을 뜻한다 [5]. 기본적으로



<그림 2> 저장변화 메모리 소자를 위한 원자/박막-소자-회로-시스템 수준 전자설계자동화

회로 설계를 위해 SPICE에 탑재되어 활용된다. 압축 모델은 주로 Verilog-A 언어로 표현되며^[6], 주로 닫힌 형태의 방정식(closed-form equation)이 빠른 회로 nodal analysis를 가능하게 하므로 선호되지만 저항 변화 메모리와 같이 동작 원리가복잡한 소자의 경우 정확성을 확보하기 위하여 semi-analytical 하거나 for, if, 등의 분기문(branch) 형태로 표현되기도 한다.

TCAD는 유한요소해석법을 기반으로 전체 소자의 체적을 미소 체적으로 나누어 각 노드에서 결합된 방정식(coupled equations)을 자기일관성(self-consistent) 있게 해를 구하는 방법을 말한다. 1960년대 말 BJT 소자의 설계를 위해 착안되었으나, 1980년대 중반을 거치며 twin-well 기반 CMOS 기술의 디자인 룰과 기생성분(대표적으로 latch-up)의 분석을 위하여 필수적인 요소로 자리 잡았다^[7, 8].

원자 수준의 전자설계자동화는 세부적으로 순이론적(ab initio), 분자동력학(molecular dynamics), 활동적 몬테 카를로(kinetic Monte Carlo) 시뮬레이션 기법 등으로 나눌 수 있다. 수 nm 수준의 스케일링에 의한 양자역학적 현상의 반영과 새로운 재료 및 동작 원리를 기반으로 하는 신소자의 다중물리 시뮬레이션에 대한 수요가 원자 수준 전자설계자동화의 발전을 이끌고 있다.

이제 다음 장에서는 몇 가지 예를 들어 차세대지능형반도체의 저항변화 현상 기반의 메모리 소자와 집적 회로에 대한 구체적인 전자설계자동화 사례에 대해 살펴본다.

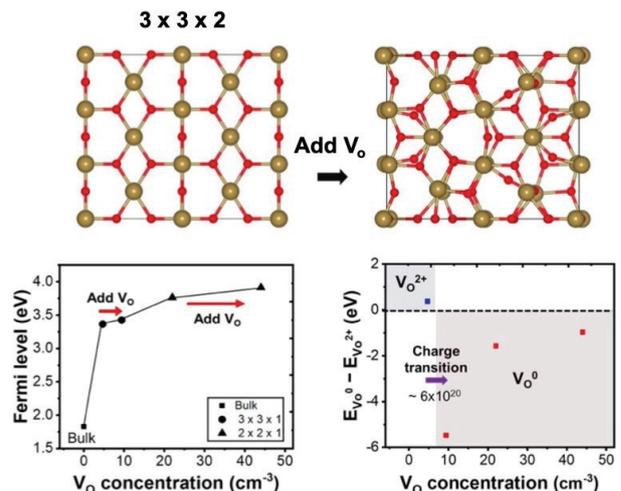
III. 전자설계자동화의 저항변화 메모리 설계 응용

1. 저항변화 현상의 원자 수준 시뮬레이션

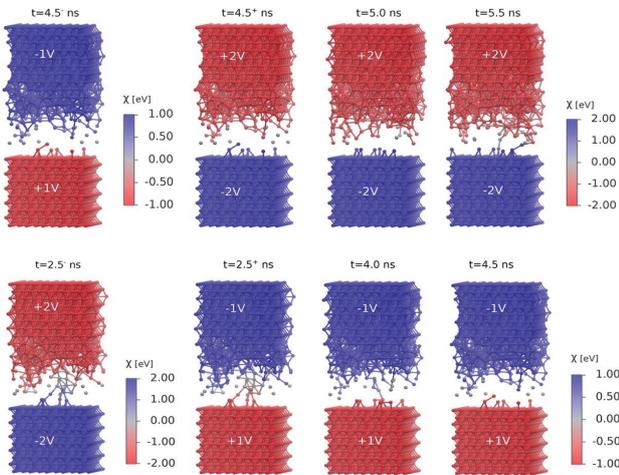
순이론적 기법은 다체 슈뢰딩거 식(many-body Schrodinger equation)을 Roothaan equation으로 표현하여 Hartree-Fock 기법을 사용해 해를 구하는 전통적인 방법과 Kohn-Sham equation으로 표현하여 밀도범함수이론(density-functional theory)을 사용해 해를 구하는 최근의 방법으로 대표 된다. 그림 3의 사례는 밀도범함수이론을 기반으로 산소 공공이 도입되었을 때의

격자 구조의 변화와 산소 공공의 농도에 따른 페르미 준위 및 화학 포텐셜 값의 변화를 보여준다^[9]. 이 같은 연구 결과는 원자 수준의 분자 동력학 시뮬레이션 혹은 활동적 몬테카를로 시뮬레이션 뿐만 아니라 소자 수준의 유한요소해석법 시뮬레이션에 사용되는 물리량을 이론적인 값으로 활용 될 수 있다.

분자 동력학 시뮬레이션 기법은 물리계에서 원자 혹은 분자들 사이의 포텐셜 혹은 힘이 주어졌을 때 이를 이용해서 뉴턴의 운동방정식을 수치적으로 풀어내는 방법을 뜻한다. 후술할 몬테 카를로 방법과는 달리 힘의 방향으로 원자를 움직여 주기 때문에 결정론적인 시뮬레이션 방법이다. 저항변화 메모리에 분자 동력학 시뮬레이션 결과를 제시한 그림 4의 사례는 600개의 원자로 이루어진 활성전극과 800개의 원자로 이루어진 비활성전극 그리고 900개 원자로 구성된 비정질HfO₂ 기반 저항변화 메모리의 저항변화 구동 사이클 전체를 모사하였다^[10]. 이는 2.5x2.5 nm² 면적과 1.5 nm 두께를 갖는 비정질 HfO₂ 로 이루어진 계를 의미 한다. 해당 시뮬레이션을 통하여 valence change mechanism 으로 동작하는 저항변화 메모리 소자 내부에서 상대적인 산소 농도가 낮은 영역 감소가 Hf의 원자가 변화를 일으켜 LRS상태에서 전기전도도가 발생하는 원리를 모사한다. 특히, 저항 변화 현상의 원인을 나노미터 수준의 공간에서의 국소적인 O 이



〈그림 3〉 밀도범함수 이론기반 순이론적 계산으로 획득한 산소 공공의 농도에 따른 페르미 준위와 화학 포텐셜 값의 변화 양상^[9]

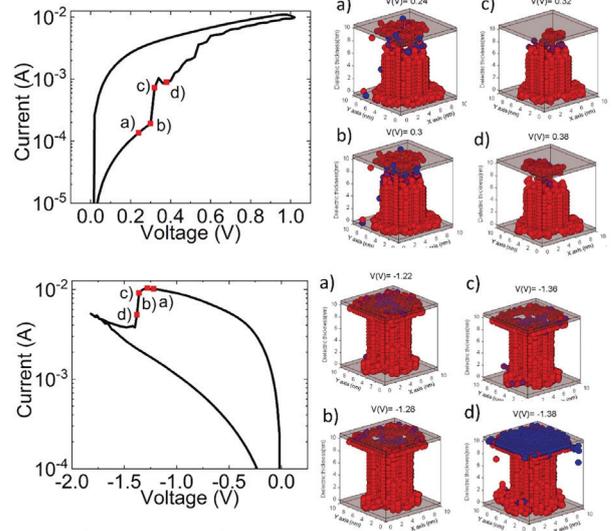


〈그림 4〉 분자 동역학 시뮬레이션 기법을 통해 계산한 SET(상)과 RESET(하) 조건에서 Hf 원자 전위 값의 sub-ns 변화^[10]

동에 의해 발생한다는 이해를 할 수 있게 되었다.

저항변화 메모리에 적용되는 활동적 몬테카를로 기법은 입자의 호핑과 산화/환원 반응을 transition-state theory를 기반으로 Maxwell-Boltzmann statistic과 정합성을 갖는 Eyring-Polanyi equation으로 모사하고 전기적 전도 현상을 trap-assisted tunneling으로 모사하여 각각의 iteration time step에 대해 해를 구한다^[11]. 해당 연구는 10x10 nm² 면적과 10 nm 두께를 갖는 HfO₂ 기반 저항변화 메모리의 SET과 RESET 과정의 I-V 특성을 제시한다 〈그림 5〉. 전도성 필라멘트의 생성 및 소멸과 저항 변화 현상의 근본적인 확률론적 거동을 이해하는데 도움을 줌과 동시에, 원자/이온의 배치 상태와 I-V 특성을 연관시켜 이해 있다는 점에서 앞서 살펴본 다른 원자 수준 시뮬레이션과는 차별성이 있다.

이러한 최신의 연구 결과를 살펴보면, 현재의 컴퓨팅 자원이 하에서, 원자 수준의 시뮬레이션 기법을 활용하여 저항변화 유전 박막 수준 혹은 일부 소자 수준까지 시뮬레이션이 가능함을 파악할 수 있다. 하지만, 두께와 면적으로 대변되는 구조체의 크기가 커질 경우에도 소자 수준의 시뮬레이션에 필요한 컴퓨팅 자원이 비합리적인 수준으로 커질 수 있음을 유념하여야 하며, 여전히 해당 기법을 회로 수준으로 확장하는 것에서는 분명한 제약이 따른다. 그럼에도 불구하고, 낮은 스케일에서 획득한 저



〈그림 5〉 활동적 몬테 카를로 시뮬레이션 기법을 통해 계산된 SET(상)과 RESET(하) 저항 변화 특성과 원자/이온 배치 상태(a-d)^[11]

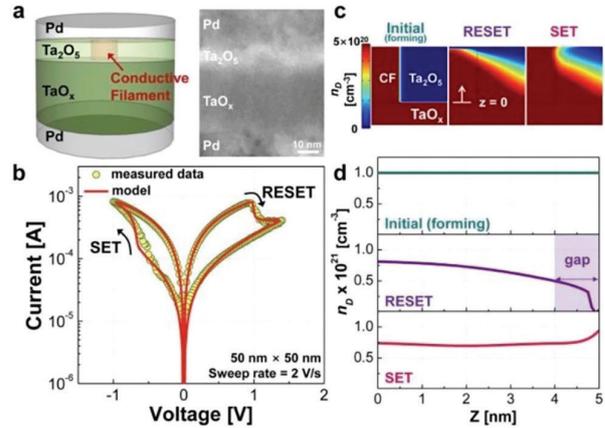
항 변화 현상에 대한 이해는 재료 물성 최적화와 최종적으로 소자의 스케일링 및 대규모 CBA 회로 구현을 위해 선결되어야 할 확률론적 거동의 제어에 기여 할 수 있을 것이다.

2. 저항변화 메모리 소자 유한요소해석법 시뮬레이션

저항변화 메모리 소자의 유한요소해석법 시뮬레이션은 산화-환원 반응 방정식 (redox reaction equation), 이온과 전자 분포를 묘사하는 포와송 방정식(Poisson equation), 이온과 전자의 연속 방정식(ion continuity equation and electron continuity equation), 저항에 의한 발열과 열 흐름을 묘사하는 격자 열 유동 방정식 (lattice heat flow equation)을 풀게 된다. 이 같은 유한요소해석법 기반 시뮬레이션 기능을 제공하는 상용 전자설계자동화 프로그램으로는 Altsoft사의 COMSOL Multiphysics와 Silvaco사의 Victory Device가 대표적이다.

저항변화 메모리 소자 유한요소해석법 시뮬레이션을 COMSOL Multiphysics를 통해 구현한 연구 결과〈그림 6〉로 김성호 교수가 초기에 발표한 두 편의 논문을 주목할 필요가 있다^[12, 13]. 2013년의 논문에서는 산소 공공의 이동 방정식에 Fick 확산과 표류 및 생성을 다루었

<p>• Dependent variables</p> <p>n_D Concentration of V_o [cm^{-3}]</p> <p>T Temperature [K]</p> <p>ψ Potential [V]</p>	<p>• Constants</p> <p>a Hopping distance, 0.05 nm β Mesh size, 0.5 nm</p> <p>f Attempt-to-escape frequency, 10^{13} Hz</p> <p>E_a Diffusion barrier, 1.5 eV ($\psi > 0$) or 2.5 eV ($\psi < 0$)</p> <p>A Coefficient of generation rate, 10^{24} $\text{cm}^{-3}\text{s}^{-1}$</p>
<p>• Oxygen vacancy transport</p> <p>Eq.(1) $\frac{\partial n_D}{\partial t} = \nabla \cdot (D \nabla n_D - v n_D) + G$</p>	<p>Diffusivity(D), Velocity(v) → Eq.(1)</p> <p>→ n_D → Eq.(5)</p>
<p>• Current continuity</p> <p>Eq.(2) $\nabla \cdot \sigma \nabla \psi = 0$</p>	<p>Temperature(T) → Eqs.(4)</p> <p>Conductivity(σ) → Eq.(2)</p>
<p>• Heat (Joule heating)</p> <p>Eq.(3) $-\nabla \cdot k_{th} \nabla T = J \cdot E = \sigma \nabla \psi ^2$</p>	<p>E-field(E) → Eq.(3)</p> <p>Potential(ψ) → Eq.(2)</p>
<p>• Parameters from physics- Eqs.(4)</p> <p>$D = 1/2 \cdot a^2 \cdot f \cdot \exp(-E_a/kT)$ Diffusivity of V_o [cm^2s^{-1}]</p> <p>$v = a \cdot f \cdot \exp(-E_a/kT) \cdot \sinh(qaE/kT)$ Drift velocity of V_o [cm/s]</p> <p>$G = A \exp(-(E_b - q\beta E)/kT)$ Generation rate of V_o at $\psi < 0$ [$\text{cm}^{-3}\text{s}^{-1}$]</p>	
<p>• Parameters from measurements and assumptions- Eq.(5)</p> <p>$\sigma = \sigma_0 \exp(-E_{AC}/kT)$ Conductivity [$\Omega^{-1}\text{cm}^{-1}$]</p> <p>$k_{th}$ Thermal conductivity [$\text{Wm}^{-1}\text{K}^{-1}$]</p>	



〈그림 5〉 활동적 몬테 카를로 시뮬레이션 기법을 통해 계산된 SET(상)과 RESET(하) 저항 변화 특성과 원자/이온 배치 상태(a-d) ^[11]

(a) Species Continuity Equation:

$$\frac{\partial [X]}{\partial t} = -D_X \nabla \cdot \left\{ \left(\frac{z_0 q_0}{k_B T} \right) [X] \mathbf{E} + \nabla [X] \right\} + (G_X - R_X)$$

$$[X] = V_0^{2+}, I_0^{2-}, D_{WS}$$

$$D_X = D_{I_0^{2-}}$$

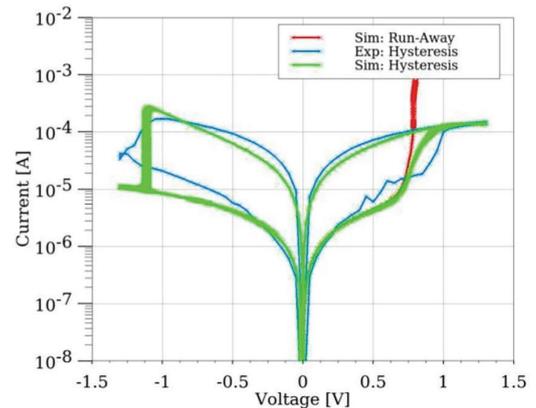
$$G_{V_0^{2+}} = G_{I_0^{2-}} = R_{D_{WS}} = k_f [D_{WS}] \exp\left(-\frac{E_{br} \pm \frac{2+k}{3} x z p_0 E}{k_B T}\right)$$

$$R_{V_0^{2+}} = R_{I_0^{2-}} = G_{D_{WS}} = k_r [V_0^{2+}] [I_0^{2-}] \exp\left(-\frac{E_{br} \mp \frac{2+k}{3} (1-x) z p_0 E}{k_B T}\right)$$

$$f_c = \exp\left(-\frac{p_1 E}{k_B T}\right)$$

(b) Extended Bridging Mobility Model:

$$\mu_{eff} = \mu_s + w(\delta, f_0) [\mu_c \exp(-E_b/k_B T) - \mu_s]$$



〈그림 7〉 Silvaco Victory를 활용한 유한요소해석법 기반 저항 변화 메모리 소자의 소자 수준 시뮬레이션 사례 : (좌) 결합된 방정식과 상수 및 매개변수 (우) 선택 소자의 유무에 따른 전류-전압 특성의 시뮬레이션 결과 차이와 실측값과의 비교 ^[14].

