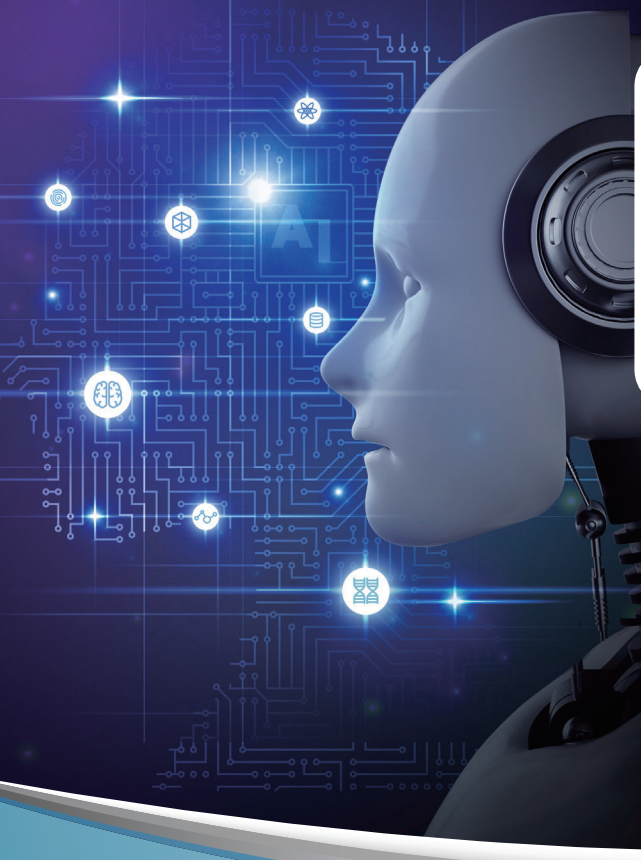
ISSN 1016-9288
제51권 5호
2024년 5월호

전자공학회지

The Magazine of the IEIE

vol.51. no.5



생성형 AI와 AI 융합 활용 기술

- GPU를 이용한 AI 활용
- 생성형 인공지능의 연구 동향
- ChatGPT와 거대언어모델
- 실물 환경에서의 데이터 기반 강화학습 적용 방안
- 반도체 공정 장비에서 인공지능 기술의 역할과 활용
- 인공지능을 활용한 기계의 건전성 예측 및 관리 방법

대한전자공학회
IEIE The Institute of Electronics and Information Engineers
www.theieie.org