고, 이후 2014년의 논문은 생성 성분(G) 대신 온도 구배에 의한 산소 공공의 Soret 확산 성분을 고려하였다. COMSOL Multiphysics는 이같이 결합된 방정식의 수정과 추가가 용이하여 저항 변화 메커니즘 연구에 도움이 된다.

한편, Silvaco사의 Victory Device의 경우 동사의 회로 시뮬레이션 환경인 SmartSPICE와 연계한 mixed-mode simulation을 통해 유한요소해석법 소자 시뮬레이션 결과를 회로 시뮬레이션에 활용하여 회로 수준의 분석이 가능하다. 그림 7의 사례는 선택 소자(select transistor)를 추가함으로써 current run-away 억제 할

수 있음을 mixed-mode simulation을 통해 제시한 결과이다 ^[14]. 해당 결과는 프랑스 그르노블 소재 CEA-Leti의 130 nm CMOS 공정을 활용해 Weebit-nano가 제작한 1 트랜지스터-1 저항(1T-1R)으로 구성된 셀의 실측값과의 비교하였을 때 준수한 정합성을 보인다. 이 같은 방법은 소자 이외의 기생 저항 혹은 축전기의 영향이나 구동 회로와 연계한 분석에 활용 될 수 있을 것이다.

유한요소해석법 기반 시뮬레이션은, 공통적으로, 원자 수준 시뮬레이션 기법과 비교했을 때 연산에 소요되는 시간과 양이 적은 장점이 있다. 따라서, 소자의 스케일링에 따른 소자 특성의 변화를 반복적인 실험 이전에 예측



해볼 수 있으며, 소자 제작 공정에 의한 미세한 구조적 변화에 따른 소자 특성의 변화 등을 가능해 볼 수 있다. 하지만, mixed-mode simulation을 활용한다고 하여도 현재 200개의 노드, 300개의 소자, 그리고 10개의 유한요소해석법의 소자로 구성된 회로 규모에 한정되어 있기 때문에^[15], 고집적 회로 수준의 설계에 활용하기에는 여전히 제약이 있다.

3. 저항변화 메모리 소자 콤팩트 모델과 회로 SPICE 시뮬레이션

저항변화 메모리의 대표적인 압축 모델로는 북경대-스탠포드대(PKU-SU)의 압축 모델^[15]과 (아르헨티나)국립기술대-(아르헨티나)국가과학기술연구협의회-바르셀로나자치대의 동적 멤다이오드 압축 모델(dynamic memdiode compact model, DMM)을 꼽을 수 있다^[17].

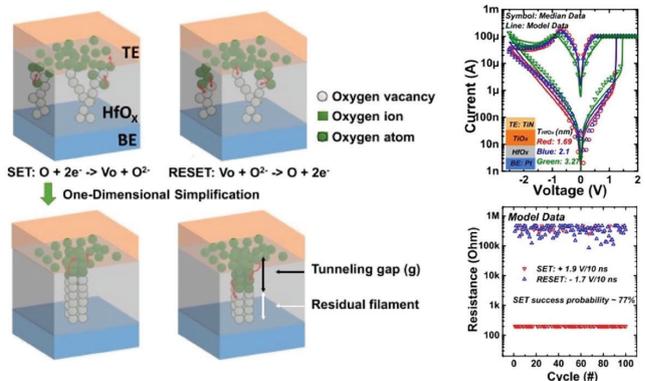
PKU-SU 모델은 그림 8에서와 같이 SET 및 RESET 상태의 원자/이온 배치를 1차원적 산소 공공으로 구성된 필라멘트와 산소 이온으로 구성된 갭의 길이(g) 변화로 묘사하고 갭의 길이의 시간에 대한 변화율을 전압에 대한 함수로 모델링한다. 이때, 잡음 특성과 확률론적 저항변화 현상을 모델링하기 위한 난수($\chi(t)$)를 사용하는 특징이 있다. 다만, 준해석적(semi-analytical) 형태와 난수의 사용과 같은 모델의 특징에 의해 복잡성이 높고 Verilog-A 언어로 작성되어 있다. 이는 SPICE 설계 시 속도를 저하시키는 요인으로 작용한다.

DMM 모델은 그림 9에서와 같이 서로 역방향으로 병

렬 연결된 두 개의 다이오드를 통해 전달 방정식을 표현하여 SET과 RESET 동작을 구분하고, 축전기의 양극 전압을 저항변화 메모리의 상태를 나타내는 형태로 표현되어 있다. 이 같은 메모리 방정식은 축전지의 양극에 연결된 두 개의 SET 및 RESET 저항과 SET 저항에 연결된 전압 소스를 통해 회로적으로 표현 될 수 있다. 따라서, PKU-SU 모델 대비 간단한 subckt 형태로 netlist에 포함되어 SPICE 시뮬레이션에 활용 될 수 있다. 이 같은 특장점으로 인해 1280개의 시냅스로 구성된 DNN기반 인공신경망이 8x8 픽셀의 MNIST 데이터 세트를 학습/추론하는 동작을 SPICE 시뮬레이션으로 직접 묘사할 수 있게 되었고, 어레이 회로의 도선 저항에 따른 쓰기 전압 및 시간의 관계 그리고 쓰기 전압에 따른 추론 정확도 등과 같은 회로 수준의 동작을 예측해 볼 수 있게 되었다.

한편, SNN기반 인공신경망의 학습 방법에 따른 동작을 예측한 결과가 Yu Wang 교수 그룹에서 발표 되었다^[4, 18]. 해당 연구는 앞서 살펴본 PKU-SU 모델의 초기 버전을 사용하였으며, 784×500×500×10 크기의 네트워크를 갖는 저항변화 메모리의 CBA와 더불어 leaky integration and fire (LIF)를 구현하는 뉴런 회로를 함께 시뮬레이션한 점에서 의의가 있다. 그 결과 다층 네트워크 구성이 용이하고 정확도가 높은 NSLS(neural sampling learning scheme)학습 방식을 사용하였을 때, 인식 테스트를 실시간 (~1μs/샘플)으로 수행 할 수 있음을 확인하였다.

- Evolution rate of gap $\frac{d(g)}{dt} = -v_0 \times \exp\left(-\frac{E_A}{kT}\right) \times \sinh\left(\gamma \times \frac{a_0}{t_{OX}} \times \frac{qV}{kT}\right)$
- Temporal noise $g|_{t+\Delta t} = \int \left(\frac{d(g)}{dt} + \delta_g \times \chi(t)\right) dt$
- Stochastic switching $\delta_g(T) = \frac{\delta_g^0}{\left[1 + \exp\left(\frac{U_{SMTH}-T}{T_{SMTH}}\right)\right]}$
- Local field enhancement factor $\gamma = \gamma_0 - \beta \cdot g^a$
- Local temperature variation $T = T_0 + V \times I \times R_{TH}$
- Current $I = I_0 \times \exp\left(-\frac{g}{g_0}\right) \times \sinh\left(\frac{V}{V_0}\right)$

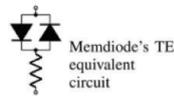


(그림 8) PKU-SU 압축 모델의 구성과 모델링 결과 : (좌) 결합된 방정식의 세부 구성 (중) SET 및 RESET 상태의 실제 소자 내부의 원자/이온 배치 상태와 1차원 근사 (하) 대표적인 I-V 특성과 확률론적인 저항변화(stochastic switching) 특성 모델링 결과^[16].

• Transport equation

$$I = I_0(\lambda) \left[e^{\beta a(\lambda)(V-IR_S(\lambda))} - e^{-(1-\beta)a(\lambda)(V-IR_S(\lambda))} \right]$$

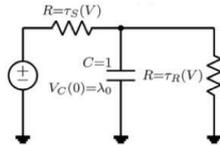
$$I_0(\lambda) = I_{min}(1-\lambda) + \lambda I_{max} \quad \text{same for } a(\lambda) \text{ and } R_S(\lambda)$$



• Memory equation

$$\frac{d\lambda}{dt} = \frac{1-\lambda}{\tau_S(V)} - \frac{\lambda}{\tau_R(V)}$$

$$\tau_{S,R}(V) = \tau_{0S,0R} e^{-\frac{V}{V_{0S,0R}}}$$

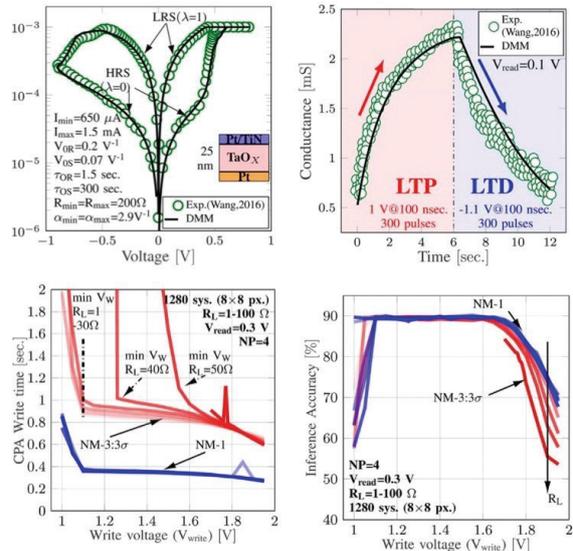


• Switching dynamics equation

$$V_{TS} = V_{0S} \ln(RR) + V_{0S} \ln\left(\frac{\tau_{0S}}{V_{0S}}\right)$$

$$\lambda_N = (\lambda_0 - 1)e^{-(NM/\tau_S)} + 1$$

$$T_{write} = -\frac{\tau_{0S} e^{-\frac{V}{V_{0S}}}}{DC} \ln\left(\frac{I_{LRS} - I_T}{I_{LRS} - I_{HRS}}\right)$$



〈그림 9〉 압축 모델의 구성과 저항 변화 소자 및 회로의 대표적인 특성 모사 결과 : (좌) 결합된 방전식의 세부 구성 (중) 전도 모델과 메모리 방전식의 기본이 되는 멤다이오드와 축전지 모델 (우) 전류-전압, 장기 소성(long-term potentiation and depression), 어레이 회로의 도선 저항에 따른 쓰기 전압 및 시간의 관계, 그리고 쓰기 전압에 따른 추론 정확도 특성 [17].

IV. 전망과 결론

전자설계자동화에 대한 관심이 증가하고 설계(design)와 기술(technology)의 개념이 종래의 독립적인 개념에서 설계-기술 공동 최적화(design-technology co-optimization) 관점의 융합된 개념으로 변화하고 있다. 이는, 전력-성능-면적의 설계 지표를 사용하여 새로운 재료, 소자 구조 및 공정 조건 등을 효율적으로 평가하고 하향식으로 선택하는 접근법과 특정 재료, 소자 구조 및 공정 조건을 가정하여 전력-성능-면적 특성을 예측하는 상향식 최적화 접근법이 새로운 기술 노드에서의 개발 비용과 제품화 소요 시간을 낮추기 위한 방안으로 매우 효과적이기 때문이다.

한편, 앞으로의 전자설계자동화는 새로운 개념의 소자가 갖는 다양한 물리적-화학적-열적인 현상을 반영하고, 극한의 스케일링에 따른 양자 역학적인 현상을 반영하기 위한 원자-소자 수준의 시뮬레이션 또한 더욱 중요해질 것으로 예상된다. 국내 파운드리 산업과 팹리스 산업의 현황은 한국에서 전자설계자동화 산업이 성장할 비옥한 토양이 될 것이며, 전자설계자동화 산업을 잘 육성한다면 궁극적으로는 한국의 반도체 산업을 떠받치는 든

든한 세 번째 축으로 자리매김 할 것으로 판단된다.

전자설계자동화 도구에 대한 연구와 개발은 여러 연구자들이 새로운 개념과 구조의 신소자 연구와 이를 활용한 회로 및 시스템 수준의 연구에 관심을 갖는 지금이 중요한 시점이라 사료된다. 한국의 파운드리와 종합 반도체 회사에서 제작한 반도체가 전세계에서 판매되는 눈부신 성과의 이면에 반도체 설계에 사용되는 전자설계자동화 도구를 전량 수입에 의존하고 있는 현실이 있다. 평판 MOSFET에서 FinFET으로 그리고 최신의 GAAFET(gate-all-around FET) 이후의 신소자 및 이를 기반으로 하는 회로 및 시스템 기술을 개척함에 있어 한국에서 개발한 전자설계자동화 도구가 설계에 널리 활용되길 바라며 이에 대한 연구와 지원이 확대되기를 기대한다.

참고 문헌

[1] Vial G 2019 Understanding digital transformation: A review and a research agenda J. Strateg. Inf. Syst. 28 118-44
 [2] Sze V, Chen Y-H, Yang T-J and Emer J 2017 Efficient Processing of Deep Neural Networks: A Tutorial and Survey
 [3] Silver D, Huang A, Maddison C J, Guez A, Sifre L, van den



- Driessche G, Schrittwieser J, Antonoglou I, Panneershelvam V, Lanctot M, Dieleman S, Grewe D, Nham J, Kalchbrenner N, Sutskever I, Lillicrap T, Leach M, Kavukcuoglu K, Graepel T and Hassabis D 2016 Mastering the game of Go with deep neural networks and tree search *Nature* 529 484–9
- [4] Tianqi T, Xia L, Li B, Luo R, Chen Y, Wang Y and Yang H 2015 Spiking Neural Network with RRAM: Can We Use It for Real-World Application? Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition (DATE) (Grenoble, France: IEEE) pp 860–5
- [5] Goldenblat G 2010 Compact Modeling (Dordrecht: Springer Netherlands)
- [6] Mijalković S 2022 A Practical Guide to Verilog-A (Berkeley, CA: Apress)
- [7] Dutton R W 1986 Modeling and Simulation for VLSI Technical Digest – International Electron Devices Meeting pp 1–7
- [8] Kit Man Cham K M 1985 Computer-aided design and VLSI device development / Kit Man Kit Man Cham.
- [9] Park J, Choi J, Kim G, Kim G, Kim G S, Song H, Kim Y S, Lee Y, Rhee H, Lee H M, Hwang C S and Kim K M 2022 Modified Dynamic Physical Model of Valence Change Mechanism Memristors,” *ACS Appl. Mater. Interfaces* 14(31) 35949–35958.
- [10] Urquiza M L, Islam M M, van Duin A C T, Cartoixa X and Strachan A 2021 Atomistic Insights on the Full Operation Cycle of a HfO₂-Based Resistive Random Access Memory Cell from Molecular Dynamics *ACS Nano* 15 12945–54
- [11] Maldonado D, Aldana S, González M B, Jiménez-Molinos F, Ibáñez M J, Barrera D, Campabadal F and Roldán J B 2022 Variability estimation in resistive switching devices, a numerical and kinetic Monte Carlo perspective *Microelectron. Eng.* 257 111736
- [12] Kim S, Kim S J, Kim K M, Lee S R, Chang M, Cho E, Kim Y B, Kim C J, In Chung U and Yoo I K 2013 Physical electro-thermal model of resistive switching in bi-layered resistance-change memory *Sci. Rep.* 3
- [13] Kim S, Choi S and Lu W 2014 Comprehensive physical model of dynamic resistive switching in an oxide memristor *ACS Nano* 8 2369–76
- [14] Goes W, Green D, Blaise P, Piccolboni G, Bricalli A, Regev A, Molas G and Nodin J F 2021 A comprehensive oxide-based ReRAM TCAD model with experimental verification 2021 IEEE International Memory Workshop, IMW 2021 – Proceedings (Institute of Electrical and Electronics Engineers Inc.)
- [15] 2014 Atlas User’s Manual (Silvaco)
- [16] Jiang Z, Wu Y, Yu S, Yang L, Song K, Karim Z and Wong H-S P 2016 A Compact Model for Metal-Oxide Resistive Random Access Memory With Experiment Verification *IEEE Trans. Electron Devices* 63 1884–92
- [17] Aguirre F L, Pazos S M, Palumbo F, Suñé J and Miranda E 2021 SPICE Simulation of RRAM-Based Cross-Point Arrays Using the Dynamic Memdiode Model *Front. Phys.* 9 1–19
- [18] Wang Y, Tang T, Xia L, Li B, Gu P, Li H, Xie Y and Yang H 2015 Energy efficient RRAM spiking neural network for real time classification *Proc. ACM Gt. Lakes Symp. VLSI, GLSVLSI* 20–22-May-2015 189–94



정성엽

- 2011년 2월 경희대학교 이과대학 정보디스플레이학
학사
- 2014년 2월 경희대학교 이과대학 정보디스플레이학
석사
- 2014년 6월 에콜 폴리테크닉 혁신기술 공학석사
- 2017년 10월 에콜 폴리테크닉 물리학 박사
- 2015년 11월 ~ 2016년 2월 실바코 유럽 마리퀴리
펠로우
- 2016년 6월 ~ 2016년 8월 컬럼비아대학교 알리앙스
펠로우
- 2016년 12월 ~ 2019년 11월 포항공과대학교
박사후연구원
- 2019년 12월 ~ 2020년 7월 포항공과대학교 연구조교수
- 2020년 8월 ~ 현재 차세대융합기술연구원 선임연구원
- 2021년 8월 ~ 현재 차세대융합기술연구원 반도체 소자
및 회로 연구실장

〈관심 분야〉

전자설계자동화, 압축 모델, SPICE, 반도체 소자 및 회로

Kubernetes 기반 오픈소스 AI/ML 플랫폼

I. 서론

실제 상용 환경에서의 AI/ML Application Lifecycle 은 <그림 1> 과 같이 4단계로 구성된다. 먼저 Data 단계에서는 학습을 위한 data 를 받아 오고 (ingestion), 결측치, 오류 등을 시정하고 (cleansing), 분석하고 (analysis), 필요에 따라 변환이나 복합 필드를 만들고 (transformation), 오류 없는지 최종 확인 (validation) 하게 된다.

그 다음 Training 단계에서는 받아온 data 를 분리하고 (splitting), 유의미한 field 들을 선택하고 (feature engineering), ML model을 설계하고 (model development), 기 확보된 data로 ML model 을 학습 시키고, model 크기를 최소화시키고 (optimization), test data로 학습된 model의 유효성을 확인 (validation) 하게 된다.

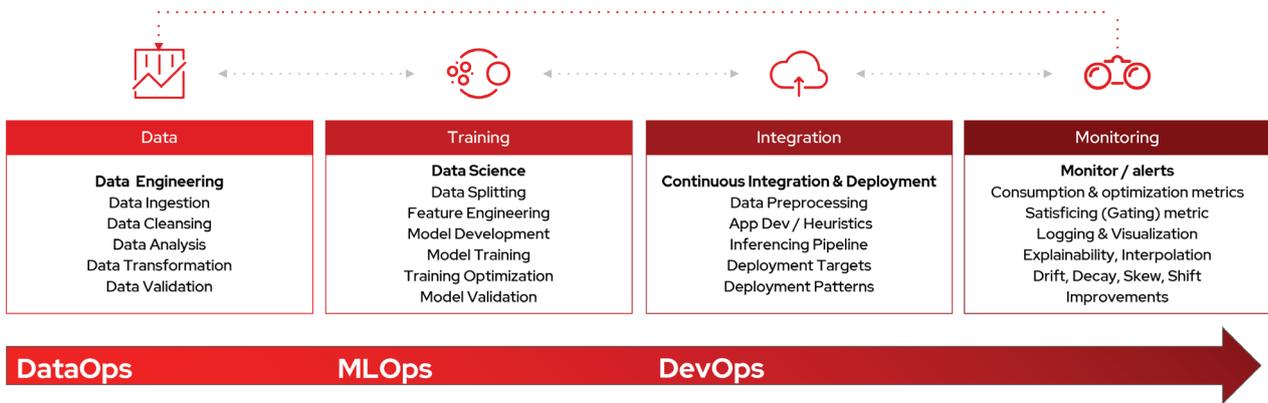
그 다음 Integration 단계에서는 data 수집 및 처리 프로세스를 구축하고 (data preprocessing), frontend, backend app 개발을 하고 (App Dev/ Heuristics), inference pipeline 을 구축하고, AI/ML app 을 배포할 장소 (deployment target) 선정 후 설계된 canary, rolling, blue/green 등 배포 pattern 에 따라 배포하게 된다.

다음 Monitoring 단계에서는 AI/ML model 이 얼마나 business impact 가 있는지 (consumption metric), 성능은 어떤지 (optimization metric), 목표를 달성하는지 (satisfying metric) 를 지속적으로 모니터링하고, 주요 이벤트를 수집하고 시각화하며 (logging & visualization), 결과를 설명 (explain) 또는 해석 (interpretation) 하며, 운용 환경 변화 감지 시 (Drift, Decay, Skew, Shift) 재학습 또는 model 변경하며 전체 추론 결과를 개선하게 (improvements) 된다.

이러한 AI/ML 프로젝트를 실제로 진행할 때 일반적으로 아래와 같은 어려움에 봉착하게 된다.



박용석
Red Hat



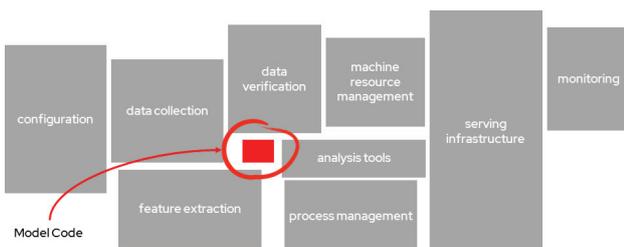
〈그림 1〉 AI/ML Lifecycle

- 전문 인력 부재: 프로젝트 수행에 필요한 주요 기술 보유 인력을 찾기도 유지하기도 어려움.
- AI/ML 도구 및 인프라를 쉽게 사용할 수 있게 해 주는 셀프 서비스 환경 부족: 데이터 과학자들과 개발자들의 업무 수행속도를 느리게 만듦
- AI/ML 프로젝트 운용 복잡성: 수동으로 진행되고 부서간 소통부재로 인해 느려진 운용에 따라 AI/ML lifecycle 전체가 느려짐

이런 운용상의 어려움들이 프로젝트 성공을 좌우하는 이유는 [1]에서 지적한 바와 같이 AI/ML model을 개발하고 coding 하는 부분이 전체 프로젝트의 극히 일부분에 불과하기 때문이다 〈그림 2〉. 잘 설계되지 않은 대부분의 프로젝트들이 실패하게 되는 이유이다.

이러한 AI/ML lifecycle 의 효율적인 운용을 지원하기 위한 체계적인 기능을 갖춘 시스템 (MLOps 로 불려지고 있음^[2]) 들이 최근 많이 등장하고 있다.

대표적인 시스템들로 public cloud 쪽에는 AWS



〈그림 2〉 AI/ML 프로젝트 구성 요소^[1]

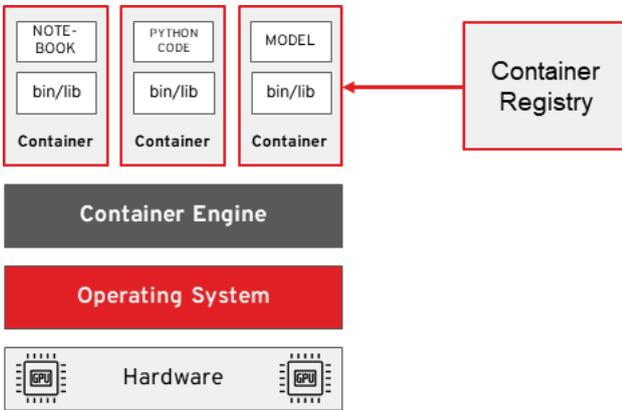
Sagemaker^[3], Azure Machine Learning^[4], Google Vertex AI^[5], IBM Watson Studio^[6] 등이 있으며 on-premise 쪽에는 metaflow^[7], mlflow^[8], kubeflow^[9] 등이 있다.

AI/ML 프로젝트에서 이러한 MLOps Platform 을 활용하게 되면 아래와 같은 변화가 가능하게 된다.

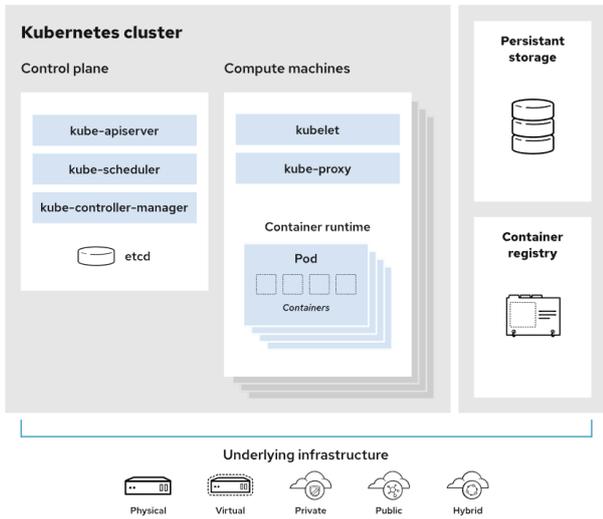
- 모델 학습 가속화
- 컴퓨팅 리소스 효율화
- 지속적으로 반복적인 결과
- 협력 용이
- 쉬운 업무 인계
- 쉬운 상용 제품화
- 쉬운 유지 보수
- 자동화

II. Container, Kubernetes, OpenShift

최근 Software 동향 중 가장 큰 변화로 Cloud-native 로의 전환을 들 수 있으며 이를 뒷받침하는 기술이 container 〈그림 3〉 와 Kubernetes^[12] 〈그림 4〉 라고 말할 수 있다. Container 는 OS 레벨 가상화 기술을 이용해 App binary, run-time framework, 필요 library, 설정 등을 하나의 package 에 담아 1) 어느 computing 환경에서도 Container framework 이 깔려 있는 한 동작할 수 있게 한 기술로 Docker^[11] 가 대표적이다.



〈그림 3〉 Container



〈그림 4〉 Kubernetes

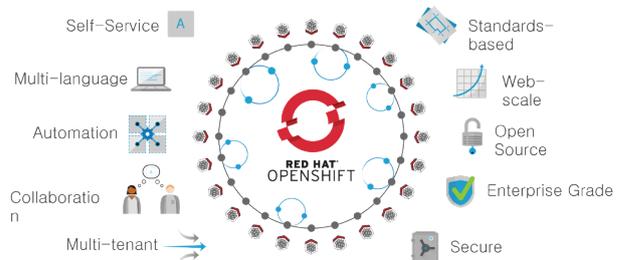
Container image는 Registry 에 저장시켜 내외부에서 해당 App 수행 시 download 해 설치하게 된다.

Kubernetes 는 여러 서버 노드로 구성된 Cluster를 만들고 Container를 적합한 노드에서 동작하게 하는 분산 Computing 플랫폼이다. 〈그림 4〉에 나와 있듯이 control plane 은 보통 3개 이상의 노드로 구성되어 있어 quorum 기반의 가용성을 보장하고 있으며 cluster lifecycle 전체에 대한 운용을 API 기반으로 제공한다. Etcd 는 cluster 상태, 설정, 인증 등 운용에 필요한 data 를 memory 상에 저장하고 있다. Pod 는 Kubernetes 에서 lifecycle 최소 단위로 하나 이상의 container 들로 구성된다. Control plane 에 속하지 않는 노드들은 worker node 로 불리우며 실제 application pod 들이 배

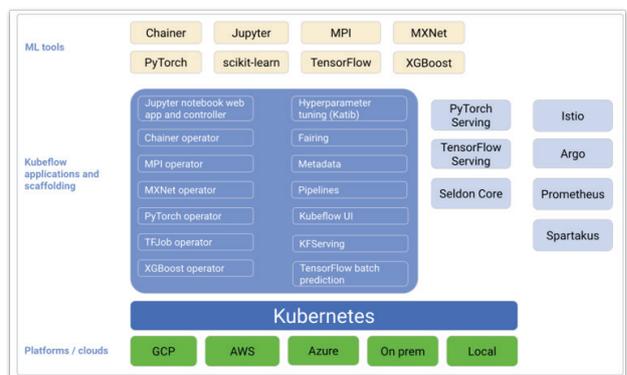
포되고 수행된다. Kubelet 은 node 에 대한 제어 및 모니터링을 수행하고 control plane 과 정보 공유 및 제어 명령 수행을 하는 agent 기능을 한다. Kubernetes 에서 기본으로 주어지는 가상 저장소는 휘발성 성격을 갖고 있어 pod 재기동 시 data 가 유실된다. 이를 방지하기 위해서는 persistent storage 를 활용한 persistent volume을 만들어 사용해야 한다. Container Registry 는 pod 에서 사용하는 container 들의 저장소 역할을 한다.

Kubernets 는 AWS, Azure, Google Cloud Platform, IBM Cloud 등 거의 모든 public cloud에서 제공하고 있고 on-premise 에서는 Openshift ^[13] (Community edition: OKD ^[14]), Rancher Prime ^[15] 등의 제품들이 있다.

특히 Openshift 는 Kubernetes 에 추가적으로 self-service console, web UI 기반 workflow, application build, metrics/monitoring, real-time/batch data streaming, zero-trust security, GPU support, public cloud infra support 등의 기능을 제공한다.



〈그림 5〉 Openshift



〈그림 6〉 Kubeflow

III. Kubeflow

최근 AI/ML Application 들도 Container/ Kubernetes 환경에서 시험되고 배포되고 있는 트렌드에 따라 MLOps 전체 단계도 Container/ Kubernetes 환경에서 수행하는 것이 가용성, 확장성, 이식성 뿐 아니라 플랫폼/인력 효율화 측면에서도 유리하다. 최근 Kubernetes 기반 Open-source MLOps framework 이 여럿 출현하고 있는데 대표적인 것이 kubeflow^[9] 이다.