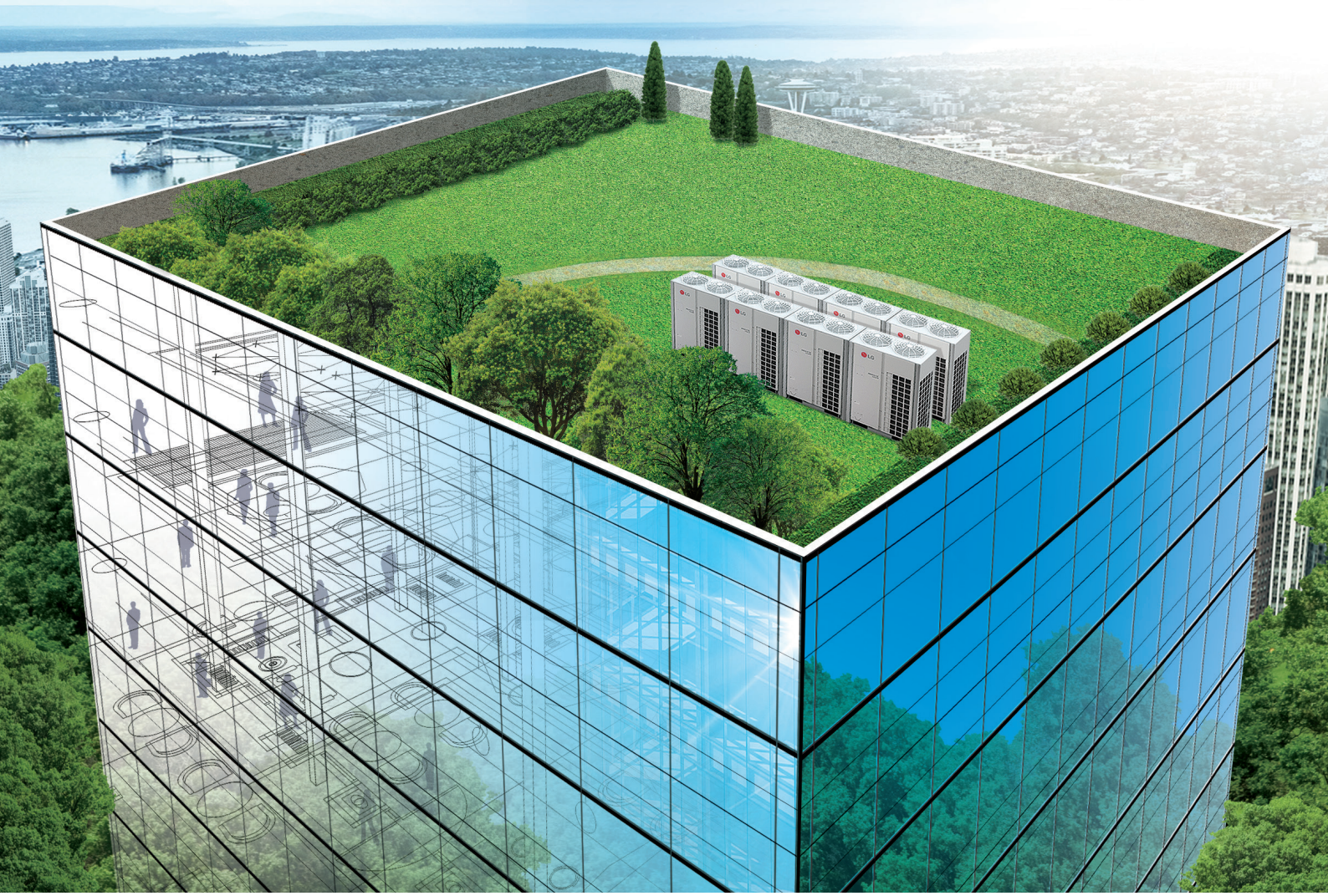
LG전자

쾌적함의 완성
그 차이는 설계입니다

서로 다른 크기, 형태, 구조의 건물들이지만 어디서나 최상의 쾌적함을 누릴 수 있는 이유
바로 LG휘센의 공기 설계가 있기 때문입니다.
공간의 온도, 습도, 청정도를 최적으로 맞춤 설계하는 LG휘센 시스템에어컨,
차원이 다른 쾌적함을 직접 경험해 보십시오.

온도 | 습도 | 공기청정

공기를 설계하다
LG WHISEN 시스템에어컨



가정용에서 산업발전용까지, 세상 모든 건물을 위한 종합공조기술은 LG휘센이 유일합니다

중대형 건물용
시스템에어컨
MULTI V. SUPER 5

중소형 상업용
냉난방 에어컨
냉난방에어컨 / 천장형에어컨

대형건물 / 발전소용
중앙공조 시스템
TURBO CHILLER

대형 상업건물용
중앙공조 시스템
ABSORPTION CHILLER

중소형건물/ 공장용/ 산업용
고효율 히트펌프 칠러 시스템
INVERTER SCROLL CHILLER
AIR WATER GEO

공기를 설계하다
LG WHISEN 시스템에어컨

구입 및 제품 문의 : 1544-8777
서비스 문의 : 1544-7777 / 1588-7777
www.lge.co.kr/kr/business

2024년도 대한전자공학회

학회상 후보자 추천

사단법인 대한전자공학회에서는 매년 전자.정보.통신 분야에 탁월한 업적을 이루고 전자공학의 발전에 크게 공헌한 분에게 학회상을 아래와 같이 시상하고 있습니다. 금년에도 회원 여러분께서 훌륭하신 후보자를 추천하여 주시면 감사하겠습니다.