Kubeflow 는 다양한 Kubernetes 플랫폼에 machine learning workflow를 간단하고 확장성 있게 제공하는 것을 목표로 하고 있으며 주요 기능은 아래와 같다.

- Interactive & customizable Jupyter notebooks
- Distributed Tensorflow model training
- Hyperparameter tuning: Katib
- Model serving: Tensorflow serving, Seldon, Nvidia Triton inference serving 등과 연동
- Kubeflow pipelines
- Multi-framework 연동: PyTorch, Apache MXNet, MPI, XGBoost, Chainer 등 ML framework; Elyra, Kale 등 workflow; Istio, Ambassador 등 Ingress framework; Nuclio 등 serverless framework

IV. Open Data Hub

Open Data Hub (ODH)^[10] 는 Kubeflow 의 개념 및 구성요소 기반으로 주요 Kubernetes 플랫폼인 OpenShift 대상으로 사용 편의성 제고 및 end-to-end ML workflow를 지원하는 Open-source MLOps 플랫폼이다. 주요 기능은 아래와 같다.

- Kubernetes Operator를 활용한 쉬운 one-stop 설치
- 모든 AI/ML 사용자 타입 및 각 단계 별 Tool 제공
- Data scientists 위한 AI/ML workflow 모니터링 제공
- Data Engineer 위한 ETL (Extract, Transform,

Load) tools 제공

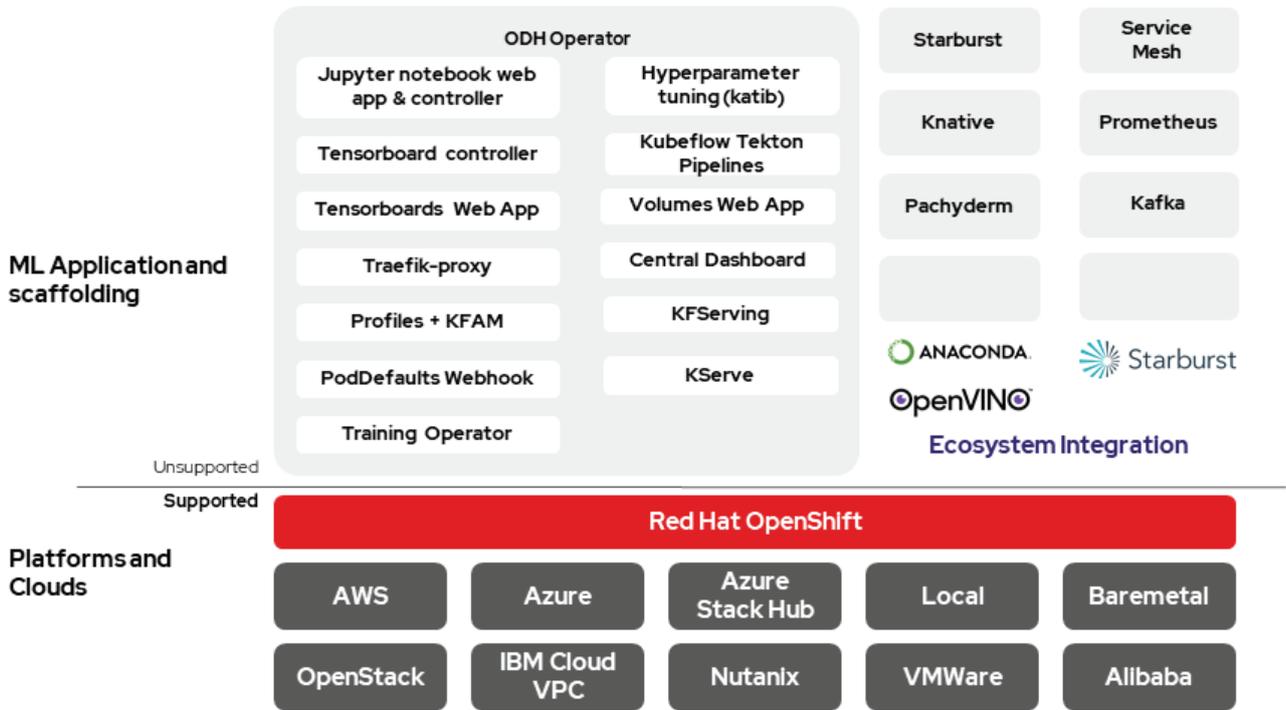
- AI/ML pipelines and long processing tasks 제공

Open Data Hub 의 구조는 <그림 7>에 나와 있으며 주요 구성 요소는 아래와 같다.

- Jupyter Notebook Controller: web UI 기반 interactive AI/ML 프로그래밍 환경을 복수의 사용자에게 제공
- Tensorboard: Tensorflow 기반 AI/ML model 의 성능 metric tracking 및 visualization 제공
- Kubeflow Access Management (KFAM): AI/ML pipeline 및 시험에 대해 multi-tenancy 격리 제공
- Traefik proxy: Secure access and load-balancing
- KFServing/KServe: AI/ML model serving
- Training Operator: Distributed training 제공
- Kubeflow Tekton Pipelines: Cloud-native workflow engine인 Tekton^[17] 기반 AI/ML workflow programming 및 수행환경 제공
- Central Dashboard: 설치된 Application 및 Tool 에 대한 shortcut, 설치 가능한 3rd party Tool list 등 제공
- Knative: Serverless engine for Kubernetes
- Pachyderm^[20]: Data Management & transformation system with governance
- Service Mesh: Microservice 들간 traffic 제어, 암호화, tracking 제공하며 Istio^[21] 가 대표적임.
- Prometheus^[22]: Pull 방식의 Metric monitoring 시스템으로 OpenShift 에도 내장되어 있음.
- Kafka^[23]: Distributed event streaming platform 으로 비주기적인 data transport에 많이 쓰임

이 외에도 아래와 같은 상용 Framework 도 연동되어 있고 리스트는 지속적으로 확대되고 있다..

- Starburst^[18]: data query engine (Open-source: Trino^[19], 과거 PrestoSQL)
- Anaconda^[24]: 다양한 AI/ML library 들을 포함한

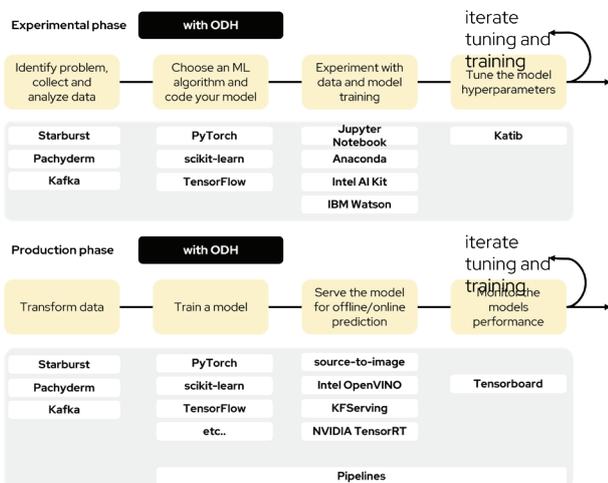


〈그림 7〉 Open Data Hub 구조

Python package

- OpenVINO ^[25]: Intel 에서 제공하는 AI/ML 성능 최적화 framework

AI/ML workflow 의 한 예는 〈그림 8〉에 나와 있다. Open Data Hub를 활용한 Experimental Phase 에서는



〈그림 8〉 ODH 활용 예

Starburst, Pachderm, Kafka 등을 활용 해 data를 수집, 가공하고, Pytorch, scikit-learn, Tensorflow 등을 활용해 Machine/Deep Learning model을 개발하고, Jupyter notebook 과 Katib 등을 이용해 train 및 hyperparameter tuning을 시킬 수 있다.

Production Phase 에서도 같은 tool 들을 활용해 data 수집 하고, model을 train 하고, 성능 확인 후 Application 과 연동시켜 배포 할 수 있다. 이후, 주기적 또는 성능 저하 이벤트에 따른 모델 re-training 이 지속적으로 발생하게 할 수 있다.

Open Data Hub을 활용한 Managed Service 인 Red Hat Openshift Data Science ^[16] 가 AWS 등 public cloud 에서 현재 제공되고 있으며 ODH 의 Enterprise 제품은 2023년 출시 예정임.

V. 전망과 결론

기기 및 서비스 측면에서 스마트폰 및 IoT 발전, 메타버스 및 스트리밍 서비스 확산되고 있고, 인프라 측면



에서 Container/Kubernetes 기반으로 public cloud, Multi-Access Edge Computing 기술이 확산됨에 따라 AI/ML 개발 및 운용 자동화 플랫폼도 이를 대응하고 수용하기 위해 계속 다양한 방향으로 진화할 것으로 예측된다. 이러한 변화를 선도하기 위해 다양한 open-source project 들이 생겨나고 있고 이를 수용하기 위한 Kubeflow, Open Data Hub 등 Open-source AI/MLOps platform의 활용도가 더욱 늘어날 것으로 예상된다.

참고 문헌

- [1] Sculley et al., "Hidden Technical Debt in Machine Learning Systems," NIPS 2015.
- [2] <https://en.wikipedia.org/wiki/MLOps>
- [3] <https://aws.amazon.com/sagemaker/>
- [4] <https://azure.microsoft.com/en-us/products/machine-learning/>
- [5] <https://cloud.google.com/vertex-ai-workbench>
- [6] <https://www.ibm.com/cloud/watson-studio>
- [7] <https://metallflow.org>
- [8] <https://mllflow.org>
- [9] <https://kubeflow.org>
- [10] <https://opendatahub.io>
- [11] <https://docker.com>
- [12] <https://kubernetes.io>
- [13] <https://www.redhat.com/en/technologies/cloud-computing/openshift/container-platform>
- [14] <https://okd.io>
- [15] <https://www.rancher.com/products/rancher>
- [16] <https://www.redhat.com/en/technologies/cloud-computing/openshift/openshift-data-science>
- [17] <https://tekton.dev>
- [18] <https://starburst.io>
- [19] <https://trino.io>
- [20] <https://pachyderm.com>
- [21] <https://istio.io>
- [22] <https://prometheus.io>
- [23] <https://kafka.apache.org>
- [24] <https://anaconda.com>
- [25] <https://www.intel.com/content/www/us/en/developer/tools/opencvino-toolkit/overview.html>



박용석

- 1986년 2월 서울대학교 전자공학과 학사
- 1988년 2월 서울대학교 전자공학과 석사
- 1996년 5월 퍼듀대학교 전기공학과 박사
- 1996년 5월 ~ 1996년 8월 IBM Austin Research Lab, Visiting Scientist
- 1996년 8월 ~ 2000년 10월 AT&T, Senior Member of Technical Staff
- 2002년 2월 ~ 2020년 7월 삼성전자 네트워크 사업부, 수석연구원
- 2020년 9월 ~ 2021년 7월 바이랩, 개발팀장
- 2021년 7월 ~ 현재 Red Hat, Principal Solution Architect

〈관심 분야〉

Cloud Platform, Data Analytics, MLOps, AIOps

통신 분야

[통신]

- 가상 배열 기반 광대역 DOA 추정
신익호, 정용식, 김신교, 박철순, 오정석
- 누수 잡음 크기 스펙트럼을 이용한 SVM 기반의 상수관로 누수 감지 및 분류
최준규, 임성빈

[ITS]

- 트램용 비접촉 과금을 위한 BLE 기반 승객 승하차 판단 시스템의 개발 및 평가
백지현, 곽재호, 황현철, 전희균

반도체 분야

[반도체 소자 및 재료]

- 인공신경망을 이용한 실리콘 태양전지 I-V 특성 예측
김경민, 김성겸, 이정은, 이종환
- RF 인덕티브 효과를 가진 Body-Contact PD-SOI MOSFET의 개선된 소신호 등가회로 파라미터 추출 방법
김경준, 이성현

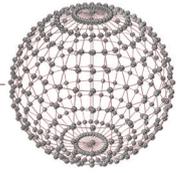
컴퓨터 분야

[M2M/IoT]

- Multi-to-One 가시광 통신을 이용한 스마트 IoT 모니터링 시스템 설계 및 구현
고나영, 최다봄, 심규성, 안병구

[인공지능 및 보안]

- 3D 모델 기반 합성 이미지 생성을 이용한 장애인 교통약자 탐지 성능 개선
이재용, 김학일



인공지능 신호처리 분야

[영상 신호처리]

- 주파수 도메인 특성과 딥러닝을 이용한 백내장 분류
신민재, 김대원
- 완전 복소 홀로그램 압축 성능 분석
서주연, 고현석

[음향 및 신호처리]

- CNN과 MKDE 모델의 직렬 연결을 기반한 합성 음성 분류
심영준, 최준규, 임성빈

산업전자 분야

[반도체재료 및 VLSI 설계]

- 피드백 구조를 이용한 외부 요인 변화에 더욱 둔감해진 기준 전원 회로의 설계
손종현, 윤홍일

[신호처리 및 시스템]

- 자율 주행 중 센서 전송처리를 위한 KC-CAR 알고리즘 제안
조도현
- 딥러닝 알고리즘을 이용한 태양광발전량 예측 모델에 대한 연구
강병복, 윤중현

이상학
LEE, SANG HAK
티맥스소프트

학위논문 제목	국문 : 공막혈관인식: 지역 구면 구조체에 기반한 새로운 매칭 기법 영문 : Sclera Recognition: Novel Matching Technique based on Local Spherical Structure
학위취득	연세대학교
취득년월	2022년 2월
지도교수	Andrew Beng Jin Teoh
KEY WORD	생체 인식, 공막혈관인식, 관심점, 혈관 모양 특징, 지역 구면 구조체, 지역 묘사자

〈논문 요약〉

COVID-19와 같은 접촉성 감염이 두드러지는 질환의 창궐은 지문 인식과 같은 기존의 접촉식 생체 인식 형태가 지닌 문제가 재고되는 한편 비접촉식 생체 인식, 즉 원격으로 획득 가능한 생체 정보에 대한 인식을 탐구하도록 촉구하고 있다.

본 논문은 인간의 눈에서 각막을 둘러싸고 있는 불투명한 흰색 영역인 공막(sclera)을 기반으로 하는 비접촉식 생체 인식 시스템을 개선하는 연구이다. 일반적으로 공막의 혈관은 조밀하게 형성된 네트워크로 구성되어 있으며 이 구조는 무작위로 구성되어 인간을 식별할 수 있을 만큼 고유한 특징을 지닌다. 기존의 연구들은 식별력이 높은 특징을 얻고자 기하학적 접근을 위한 관심점(interest point), 묘사자 생성을 위한 질감(texture), 또는 그 양쪽의 조합에 주로 의존한다. 하지만 3차원 구 형태에 가까운 인간의 안구에서 기인하는 비선형 변형(nonlinear deformation) 문제를 해결하기에는 충분하지 않다. 따라서 본 논문에서는 이러한 문제를 해결하고자 공막 혈관에 대한 2차원 영상으로부터 추출한 혈관 모양 특징(vessel shape feature)을 기반으로 하는 묘사자와 구 형태의 안구 모양을 고려한 지역 구면 구조체(local spherical structure)를 제안한다.

전처리 단계에서는 홍채 영역 분할(segmentation), 공막 영역 분할, 혈관 패턴 강화 순서대로 수행되며 특히 강화 단계에서는 형태학적(morphological) 연산을 통해 혈관의 두께 변화를 완화시킨다. 그리고 해리스 코너 검출기를 이용하여 풍부한 관심점들을 추출한 후, 각 관심점 주변의 작은 영역을 극(polar) 히스토그램 특징으로 부호화 하여 혈관 패턴 모양을 정의한다. 또한 각 관심점은 주변 관심점들과의 위상 관계(topological relation) 정보를 안구의 모양을 기반으로 정의한다. 정합(matching) 단계에서는 앞서 정의된 특징들이 검출된 관심점 개수에 무관하게 수행되도록 점수화 한 알고리즘을 적용한다. 본 논문에서 제안된 방법을 검증하기 위해 UBIRIS.v1, UTIRIS, SBVPI 데이터베이스를 활용하여 광범위한 실험을 수행하였으며 성능 평가를 위한 실험은 EER(equal error rate)을 기준으로 수행되었다.

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: biz@theieie.org

>>2023년 3월

일자	학술대회명	개최장소	홈페이지/연락처
03.01 - 03.03	2023 International Conference on Emerging Smart Computing and Informatics (ESCI)	Pune, India	https://esciioit.org/
03.02 - 03.03	2023 IEEE Power and Energy Conference at Illinois (PECI)	Champaign, Illinois, USA	https://www.peci.ece.illinois.edu/
03.02 - 03.04	2023 Second International Conference on Electronics and Renewable Systems (ICEARS)	Tuticorin, India	http://icears.com/2023/
03.03 - 03.05	2023 15th International Conference on Computer and Automation Engineering (ICCAE)	Sydney, Australia	http://www.iccae.org/
03.03 - 03.04	2023 6th International Conference on Information Systems and Computer Networks (ISCON)	Mathura, India	https://www.gla.ac.in/iscon2023/index.html
03.03 - 03.04	2023 19th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)	Kedah, Malaysia	https://sites.google.com/view/asprg-cspa/home
03.03 - 03.05	2023 International Conference on Robotics and Automation in Industry (ICRAI)	Peshawar, Pakistan	https://icrai2023.org/
03.03 - 03.05	2023 2nd International Conference for Innovation in Technology (INOCON)	Bangalore, India	http://inoconf.org/
03.04 - 03.11	2023 IEEE Aerospace Conference	Big Sky, Montana, USA	https://www.aeroconf.org/
03.05 - 03.09	2023 Optical Fiber Communications Conference and Exhibition (OFC)	San Diego, California, USA	https://www.ofcconference.org/en-us/home/
03.05 - 03.07	2023 IEEE 3rd International Symposium on Joint Communications & Sensing (JC&S)	Seefeld, Austria	https://jcs-symposium.org/
03.06 - 03.09	2023 IEEE Underwater Technology (UT)	Tokyo, Japan	http://www.ut2021.org/
03.06 - 03.07	2023 International Conference on Advances in Electronics, Control and Communication Systems (ICAECCS)	BLIDA, Algeria	https://www.univ-blida.dz/icaeccs23/
03.06 - 03.09	2023 26th Conference on Innovation in Clouds, Internet and Networks and Workshops (ICIN)	Paris, France	https://www.icin-conference.org/
03.08 - 03.09	2023 Conference on Information Communications Technology and Society (ICTAS)	Event Format: Virtual	http://www.ictas.org/
03.09 - 03.11	2023 SICE International Symposium on Control Systems (SICE ISCS)	Kusatsu, Japan	http://iscs2023.sice-ctrl.jp/
03.09 - 03.10	2023 Argentine Conference on Electronics (CAE)	Cordoba, Argentina	http://eamta.ar/
03.10 - 03.12	2023 7th International Conference on Green Energy and Applications (ICGEA)	Singapore, Singapore	http://www.icgea.org/
03.11	2023 IEEE Integrated STEM Education Conference (ISEC)	Laurel, Maryland, USA	https://ewh.ieee.org/conf/stem/index.html
03.11 - 03.12	2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT)	Male, Maldives	http://globconht.org/
03.12 - 03.15	2023 IEEE PES Conference on Innovative Smart Grid Technologies - Middle East (ISGT Middle East)	Abu Dhabi, United Arab Emirates	http://ieee-isgt-me.org/

일자	학술대회명	개최장소	홈페이지/연락처
03.13 - 03.17	2023 IEEE IAS Electrical Safety Workshop (ESW)	Reno, Nevada, USA	https://electricalsafetyworkshop.com/
03.13 - 03.17	2023 IEEE 20th International Conference on Software Architecture Companion (ICSA-C)	L'Aquila, Italy	https://icsa-conferences.org/2023/
03.13 - 03.16	2023 18th ACM/IEEE International Conference on Human-Robot Interaction (HRI)	Stockholm, Sweden	https://humanrobotinteraction.org/2023/
03.14 - 03.16	2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)	Uttarakhand, India	http://icidca.com/2023/
03.14 - 03.16	2023 Systems of Signals Generating and Processing in the Field of on Board Communications	Moscow, Russia	http://media-publisher.ru/en/2023-on-board/
03.15 - 03.17	2023 22nd International Symposium INFOTEH-JAHORINA (INFOTEH)	East Sarajevo, Bosnia and Herzegovina	https://infoteh.etf.ues.rs.ba/indexe.php
03.15 - 03.17	2023 10th International Conference on Computing for Sustainable Global Development (INDIACom)	New Delhi, India	http://bvicam.ac.in/indiacom/
03.15 - 03.16	2023 International Conference on Communication Technologies (ComTech)	Rawalpindi, Pakistan	http://conferences.mcs.nust.edu.pk/comtech2023/cfp.html
03.15 - 03.16	2023 6th International Conference on Energy Conservation and Efficiency (ICECE)	Lahore, Pakistan	https://icece.kics.edu.pk/2023/
03.15 - 03.17	2023 IEEE International Conference on Mechatronics (ICM)	Loughborough, United Kingdom	https://www.imagovenues.co.uk/icm2023/
03.15 - 03.16	2023 International Conference on IT Innovation and Knowledge Discovery (ITIKD)	Manama, Bahrain	https://tikd.ahlia.edu.bh/
03.15 - 03.17	2023 Winter Summit on Smart Computing and Networks (WISSCoN)	Chennai, India	http://www.wisscon.in/
03.15 - 03.17	2023 IEEE 15th International Symposium on Autonomous Decentralized System (ISADS)	Mexico City, Mexico	https://www.isads2023.org/
03.16 - 03.18	2023 IEEE 12th International Conference on Educational and Information Technology (ICEIT)	Chongqing, China	http://www.iceit.org/
03.16 - 03.18	2023 5th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)	Moscow, Russia	http://reepe.mpei.ru/IEEE/Pages/default.aspx
03.17 - 03.18	2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)	Coimbatore, India	https://icaccs.sece.ac.in/
03.17 - 03.18	2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT)	Dehradun, India	https://dicct.geu.ac.in/
03.17 - 03.18	2022 International Conference on Digital Management, Information Systems and Technologies (DMIST)	Shenyang, China	http://www.icdmist.com/
03.18 - 03.20	2023 3rd International conference on Artificial Intelligence and Signal Processing (AISP)	VIJAYAWADA, India	https://www.aisp.in/
03.18 - 03.20	2023 11th International Conference on Information and Education Technology (ICIET)	Fujisawa, Japan	http://www.iciet.org/
03.19 - 03.23	2023 IEEE Applied Power Electronics Conference and Exposition (APEC)	Orlando, Florida, USA	http://www.apec-conf.org/
03.20 - 03.22	2023 6th Conference on Cloud and Internet of Things (CIoT)	Lisbon, Portugal	https://ciot.dnac.org/
03.20 - 03.22	2023 International Conference on Information Technology, Applied Mathematics and Statistics (ICITAMS)	Al-Qadisiya, Iraq	https://qu.edu.iq/ICITAMS/
03.20 - 03.24	2023 IEEE 20th International Conference on Software Architecture (ICSA)	L'Aquila, Italy	https://icsa-conferences.org/2023/
03.21 - 03.23	2023 IEEE International Symposium on Power Line Communications and its Applications (ISPLC)	Manchester, United Kingdom	https://isplc2023.ieee-isplc.org/
03.21 - 03.24	2023 Data Compression Conference (DCC)	Snowbird, Utah, USA	https://www.cs.brandeis.edu/~dcc/
03.21 - 03.24	2023 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)	Taipa, Macao	https://saner2023.must.edu.mo/
03.22 - 03.24	2023 57th Annual Conference on Information Sciences and Systems (CISS)	Baltimore, Maryland, USA	https://ciis.jhu.edu/

일자	학술대회명	개최장소	홈페이지/연락처
03.23 - 03.26	2023 Young Professionals Virtual Control Conference (YPVCC)	Event Format: Virtual	http://ypvcc2023.ieeecss.org/
03.23 - 03.25	2023 13th International Symposium on Advanced Topics in Electrical Engineering (ATEE)	Bucharest, Romania	http://www.atee.pub.ro/atee2023/
03.23 - 03.25	2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)	Erode, India	http://icscnds.com/2023/
03.23 - 03.26	2023 5th Asia Energy and Electrical Engineering Symposium (AEEES)	Chengdu, China	http://www.aeees.org/
03.23 - 03.24	2023 10th International Conference on Signal Processing and Integrated Networks (SPIN)	Noida, India	https://www.amity.edu/spin2023/
03.24 - 03.25	2023 IEEE Bangalore Humanitarian Technology Conference (B-HTC)	Mysore, India	https://www.bhtc-2023.ieeebangalore.org/
03.24 - 03.25	2023 Second International Conference On Smart Technologies For Smart Nation (SmartTechCon)	Singapore, Singapore	http://smarttech-conference.org/2023/
03.24 - 03.25	2023 Somaiya International Conference on Technology and Information Management (SICTIM)	Mumbai, India	https://simsr.somaiya.edu/en/conference/SICTIM
03.24 - 03.26	2023 3rd International Conference on Computer, Control and Robotics (ICCCR)	Shanghai, China	http://www.icccr.org/
03.26 - 03.29	2023 IEEE Wireless Communications and Networking Conference (WCNC)	Glasgow, United Kingdom	https://wcnc2023.ieee-wcnc.org/
03.26 - 03.31	2023 17th European Conference on Antennas and Propagation (EuCAP)	Florence, Italy	https://www.eucap2023.org/
03.27 - 03.31	2023 International Russian Smart Industry Conference (SmartIndustryCon)	Sochi, Russia	https://smartindustrycon.ru/index-eng.html
03.27 - 03.30	2023 35th International Conference on Microelectronic Test Structure (ICMETS)	Tokyo, Japan	http://icmts.if.t.u-tokyo.ac.jp/icmts.if.t.u-tokyo.ac.jp_8080/index.html
03.29 - 03.31	2023 IEEE International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles & International Transportation Electrification Conference (ESARS-ITEC)	Venice, Italy	http://www.esars-itec.eu/
03.29 - 03.31	2023 25th International Conference on Digital Signal Processing and its Applications (DSPA)	Moscow, Russia	http://dspa-conf.org/pages/home

>>2023년 4월

04.04 - 04.07	2023 IEEE Silicon Photonics Conference (SiPhotonics)	Washington, District of Columbia, USA	https://ieee-gfp.org/
04.05 - 04.07	2023 Second International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)	Trichirappalli, India	http://iceict.in/
04.05 - 04.07	2023 24th International Symposium on Quality Electronic Design (ISQED)	San Francisco, California, USA	https://www.isqed.org/
04.07 - 04.09	2023 IEEE 8th International Conference for Convergence in Technology (I2CT)	Lonavla, India	https://ieeepune.i2ct.in/
04.08 - 04.09	2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT)	Bhopal, India	http://csnt.in/
04.11 - 04.13	2023 7th International Conference on Trends in Electronics and Informatics (ICOEI)	Tirunelveli, India	http://icoei.com/2023/
04.13 - 04.16	SoutheastCon 2023	Orlando, Florida, USA	https://ieeesoutheastcon.org/
04.14 - 04.16	2023 8th Asia Conference on Power and Electrical Engineering (ACPEE)	Tianjin, China	https://acpee.net/
04.14 - 04.16	2023 8th International Conference on Control and Robotics Engineering (ICCRE)	Niigata, Japan	http://www.iccre.org/
04.14 - 04.16	2023 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC)	Dalian, China	http://www.acipec.org/
04.14 - 04.16	2023 IEEE 3rd International Conference on Electronic Communications, Internet of Things and Big Data (ICEIB)	Taichung, Taiwan	http://www.iceib.asia/

일자	학술대회명	개최장소	홈페이지/연락처
04.16 - 04.20	2023 IEEE Conference on Software Testing, Verification and Validation (ICST)	Dublin, Ireland	https://conf.researchr.org/home/icst-2023
04.17 - 04.20	2023 IEEE International Systems Conference (SysCon)	Vancouver, British Columbia, Canada	https://ieeesystemscouncil.org/
04.17 - 04.20	2023 International VLSI Symposium on Technology, Systems and Applications (VLSI-TSA/VLSI-DAT)	HsinChu, Taiwan	https://expo.itri.org.tw/2023VLSITSADAT/
04.17 - 04.21	2023 IX International Conference on Information Technology and Nanotechnology (ITNT)	Event Format: Virtual	http://itnt-conf.org/index.php/en/
04.17 - 04.18	2023 IEEE Wireless and Microwave Technology Conference (WAMICON)	Melbourne, Florida, USA	https://www.ieewamicon.org/
04.17 - 04.19	2023 Design, Automation & Test in Europe Conference & Exhibition (DATE)	Antwerp, Belgium	https://www.date-conference.com/
04.18 - 04.20	2023 INSUCON - 14th International Electrical Insulation Conference (INSUCON)	Birmingham, United Kingdom	https://insucon.org/
04.18 - 04.21	2023 IEEE 16th Pacific Visualization Symposium (PacificVis)	Seoul, Korea (South)	https://pvis2023.github.io/pvis2023/
04.18 - 04.20	2023 Integrated Communication, Navigation and Surveillance Conference (ICNS)	Herndon, Virginia, USA	https://i-cns.org/
04.19 - 04.21	2023 Wireless Telecommunications Symposium (WTS)	Boston, Massachusetts, USA	https://www.cpp.edu/~wts/
04.19 - 04.21	2023 International Symposium on Medical Robotics (ISMR)	Atlanta, Georgia, USA	https://ismr.gatech.edu/
04.19 - 04.21	2023 IEEE Symposium in Low-Power and High-Speed Chips (COOL CHIPS)	Tokyo, Japan	https://www.coolchips.org/2023/
04.19 - 04.20	2023 33rd International Conference Radioelektronika (RADIOELEKTRONIKA)	Pardubice, Czech Republic	https://www.marew.cz/
04.19 - 04.21	2023 International Conference on Power, Instrumentation, Control and Computing (PICC)	Thrissur, India	http://icetec.gectcr.ac.in/picc-2023/
04.19 - 04.22	2023 IEEE Conference on Technologies for Sustainability (SusTech)	Portland, Oregon, USA	https://ieee-sustech.org/
04.20 - 04.22	2023 IEEE European Technology and Engineering Management Summit (E-TEMS)	Kaunas, Lithuania	https://etems.digital/2023/
04.21 - 04.23	2023 IEEE Region 5 Annual Meeting (R5)	Denver, Colorado, USA	https://r5conferences.org/
04.21 - 04.22	2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1)	Bangalore, India	https://www.alliance.edu.in/icaia-2023/
04.21 - 04.23	2023 8th International Conference on Computer and Communication Systems (ICCCS)	Guangzhou, China	http://www.icccs.org/
04.21 - 04.25	2023 9th International Conference on Applied System Innovation (ICASI)	Chiba, Japan	https://2023.icasi-conf.net/
04.21 - 04.23	2023 11th International Conference on Bioinformatics and Computational Biology (ICBCB)	Hangzhou, China	http://www.icbc.org/index.html
04.21 - 04.22	2023 IEEE International Conference on Contemporary Computing and Communications (InC4)	Bangalore, India	https://ic4.co.in/
04.22	2023 International Conference on Code Quality (ICCCQ)	St. Petersburg, Russia	https://www.icccq.ru/2023.html
04.23 - 04.28	2023 IEEE Information Theory Workshop (ITW)	Saint-Malo, France	https://itw2023.org/
04.23 - 04.27	2023 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA)	Dallas, Texas, USA	https://www.cementconference.org/event/9c9cbcec-fc6a-4fa8-a1ef-b224d1bc6743/summary
04.23 - 04.25	2023 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)	Raleigh, North Carolina, USA	https://ispass.org/ispass2023/
04.23 - 04.26	2023 IEEE Custom Integrated Circuits Conference (CICC)	San Antonio, Texas, USA	https://www.ieee-cicc.org/
04.24 - 04.27	2023 IEEE/ION Position, Location and Navigation Symposium (PLANS)	Monterey, California, USA	https://www.ion.org/plans/index.cfm

일자	학술대회명	개최장소	홈페이지/연락처
04.24 - 04.26	2023 IEEE International Workshop of Electronics, Control, Measurement, Signals and their application to Mechatronics (ECMSM)	Donosti/San Sebastian, Spain	https://www.ecmsm.eu/
04.25 - 04.28	2023 24th International Vacuum Electronics Conference (IVEC)	Chengdu, China	https://ivec2023.org/
04.25 - 04.27	2023 International Conference on Data Science, E-learning and Information Systems (Data)	Dubai, United Arab Emirates	https://iares.net/Conference/DATA2022
04.26 - 04.28	2023 8th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)	Chengdu, China	http://www.icccbda.com/
04.27 - 04.29	2023 1st International Conference on Cognitive Computing and Engineering Education (ICCCCE)	Pune, India	https://www.mitadtsocce-iccmr.org/
04.27 - 04.30	2023 Panda Forum on Power and Energy (PandaFPE)	Chengdu, China	https://www.pandafpe.org/
04.27 - 04.28	2023 IEEE Kansas Power and Energy Conference (KPEC)	Event Format: Virtual	https://www.kpec-ksu.org/
04.28 - 04.30	2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT)	Event Format: Virtual	http://www.iccect.com/
04.29 - 05.01	2023 IEEE International Conference on Advanced Systems and Emergent Technologies (IC_ASET)	Hammamet, Tunisia	https://attend.ieee.org/aset/
04.29 - 04.30	2023 International Conference on Mechanical, Automotive and Mechatronics Engineering (ICMAME)	Dubai, United Arab Emirates	http://icmame.org/
04.29 - 04.30	2023 International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE)	Ballar, India	https://icdcece.in/

>>2023년 5월

05.01 - 05.04	2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)	Dubai, United Arab Emirates	https://icbc2023.ieee-icbc.org/
05.01 - 05.05	2023 IEEE Radar Conference (RadarConf23)	San Antonio, Texas, USA	https://radar2023.ieee-radarconf.org/
05.01 - 05.03	2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON)	New Delhi, India	https://reedconjmi.com/
05.01 - 05.04	2023 34th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC)	Saratoga Springs, New York, USA	https://semi.org/en/connect/events/advanced-semiconductor-manufacturing-conference-asmc
05.01 - 05.04	2023 IEEE 7th International Conference on Fog and Edge Computing (ICFEC)	Bangalore, India	http://icfec2023.ontariotechu.ca/
05.01 - 05.04	2023 IEEE/ACM 23rd International Symposium on Cluster, Cloud and Internet Computing (CCGrid)	Bangalore, India	https://ccgrid2023.iisc.ac.in/
05.01 - 05.04	2023 IEEE Radio and Antenna Days of the Indian Ocean (RADIO)	Balaclava, Mauritius	http://www.radiosociety.org/radio2023/
05.01 - 05.04	2023 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)	San Jose, California, USA	http://www.hostsymposium.org/
05.03 - 05.05	2023 International Interdisciplinary PhD Workshop (IIPhDW)	Wismar, Germany	https://www.hs-wismar.de/iiphdw-2023
05.03 - 05.04	2023 9th International Conference on Web Research (ICWR)	Tehran, Iran	https://iranwebconf.ir/
05.03 - 05.05	2023 26th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)	Tallinn, Estonia	https://ddecs2023.taltech.ee/
05.04 - 05.06	2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAIC)	Salem, India	http://icaic.com/2023/
05.04 - 05.06	2023 IEEE International Conference on Smart Information Systems and Technologies (SIST)	Astana, Kazakhstan	https://sist.astanait.edu.kz/
05.05	2023 IEEE Vision, Innovation, and Challenges Summit and Honors Ceremony (VIC Summit)	Atlanta, Georgia, USA	https://corporate-awards.ieee.org/event/2023-vic-summit-honors-ceremony-gala/
05.05 - 05.06	2023 International Conference on Advancement in Computation & Computer Technologies (InCACCT)	Gharuan, India	https://icacct-cu.com/

일자	학술대회명	개최장소	홈페이지/연락처
05.05 - 05.06	2023 32nd Wireless and Optical Communications Conference (WOCC)	Newark, New Jersey, USA	http://www.wocc.org/
05.05	2023 IEEE Long Island Systems, Applications and Technology Conference (LISAT)	Old Westbury, New York, USA	https://www.ieee.li/lisat/
05.05 - 05.06	2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (VITECoN)	Vellore, India	https://vit.ac.in/VITECoN2023/
05.06	2023 IEEE Workshop on Industry and Innovations enabled by Industry 4.0 Technologies (USEI40)	San Jose, California, USA	http://usei40.org/
05.06 - 05.08	2023 1st International Conference on Renewable Solutions for Ecosystems: Towards a Sustainable Energy Transition (ICRSEtoSET)	Djelfa, Algeria	http://www.univ-djelfa.dz/icrsetoset/
05.06 - 05.08	2023 IEEE 9th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)	New York, New York, USA	http://www.cloud-conf.net/datasec/2023/index.html
05.07 - 05.10	2023 IEEE 27th Workshop on Signal and Power Integrity (SPI)	Aveiro, Portugal	https://spi2023.av.it.pt/
05.07 - 05.10	2023 IEEE 31st Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)	Los Angeles, California, USA	https://www.fccm.org/
05.07 - 05.10	2023 IEEE 4th International Conference on Electrical Materials and Power Equipment (ICEMPE)	Shanghai, China	http://www.icempe.org/
05.07 - 05.12	2023 Conference on Lasers and Electro-Optics (CLEO)	San Jose, California, USA	https://www.cleoconference.org/home/
05.08 - 05.12	NOMS 2023-2023 IEEE/IFIP Network Operations and Management Symposium	Miami, Florida, USA	https://noms2023.ieee-noms.org/
05.08 - 05.11	2023 International Conference on Optical Network Design and Modeling (ONDM)	Coimbra, Portugal	https://ondm2023.inescc.pt/
05.08 - 05.10	2023 IEEE 32nd Microelectronics Design & Test Symposium (MDTS)	Albany, New York, USA	https://mdts.ieee.org/
05.08 - 05.09	2023 8th International Engineering Conference on Renewable Energy & Sustainability (ieCRES)	Gaza, Palestine	http://engconf.iugaza.edu.ps/
05.09 - 05.12	2023 IEEE 29th Real-Time and Embedded Technology and Applications Symposium (RTAS)	San Antonio, Texas, USA	http://2023.rtas.org/
05.09 - 05.11	2023 4th International Conference on Artificial Intelligence, Robotics and Control (AIRC)	Cairo, Egypt	http://www.airc.org/
05.09 - 05.12	2023 20th International Conference on Electrical Engineering/ Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)	Nakhon Phanom, Thailand	https://ecticon2023.ecticon.org/
05.09 - 05.12	2023 IEEE/ACM Eighth International Conference on Internet-of-Things Design and Implementation (IoTDI)	San Antonio, Texas, USA	https://conferences.computer.org/iotDI/2023/
05.09	2023 11th Workshop on Modeling and Simulation of Cyber-Physical Energy Systems (MSCPES)	San Antonio, Texas, USA	https://palensky.org/mscpes/2023/
05.10 - 05.12	2023 International Conference on Control, Automation and Diagnosis (ICCAD)	Rome, Italy	https://www.iccad-conf.com/
05.10 - 05.14	2023 46th International Spring Seminar on Electronics Technology (ISSE)	Timisoara, Romania	https://isse-conf.eu/
05.10 - 05.11	2023 3rd International Conference on Computing and Information Technology (ICCIT)	Tabuk, Saudi Arabia	https://ut-iccit.org/ICCIT/
05.10 - 05.12	2023 IEEE International 3D Systems Integration Conference (3DIC)	Cork, Ireland	https://3dic-conf.org/
05.12 - 05.13	2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)	Greater Noida, India	http://icacite.com/index.php
05.12 - 05.14	2023 9th International Conference on Virtual Reality (ICVR)	Xianyang, China	http://www.icvr.org/
05.12 - 05.14	2023 IEEE 6th International Electrical and Energy Conference (CIEEC)	Hefei, China	https://www.cieec.com.cn/

일자	학술대회명	개최장소	홈페이지/연락처
05.12 - 05.14	2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS)	Xiangtan, China	https://ddcls23.hnust.edu.cn/index.html
05.12 - 05.13	2023 Intermountain Engineering, Technology and Computing (IETC)	Provo, Utah, USA	https://www.uvu.edu/cet/i-etc/
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)	Melbourne, Australia	http://www.icse-conferences.org/
05.14 - 05.17	2023 IEEE 18th International Conference on Nano/Micro Engineered and Molecular Systems (NEMS)	Jeju Island, Korea (South)	https://ieee-nems2023.org/
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)	Melbourne, Australia	https://conf.researchr.org/track/icse-2023/icse-2023-SEET
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)	Melbourne, Australia	https://conf.researchr.org/track/icse-2023/icse-2023-SEIS
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)	Melbourne, Australia	https://conf.researchr.org/track/icse-2023/icse-2023-technical-track
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)	Melbourne, Australia	https://conf.researchr.org/track/icse-2023/icse-2023-SEIP
05.14 - 05.20	2023 IEEE/ACM 45th International Conference on Software Engineering: New Ideas and Emerging Results (ICSE-NIER)	Melbourne, Australia	https://conf.researchr.org/track/icse-2023/icse-2023-NIER
05.15 - 05.18	2023 IEEE International Electric Machines & Drives Conference (IEMDC)	San Francisco, California, USA	https://www.iemdc.org/
05.15 - 05.19	2023 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)	Sochi, Russia	http://icieam.su-ieee.ru/
05.15 - 05.19	2023 IEEE International Parallel and Distributed Processing Symposium (IPDPS)	St. Petersburg, Florida, USA	http://www.ipdps.org/
05.15 - 05.17	2023 International Workshop on Antenna Technology (IWAT)	Aalborg, Denmark	http://iwat2023.org/
05.15 - 05.17	2023 IEEE Ural-Siberian Conference on Biomedical Engineering, Radioelectronics and Information Technology (USBEREIT)	Yekaterinburg, Russia	http://usbereit.ieeesiberia.org/
05.16 - 05.19	2023 IEEE MTT-S International Wireless Symposium (IWS)	Qingdao, China	http://www.em-conf.com/iws2023/
05.17 - 05.19	2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)	Madurai, India	http://iciccs.com/2023/
05.17 - 05.20	IEEE INFOCOM 2023 - IEEE Conference on Computer Communications	New York City, New York, USA	https://infocom2023.ieee-infocom.org/
05.17 - 05.19	2023 4th Information Communication Technologies Conference (ICTC)	Nanjing, China	http://www.ictc.net/index.html
05.17 - 05.19	2023 IEEE International Conference on Mobility, Operations, Services and Technologies (MOST)	Detroit, Michigan, USA	http://ieeemobility.org/
05.17 - 05.18	2023 International Conference on Communication, Computing and Digital Systems (C-CODE)	Islamabad, Pakistan	https://ccode.bahria.edu.pk/
05.17 - 05.19	2023 Joint Urban Remote Sensing Event (JURSE)	Heraklion, Greece	http://jurse2023.org/
05.17 - 05.19	2023 IEEE Renewable Energy and Sustainable E-Mobility Conference (RESEM)	Bhopal, India	https://conf.manit.ac.in/resem2023/important_dates.php
05.18 - 05.19	2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)	Mohammedia, Morocco	http://iraset.org/2023/
05.18 - 05.20	2023 IEEE International Conference on Electro Information Technology (eIT)	Romeoville, Illinois, USA	http://www.eit-conference.org/eit2023/
05.18 - 05.20	2023 3rd International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)	Event Format: Virtual	https://sites.google.com/adishankara.ac.in/access-23?pli=1
05.18 - 05.19	2023 8th International Conference on Business and Industrial Research (ICBIR)	Bangkok, Thailand	https://icbir.tni.ac.th/
05.19 - 05.21	2023 IEEE IAS Global Conference on Emerging Technologies (GlobConET)	Warsaw, Poland	https://www.globconet.org/

일자	학술대회명	개최장소	홈페이지/연락처
05.19 - 05.21	2023 International Conference on Control, Communication and Computing (ICCC)	Thiruvananthapuram, India	http://iccc2023.cet.ac.in/
05.19 - 05.21	2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)	Chengdu, China	http://www.icmsp.net/
05.20	IEEE INFOCOM 2023 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPs)	Hoboken, New Jersey, USA	https://infocom2023.ieee-infocom.org/
05.20 - 05.22	2023 35th Chinese Control and Decision Conference (CCDC)	Yichang, China	http://www.ccdc.neu.edu.cn/
05.21 - 05.24	2023 IEEE International Memory Workshop (IMW)	Monterey, California, USA	https://www.ewh.ieee.org/soc/eds/imw/
05.21 - 05.23	2023 IEEE 3rd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)	Benghazi, Libya	https://mista-con.org/
05.21 - 05.25	2023 IEEE/IAS 59th Industrial and Commercial Power Systems Technical Conference (I&CPS)	Las Vegas, Nevada, USA	https://site.ieee.org/ias-icps/2023-conference/
05.21 - 05.25	2023 IEEE International Conference on Plasma Science (ICOPS)	Santa Fe, New Mexico, USA	http://ece-events.unm.edu/icops2023/index.html
05.21 - 05.25	2023 IEEE Symposium on Security and Privacy (SP)	San Francisco, California, USA	https://www.ieee-security.org/TC/SP2023/
05.21 - 05.25	2023 IEEE International Symposium on Circuits and Systems (ISCAS)	Monterey, California, USA	https://iscas2023.org/
05.22 - 05.25	2023 IEEE PES GTD International Conference and Exposition (GTD)	Istanbul, Turkey	https://ieee-gtd.org/
05.22 - 05.26	2023 24th International Conference on the Computation of Electromagnetic Fields (COMPUMAG)	Kyoto, Japan	https://www.compumag2023.com/content/
05.22 - 05.26	2023 46th MIPRO ICT and Electronics Convention (MIPRO)	Opatija, Croatia	http://www.mipro.hr/
05.22 - 05.26	2023 URSI International Symposium on Electromagnetic Theory (EMTS)	Vancouver, British Columbia, Canada	https://www.emts2023.org/
05.22 - 05.25	2023 11th International Conference on Power Electronics and ECCE Asia (ICPE 2023 - ECCE Asia)	Jeju Island, Korea (South)	https://www.icpe-conf.org/
05.22 - 05.25	2023 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)	Kuala Lumpur, Malaysia	https://i2mtc2023.ieee-ims.org/
05.22 - 05.24	2023 IEEE 53rd International Symposium on Multiple-Valued Logic (ISMVL)	Matsue, Japan	https://mvl.jp.n.org/ISMVL2023/
05.22 - 05.25	2023 IEEE International Interconnect Technology Conference (IITC) and IEEE Materials for Advanced Metallization Conference (MAM)(IITC/MAM)	Dresden, Germany	https://iitc-conference.org/
05.22 - 05.26	2023 IEEE European Test Symposium (ETS)	Venezia, Italy	https://cas.polito.it/ETS23/
05.23 - 05.26	2023 IEEE 17th International Symposium on Applied Computational Intelligence and Informatics (SACI)	Timisoara, Romania	http://conf.uni-obuda.hu/saci2023/
05.23 - 05.25	2023 XIX International Scientific Technical Conference Alternating Current Electric Drives (ACED)	Event Format: Virtual	https://aced2023.ru/
05.24 - 05.26	2023 26th International Conference on Computer Supported Cooperative Work in Design (CSCWD)	Rio de Janeiro, Brazil	http://2023.cscwd.org/
05.24 - 05.26	2023 24th International Radar Symposium (IRS)	Berlin, Germany	https://www.dgon-irs.org/home/
05.24 - 05.26	2023 23rd International Scientific Conference on Electric Power Engineering (EPE)	Brno, Czech Republic	http://www.epe-conference.eu/
05.24 - 05.26	2023 XXVI International Conference on Soft Computing and Measurements (SCM)	Saint Petersburg, Russia	https://etu.ru/en/university/conferences/scm2023
05.25 - 05.26	2023 Smart City Symposium Prague (SCSP)	Prague, Czech Republic	https://akce.fd.cvut.cz/en/scsp2023
05.26 - 05.29	2023 6th International Conference on Artificial Intelligence and Big Data (ICAIBD)	Chengdu, China	http://www.icaibd.org/

일자	학술대회명	개최장소	홈페이지/연락처
05.26 - 05.28	2023 2nd International Conference on Innovations and Development of Information Technologies and Robotics (IDITR)	Chengdu, China	https://iditr.org/
05.26 - 05.28	2023 IEEE 3rd International Conference on Electronic Technology, Communication and Information (ICETCI)	Changchun, China	http://www.icetci.org/
05.26 - 05.28	2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)	Chongqing, China	http://www.iciba.org/
05.26 - 05.28	2023 International Conference on Microwave, Optical, and Communication Engineering (ICMOCE)	Bhubaneswar, India	https://www.iitbbs.ac.in/conference/icmoce-2023/
05.26 - 05.28	2023 4th International Conference for Emerging Technology (INCET)	Belgaum, India	http://www.incet.org/
05.26 - 05.27	2023 17th International Conference on Engineering of Modern Electric Systems (EMES)	Oradea, Romania	http://www.icemes.uoradea.ro/icemes2023/
05.26 - 05.28	2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC)	Jalandhar, India	https://www.nitj.ac.in/icsc2023/
05.28 - 06.01	2023 IEEE International Conference on Communications Workshops (ICC Workshops)	Rome, Italy	https://icc2023.ieee-icc.org/authors/call-workshop-papers
05.28 - 06.01	ICC 2023 - IEEE International Conference on Communications	Rome, Italy	https://icc2023.ieee-icc.org/
05.29 - 05.31	2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)	Milano, Italy	https://www.metrolivenv.org/
05.29 - 06.02	2023 IEEE International Conference on Robotics and Automation (ICRA)	London, United Kingdom	https://www.icra2023.org/
05.29 - 06.02	2023 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)	St. Petersburg, Russia	http://media-publisher.ru/en/about-weconf-2023/
05.29 - 05.31	2023 IEEE/ACIS 21st International Conference on Software Engineering Research, Management and Applications (SERA)	Orlando, Florida, USA	https://acisinternational.org/conferences/sera-2023/
05.29 - 06.03	2023 IEEE 73rd Electronic Components and Technology Conference (ECTC)	Orlando, Florida, USA	https://www.ectc.net/
05.29 - 05.31	2023 Zooming Innovation in Consumer Technologies Conference (ZINC)	Novi Sad, Serbia	https://www.gozinc.org/
05.30 - 06.02	2023 22nd IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)	Orlando, Florida, USA	http://www.ieee-itherm.net/
05.31 - 06.02	2023 American Control Conference (ACC)	San Diego, California, USA	https://acc2023.a2c2.org/
05.31 - 06.01	2023 International Conference on Engineering, Science and Advanced Technology (ICESAT)	Mosul, Iraq	https://icesat.org/
05.31 - 06.02	2023 Prognostics and Health Management Conference (PHM)	Paris, France	http://www.phmice.org/

>>2023년 6월

06.01. - 06.03.	2023 8th International Conference on Communication and Electronics Systems (ICCES)	Coimbatore, India	http://icoecs.org/2023/
06.01. - 06.04.	2023 IEEE International Conference on Predictive Control of Electrical Drives and Power Electronics (PRECEDE)	Wuhan, China	http://www.precede2023.com/
06.01. - 06.04.	2023 9th International Conference on Engineering, Applied Sciences, and Technology (ICEAST)	Vientiane, Laos	https://iceast.kmitl.ac.th/2023/
06.02. - 06.04.	2023 IEEE 5th Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)	Tainan, Taiwan	https://www.ecbios.asia/
06.04. - 06.08.	2023 IEEE Wireless Power Technology Conference and Expo (WPTCE)	San Diego, California, USA	https://ieee-wptce.org/
06.04. - 06.05.	2023 6th International Conference on Engineering Technology and its Applications (IICETA)	Al-Najaf, Iraq	https://iiceta.com/
06.04. - 06.09.	2023 IEEE Magnetic Society Summer School (MSSS)	Bari, Italy	https://ieemagnetics.org/

일자	학술대회명	개최장소	홈페이지/연락처
06.04. - 06.07.	2023 11th International Conference on Smart Grid (icSmartGrid)	Paris, France	http://www.icsmartgrid.com/index.php?id=main
06.04. - 06.07.	2023 IEEE Intelligent Vehicles Symposium (IV)	Anchorage, Alaska, USA	https://2023.ieee-iv.org/
06.04. - 06.10.	ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	Rhodes Island, Greece	https://2023.ieeeicassp.org/
06.05. - 06.08.	OCEANS 2023 - Limerick	Limerick, Ireland	https://limerick23.oceansconference.org/
06.05. - 06.07.	2023 IEEE 24th International Conference on High Performance Switching and Routing (HPSR)	Albuquerque, New Mexico, USA	https://hpsr2023.ieee-hpsr.org/
06.05. - 06.07.	2023 33rd ACM Great Lakes Symposium on VLSI (GLSVLSI)	Knoxville, Tennessee, USA	http://www.glsvlsi.org/
06.05. - 06.09.	2023 Days on Diffraction (DD)	Event Format: Virtual	http://www.pdmi.ras.ru/~dd/
06.05. - 06.07.	2023 IEEE International Conference on Advanced Robotics and Its Social Impacts (ARSO)	Berlin, Germany	https://ieee-arso.org/
06.05. - 06.06.	2023 IEEE Conference on Artificial Intelligence (CAI)	Santa Clara, California, USA	https://cai.ieee.org/2023/
06.05. - 06.07.	2023 IEEE International Conference on Prognostics and Health Management (ICPHM)	Montreal, Quebec, Canada	https://pnmconf.org/
06.06. - 06.08.	2023 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)	Brescia, Italy	https://www.metroind40iot.org/
06.06. - 06.09.	2023 IEEE International Conference on Environment and Electrical Engineering and 2023 IEEE Industrial and Commercial Power Systems Europe (IEEEIC / I&CPS Europe)	Madrid, Spain	https://www.eeeic.net/
06.06. - 06.09.	2023 International Conference on Unmanned Aircraft Systems (ICUAS)	Warsaw, Poland	http://www.uasconferences.com/2023_icuas/
06.06. - 06.09.	2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)	Gothenburg, Sweden	https://www.eucnc.eu/
06.06. - 06.09.	2023 24th International Conference on Process Control (PC)	Strbske Pleso, Slovakia	https://www.process-control.sk/
06.06. - 06.10.	2023 12th Mediterranean Conference on Embedded Computing (MECO)	Budva, Montenegro	https://mecoconference.me/meco2023/
06.06. - 06.08.	2023 International Conference on Localization and GNSS (ICL-GNSS)	Castellón, Spain	https://events.tuni.fi/icl-gnss2023/
06.07. - 06.09.	2023 10th International Conference on Recent Advances in Air and Space Technologies (RAST)	Istanbul, Turkey	https://www.rast.org.tr/
06.07. - 06.10.	2023 IEEE World AI IoT Congress (AllIoT)	Event Format: Virtual	https://worldaiotcongress.org/
06.08. - 06.10.	2023 11th International Conference on Thermal Equipment, Renewable Energy and Rural Development (TE-RE-RD)	Bucharest, Romania	https://www.tererd.upb.ro/
06.08. - 06.10.	2023 IEEE Sustainable Smart Lighting World Conference & Expo (LS18)	Mumbai, India	https://ieeesmartlightingworld.org/
06.08. - 06.10.	2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)	Istanbul, Turkey	http://www.horacongress.com/
06.08. - 06.09.	2023 9th International Workshop on Advances in Sensors and Interfaces (WASI)	Monopoli (Bari), Italy	https://mclabservices.di.uniroma1.it/iwasi/2023/aimsandscope.php
06.09. - 06.12.	2023 IEEE 14th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)	Shanghai, China	http://www.ieee-pedg2023.org/index.html
06.09. - 06.10.	2023 International Conference on Advanced & Global Engineering Challenges (AGEC)	Surampalem, Kakinada, India	http://aec.edu.in/agec/
06.09. - 06.10.	2023 17th International Conference on Engineering of Modern Electric Systems (EMES)	Oradea, Romania	http://www.icemes.uoradea.ro/icemes2023/
06.10. - 06.11.	2023 9th International Symposium on System Security, Safety, and Reliability (ISSSR)	Hangzhou, China	https://issr23.techconf.org/

일자	학술대회명	개최장소	홈페이지/연락처
06.11. - 06.13.	2023 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)	San Diego, California, USA	https://rfic-ieee.org/
06.11. - 06.15.	2023 IEEE IAS Pulp and Paper Industry Conference (PPIC)	Spokane, Washington, USA	https://www.pulppaper.org/about/ppic-info/
06.11. - 06.14.	2023 XXIX International Conference on Information, Communication and Automation Technologies (ICAT)	Sarajevo, Bosnia and Herzegovina	https://icat.etf.unsa.ba/
06.11. - 06.16.	2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)	Kyoto, Japan	https://www.vlssymposium.org/
06.11. - 06.13.	2023 IEEE 5th International Conference on Artificial Intelligence Circuits and Systems (AICAS)	Hangzhou, China	http://www.aicas2023.org/
06.11. - 06.16.	2023 IEEE/MTT-S International Microwave Symposium - IMS 2023	San Diego, California, USA	https://ims-ieee.org/
06.12. - 06.16.	2023 International Conference on Advancements in Nuclear Instrumentation Measurement Methods and their Applications (ANIMMA)	Viareggio/Pisa, Italy	https://animma.com/
06.12. - 06.14.	2023 24th International Carpathian Control Conference (ICCC)	Miskolc-Szilvásvárad, Hungary	http://mzsola.iit.uni-miskolc.hu/iccc2023/
06.12. - 06.15.	2023 12th Asia-Pacific International Conference on Lightning (APL)	Langkawi, Malaysia	https://attend.ieee.org/apl-2023/
06.12. - 06.14.	2023 International Conference on Future Energy Solutions (FES)	Vaasa, Finland	https://sites.uwasa.fi/fes2023/
06.12. - 06.14.	2023 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)	Gammarth, Tunisia	https://civemsa2023.ieee-ims.org/
06.12. - 06.15.	2023 IEEE 24th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)	Boston, Massachusetts, USA	https://coe.northeastern.edu/Groups/wowmom2023/
06.12. - 06.15.	2023 Photonics North (PN)	Montreal, Quebec, Canada	https://www.photonicsnorth.com/en
06.13. - 06.14.	2023 Multimedia Innovation and Digital Humanities International Conference (MIDHIC)	Event Format: Virtual	http://smmtc.uum.edu.my/MIDHIC2022
06.13. - 06.16.	2023 European Control Conference (ECC)	Bucharest, Romania	https://ecc23.euca-ecc.org/
06.13. - 06.15.	2023 IEEE International Conference on RFID (RFID)	Seattle, Washington, USA	https://2023.ieee-rfid.org/
06.14. - 06.16.	2023 32nd Annual Conference of the European Association for Education in Electrical and Information Engineering (EAEEIE)	Eindhoven, Netherlands	https://fontys.nl/EAEEIE-2023.htm
06.14. - 06.16.	2023 8th International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)	Nice, France	https://mt-its2023.eurecom.fr/
06.14. - 06.16.	2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)	Jeju, Korea (South)	https://memea2023.ieee-ims.org/
06.14. - 06.16.	2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)	Event Format: Virtual	http://icscss.com/
06.14. - 06.16.	2023 IEEE 17th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)	Tallinn, Estonia	https://taltech.ee/en/cpe-powereng2023
06.14. - 06.16.	2023 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)	Beijing, China	https://www.bmsb2023.com/index.html
06.14. - 06.16.	2023 AIAA/IEEE Electric Aircraft Technologies Symposium (EATS)	San Diego, California, USA	https://www.aiaa.org/aviation/presentations-papers
06.15. - 06.16.	2023 IEEE Women in Engineering International Leadership Conference (WIE ILC)	San Diego, California, USA	https://ieee-wie-ilc.org/
06.15. - 06.17.	2023 5th International Conference on Energy, Power and Environment: Towards Flexible Green Energy Technologies (ICEPE)	Shillong, India	http://nitm.ac.in/icepe2023/
06.16. - 06.23.	2023 IEEE/ACM 31st International Symposium on Quality of Service (IWQoS)	Orlando, Florida, USA	https://iwqos2023.ieee-iwqos.org/
06.16. - 06.18.	2023 11th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC)	Chongqing, China	http://www.icwoc.org/

일자	학술대회명	개최장소	홈페이지/연락처
06.16. - 06.19.	2023 International Conference Automatics, Robotics and Artificial Intelligence (ICARAI)	Sozopol, Bulgaria	http://e-university.tu-sofia.bg/e-conf/?konf=198
06.16	2023 101st ARFTG Microwave Measurement Conference (ARFTG)	San Diego, California, USA	https://www.arftg.org/
06.16	2023 IV International Conference on Neural Networks and Neurotechnologies (NeuroNT)	Saint Petersburg, Russia	https://etu.ru/en/university/conferences/neuront2023
06.16. - 06.17.	2023 International Conference on Applied Intelligence and Sustainable Computing (ICAISC)	Dharwad, India	https://icaisc.in/
06.17. - 06.24.	2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	Vancouver, British Columbia, Canada	https://cvpr.thecvf.com/
06.17	2023 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)	Shah Alam, Malaysia	https://sites.google.com/view/i2cacis/home?pli=1
06.17. - 06.21.	2023 ACM/IEEE 50th Annual International Symposium on Computer Architecture (ISCA)	Orlando, Florida, USA	https://iscaconf.org/isca2023/
06.18. - 06.21.	2023 IEEE 97th Vehicular Technology Conference (VTC2023-Spring)	Florence, Italy	https://events.vtsociety.org/vtc2023-spring/
06.18. - 06.19.	2023 International Conference on Sustaining Heritage: Innovative and Digital Approaches (ICSH)	Event Format: Virtual	https://heritage.uob.edu.bh/
06.19. - 06.22.	2023 IEEE 10th International Workshop on Metrology for AeroSpace (MetroAeroSpace)	Milan, Italy	https://www.metroaerospace.org/
06.19. - 06.23.	2023 International Wireless Communications and Mobile Computing (IWCMC)	Marrakesh, Morocco	https://iwcmc.org/2023/
06.19. - 06.22.	2023 ARC Workshop on Electrochemical Energy Conversion and Power Electronics (WEEPE)	Narvik, Norway	https://uit.no/tavla/artikkel/793268/2023_ieee_arc_workshop_on_electrochemical_energy
06.19. - 06.21.	2023 27th International Conference Electronics	Palanga, Lithuania	http://electronicsconf.ktu.edu/index.php/elc
06.19. - 06.23.	2023 IEEE 9th International Conference on Network Softwarization (NetSoft)	Madrid, Spain	https://netsoft2023.ieee-netsoft.org/
06.20. - 06.23.	2023 8th International Conference on Smart and Sustainable Technologies (SpliTech)	Split/Bol, Croatia	https://2023.splitech.org/home
06.20. - 06.23.	2023 18th Iberian Conference on Information Systems and Technologies (CISTI)	Aveiro, Portugal	http://cisti.eu/index.php/en/
06.20. - 06.22.	2023 IEEE Conference on Innovation Management (INNOCONF)	Los Angeles, California, USA	https://2023.innoconf.org/
06.20. - 06.22.	2023 15th International Conference on Quality of Multimedia Experience (QoMEX)	Ghent, Belgium	https://sites.google.com/view/qomex2023
06.21. - 06.23.	2023 IEEE Transportation Electrification Conference & Expo (ITEC)	Detroit, Michigan, USA	https://itec-conf.com/
06.21. - 06.22.	2023 IEEE Cognitive Communications for Aerospace Applications Workshop (CCAAW)	Cleveland, Ohio, USA	https://ieee-ccaa.com/
06.21. - 06.23.	2023 19th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)	Montreal, Quebec, Canada	http://www.wimob.org/wimob2023/
06.21. - 06.23.	2023 IEEE 25th Conference on Business Informatics (CBI)	Prague, Czech Republic	https://cbi2023.org/index.php
06.22. - 06.24.	2023 IEEE 36th International Symposium on Computer-Based Medical Systems (CBMS)	L'Aquila, Italy	https://2023.cbms-conference.org/
06.22. - 06.23.	2023 IEEE 7th Portuguese Meeting on Bioengineering (ENBENG)	Porto, Portugal	http://embs.ieee-pt.org/7th-enbeng-2023/
06.23. - 06.25.	2023 IEEE/ACIS 23rd International Conference on Computer and Information Science (ICIS)	Wuxi, China	https://acisinternational.org/conferences/icis-2023/
06.24. - 06.29.	2023 38th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)	Boston, Massachusetts, USA	https://lics.siglog.org/lics23/
06.25. - 06.30.	2023 IEEE International Symposium on Information Theory (ISIT)	Taipei, Taiwan	https://isit2023.org/

일자	학술대회명	개최장소	홈페이지/연락처
06.25. - 06.30.	2023 30th International Symposium on Discharges and Electrical Insulation in Vacuum (ISDEIV)	Okinawa, Japan	http://isdeiv2023.w3.kanazawa-u.ac.jp/
06.25. - 06.29.	2023 IEEE Belgrade PowerTech	Belgrade, Serbia	https://powertech2023.com/
06.25. - 06.29.	2023 IEEE Pulsed Power Conference (PPC)	San Antonio, Texas, USA	http://www.ppc2023.org/
06.25. - 06.28.	2023 20th International Conference on Ubiquitous Robots (UR)	Honolulu, Hawaii, USA	https://2023.ubiquitousrobots.org/
06.25. - 06.30.	2023 International Forum on MPSoC for Software-Defined Hardware (MPSoC)	Helena, Montana, USA	http://mpsoc-forum.org/
06.25. - 06.28.	2023 Device Research Conference (DRC)	Santa Barbara, California, USA	https://www.mrs.org/drc-2023
06.25. - 06.29.	2023 22nd International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers)	Kyoto, Japan	https://www.transducers2023.org/
06.25. - 06.28.	2023 IEEE 24th Workshop on Control and Modeling for Power Electronics (COMPEL)	Ann Arbor, Michigan, USA	https://ieee-compel.org/
06.26. - 06.30.	2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)	Torino, Italy	https://ieeecompsac.computer.org/2023/
06.26. - 06.28.	2023 IEEE International Conference on Metaverse Computing, Networking and Applications (MetaCom)	Kyoto, Japan	https://www.ieee-metacom.org/2023/
06.26. - 06.28.	2023 21st IEEE Interregional NEWCAS Conference (NEWCAS)	Edinburgh, United Kingdom	https://2023.ieee-newcas.org/
06.26. - 06.30.	2023 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)	Munich, Germany	https://www.cleo-europe.org/
06.26. - 06.30.	2023 IEEE International Conference on Smart Computing (SMARTCOMP)	Nashville, Tennessee, USA	https://smartcomp.isis.vanderbilt.edu/
06.26. - 06.30.	2023 Radiation and Scattering of Electromagnetic Waves (RSEMW)	Divnomorskoe, Russia	http://rsemw.sfedu.ru/
06.26. - 06.27.	2023 China Semiconductor Technology International Conference (CSTIC)	Shanghai, China	https://www.semiconchina.org/en/5
06.26. - 06.29.	2023 IEEE 11th International Conference on Healthcare Informatics (ICHI)	Houston, Texas, USA	https://ieeichi.github.io/ICHI2023/
06.26. - 06.30.	2023 ACM/IEEE Joint Conference on Digital Libraries (JCDL)	Santa Fe, New Mexico, USA	https://2023.jcdl.org/
06.27. - 06.30.	2023 53rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)	Porto, Portugal	https://dsn2023.dei.uc.pt/
06.27. - 06.29.	2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP)	Ohrid, Macedonia	http://iwSSIP.feit.ukim.edu.mk/
06.27. - 06.30.	2023 26th International Conference on Information Fusion (FUSION)	Charleston, South Carolina, USA	https://www.fusion2022.se/
06.28. - 06.30.	2023 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)	Winnipeg, Manitoba, Canada	https://nemo-ieee.org/
06.28. - 06.30.	2023 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)	Seattle, Washington, USA	http://aim2023.org/
06.28. - 07.01.	2023 IEEE International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE)	Paris, France	http://www.olab-dynamics.net/wetice2023/index.html
06.28. - 06.30.	2023 IEEE International Workshop on Metrology for Automotive (MetroAutomotive)	Modena, Italy	https://www.metroautomotive.org/
06.28. - 06.30.	2023 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)	Pskov, Russia	http://media-publisher.ru/en/about-synchroinfo-2023/
06.28. - 06.30.	2023 12th International Conference on Modern Circuits and Systems Technologies (MOCAS)	Athens, Greece	http://mocas.physics.auth.gr/
06.28. - 06.30.	2023 Power Quality and Electromagnetic Compatibility at Low Frequency (PQEMC-LF)	Craiova, Romania	http://pqemc-lf.ucv.ro/

일자	학술대회명	개최장소	홈페이지/연락처
06.29. - 06.30.	2023 19th International Conference on Intelligent Environments (IE)	Uniciti, Mauritius	https://ie2023.mdxmru.com/
06.29. - 06.30.	2023 15th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)	Bucharest, Romania	https://ecai.ro/
06.29. - 07.03.	2023 IEEE 24th International Conference of Young Professionals in Electron Devices and Materials (EDM)	Novosibirsk, Russia	https://edm.ieeesiberia.org/
06.29. - 07.01.	2023 18th Conference on Electrical Machines, Drives and Power Systems (ELMA)	Varna, Bulgaria	https://elma.tu-varna.bg/
06.29. - 07.01.	2023 58th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)	Nis, Serbia	https://icestconf.org/
06.30. - 07.03.	2023 Sixth International Symposium on Computer, Consumer and Control (IS3C)	Taichung, Taiwan	http://is3c2023.ncuteecs.org/

The Magazine of the IEIE

특별회원사 명단

회원사	대표자	주소	전화	홈페이지
(주)디비하이텍	최창식	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)에어포인트	백승준	대전광역시 유성구 테크노2로 187, 204호(용산동, 미건테크노월드 2차)	042-484-5460	http://www.airpoint.co.kr
(주)와이솔	염상덕	경기도 오산시 가장로 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	구현모, 박종욱	경기도 성남시 분당구 정자동 206	031-727-0114	http://www.kt.com
LG이노텍(주)	정철동	서울시 강서구 마곡중앙10로 30	02-3777-1114	www.lginnotek.com
LG전자(주)	조주완, 배두용	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, 벤델레코츠마티아스	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)	박정호, 이석희	경기도 이천시 부발읍 아미리 산 136-1	031-630-4114	http://www.skhynix.com
네이버(주)	한성숙	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	임헌문	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
리얼텍코리아 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
삼성전자(주)	한중희	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)	박영우	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.airsmec.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
정보통신정책연구원	권호열	충북 진천군 덕산읍 정통로 18	043-531-4389	www.kisdi.re.kr
㈜LX세미콘	손보익	대전시 유성구 탑립동 707	042-712-7700	www.lxsemicon.com

회원사	대표자	주소	전화	홈페이지
㈜넥스틴	박태훈	경기도 화성시 동탄면 동탄산단9길 23-12	031-629-2300	http://www.nextinsol.com
㈜더즈텍	김태진	경기도 안양시 동안구 학의로 292 금강펜테리움IT타워 A동 1061호	031-450-6300	http://www.doestek.co.kr
㈜만도	정동원, 조성현, 김광현	경기도 평택시 포승읍 하만호길 32	02-6244-2997	www.mando.com
㈜빅텍	임만규	경기도 이천시 마장면 덕이로 180-31	031-631-7301	http://www.vitek.co.kr
㈜스프링클라우드	송영기	경기도 성남시 창업로 42	031-778-8328	www.aspringcloud.com
㈜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
㈜실리콘마이터스	허염	경기도 성남시 분당구 대왕판교로 660 유스페이스-1 A동 8층	1670-7665	http://www.siliconmitus.com
㈜싸이몬	정창호	경기도 성남시 분당구 별말로48(구 야탑동 272-1 케이디티빌딩)	02-480-8580	http://www.cimon.com
㈜싸인텔레콤	박영기	서울시 영등포구 경인로 775, 문래동 3가 에이스하이테크시티 1동 119호	02-3439-0033	http://www.signtelecom.com
㈜솔리드	정준, 이승희	경기도 성남시 분당구 판교역로 220 솔리드스페이스	031-627-6000	http://www.st.co.kr
㈜와이솔	염상덕	경기도 오산시 가장로 531-7	070-7837-2730	http://www.wisol.co.kr
㈜유니트론텍	남궁 선	서울시 강남구 영동대로 638(삼도빌딩) 9층	02-573-6800	http://unitrontech.com
㈜코클리어닷컴에이아이	한윤창	서울시 강남구 봉은사로 51길 26		www.cochl.ai
㈜크레셈	오상민	대전시 유성구 대덕대로 582, 4층 402호(도룡동, 옥토빌딩)	031-427-3445	http://www.cressem.com
㈜텔레칩스	이장규	서울시 송파구 올림픽로 35다길 42(신천동 한국루터회관)19~23층	02-3443-6792	www.telechips.com
㈜티에이치엔	이광연, 채승훈	대구시 달서구 갈산동 973-3	053-583-3001	http://www.th-net.co.kr
㈜티엘아이	김달수	경기도 성남시 중원구 양현로 405번길 12 티엘아이 빌딩	031-784-6800	http://www.tli.co.kr
㈜해치텍	최성민	충북 청주시 청원구 오창읍 연구단지로 40, 스타기업관 207호	043-715-9034	http://www.haechitech.com
중소벤처기업진흥공단	김학도	경상남도 진주시 동진로 430	055-751-9380	www.kosmes.or.kr
케이케이테크(주)	다케시게신이치, 황정성	경기도 안성시 대덕면 무능로132	031-678-1586	http://www.k-kttech.co.kr
코어인사이트㈜	유용훈	경기도 성남시 중원구 갈매치로 186 반포테크노피아 5층	031-750-9200	http://www.coreinsight.co.kr
한국알박㈜	김선길	경기도 평택시 청북읍 한산길5	031-683-2922	http://www.ulvackora.co.kr
한국인터넷진흥원	이원태	서울시 송파구 중대로 135 (가락동) IT벤처타워	02-405-5118	http://www.kisa.or.kr
한국전기연구원	명성호	경남 창원시 성산구 불모산로10번길 12 (성주동)	055-280-1114	http://www.keri.re.kr
한국전자기술연구원	김영삼	경기도 성남시 분당구 새나리로 25 (야탑동)	031-789-7000	http://www.keti.re.kr
한국전자통신연구원	김명준	대전시 유성구 가정로 218	042-860-6114	http://www.etri.re.kr
한화시스템㈜	김연철	서울시 중구 청계천로 86 (장교동) 한화비딩 (19,20층)	02-729-3030	http://www.hanwhasystems.com
현대로템㈜	이용배	경기도 의왕시 철도박물관로 37	031-596-9114	http://www.hyundai-rotem.co.kr
현대모비스㈜	조성환	서울시 강남구 테헤란로 203	02-2018-5114	http://www.mobis.co.kr
현대자동차㈜	정의선, 하연태	경기도 화성시 장덕동 772-1	02-3464-1114	http://www.hyundai-motor.com
호리바에스텍코리아㈜	김성환 외 1명	경기도 용인시 수지구 디지털벨리로 98 호리바빌딩	031-6520-6500	http://www.horiba.com
히로세코리아㈜	이상엽	경기도 시흥시 정왕동 희망공원로 250	031-496-7000	http://www.hirose.co.kr
히타치하이테크코리아㈜	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 : biz@theieie.org

TEL : (02)553-0255(내선 5) FAX : (02)552-6093



대한전자공학회

The Institute of Electronics and Information Engineers

전자공학회지 <월간>

제50권 제2호(통권 제465호)

The Magazine of the IEIE

2023년 2월 20일 인쇄

2023년 2월 25일 발행

발행및

(사) 대한전자공학회

회장 이 혁 재

편집인

인쇄인

한림원(주)

대표 김 흥 중

발행인

사 단 법 인 대 한 전 자 공 학 회

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

TEL.(02)553-0255~7 FAX.(02)552-6093

E-mail : ieie@theieie.org

Homepage : http://www.theieie.org

씨티은행 102-53125-258



2023년도 회비납부 안내

1. 회비의 납부 및 유효기간

2023년도 회원 연회비는 2022년과 동일함을 알려드리며, 아직 2023년도 회비를 납부하지 않으신 회원님께서서는 납부하여 주시기 바라며, 연회비의 유효기간은 회비를 납부한 당해연도에 한합니다.

◆ 2023년도 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 : 2023년 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

첨단기술로 더 나은 환경을 만듭니다

더 나은 미래를 열어가는 기술
환경문제를 해결하는 기술
반도체의 미래를 준비하는 일
지금, SK하이닉스가 하고 있습니다
We Do **Green** Technology