1. 시상부문

시 상 부 문	인 원	시 상 자 격	시 상 내 용
대한전자공학대상	1명	전자.정보.통신 및 그 관련 분야에 탁월한 업적이 있는 자	상패 및 부상 (2,000만원)
기술혁신상	1명	전자공학 기술발전에 현저한 업적을 이룩한 자 또는 기업의 기술혁신에 기여한 자	상패 및 부상
IEIE Research Pioneer Award	1명	50세 이하로서 전자,정보,통신 및 그 관련 분야에서 새로운 연구분야를 개척, 확대시킨 업적이 있는 자	상패 및 부상 (500만원)
논문상 (TC,SD,CI,SP, SC,IE)	6명	우수한 논문을 대한전자공학회 논문지 및 해외 저명 학술지에 발표한 자로서 6개 Society (TC,SD,CI,SP,SC,IE)별 각 1인(*)	상패 및 부상

* 최근 5년간 전자공학회 논문지에 3편 이상의 저널 논문이 포함되어야 함.

2. 추천권자

가. 소속기관장 (연구소, 대학, 기업체, 행정기관 등)　　나. 개인(본인 포함)

* 단, 대한전자공학대상은 소속기관장의 추천에 한함.

3. 제출서류

* 제출서류 작성양식은 학회 홈페이지(http://www.theieie.org)를 참조하기 바람.

4. 서류 또는 이메일 접수

가. 접수마감: 2024년 7월 16일(화)

나. 접 수 처: 서울 강남구 테헤란로 7길 22(역삼동) 과학기술회관 1관 907호
대한전자공학회 사무국 이안순 부장 (02-553-0255(내선 6번), ieie@theieie.org)

5. 수상자 발표 및 시상

가. 발 표: 2024년 11월 초순

나. 시 상: 2024년 11월 22일(금) 정기총회(장소: 추후공지)

2024년도 대한전자공학회

해동상 후보자 추천

사단법인 대한전자공학회에서는 매년 전자.정보.통신 분야에 탁월한 업적을 이루고 전자공학의 발전에 크게 공헌한 분에게 해동상을 아래와 같이 시상하고 있습니다. 해동상은 대덕전자(주) 고, 김정식 회장께서 우리나라 전자공학 분야의 학문 발전과 기술 발전을 위하여 크게 업적을 쌓은 분들의 노고를 치하하고 업적을 기리기 위하여 해동과학문화재단을 설립함으로써 제정되었습니다. 금년에도 회원 여러분께서 훌륭하신 후보자를 추천하여 주시면 감사하겠습니다.

1. 시상부문 및 자격

시상부문	인원	시상자격	시상내용
해동학술상	1명	학회 정회원 혹은 평생회원(회원자격 최근 2년 이상 유지)이어야 하며, 최근 10년간 본 학회 논문지 및 SCI-E급 국제저널에 논문 게재 등 전자 · 정보 · 통신 및 관련 분야 학술 활동에 탁월한 업적이 있는 자(*)	상패 및 부상 (2,500만원)
해동기술상	1명	기업체, 공공연구기관 및 정부출연기관에 근무하는 임직원 중 우수한 기술개발 업적을 달성하여 국가적으로 전자 · 정보 · 통신 기술 발전에 크게 기여한 자	상패 및 부상 (2,500만원)
해동젊은공학인상 (학술상)	1명	학회 정회원 혹은 평생회원(회원자격 최근 2년 이상 유지)이며, 만 40세 이하인 자로 최근 5년간 본 학회 학술대회 및 학술지에 논문 게재 등 전자 · 정보 · 통신 및 관련 분야 학술 활동에 탁월한 업적이 있는 자(**)	상패 및 부상 (1,000만원)
해동젊은공학인상 (기술상)	1명	만 40세 이하인 자로, 우수한 기술개발 업적을 달성하여 국가적으로 전자 · 정보 · 통신 기술 발전에 크게 기여한 자	상패 및 부상 (1,000만원)

* 해동학술상은 최근 10년간 본 학회 학술지(전자공학회 논문지, JSTS, IEIE SPC)에 5편 이상의 논문을 게재한 자로서 10년간(2014년 7월 1일~2024년 6월 30일)의 실적을 평가함.

** 해동젊은공학인상(학술상)은 당해년도 만 40세 이하로서 최근 5년간 본 학회 학술대회(하계, 추계, ICEIC, ICCE-Asia, ITC-CSCC) 및 학술지(전자공학회논문지, JSTS, IEIE SPC)에 5편 이상의 논문을 게재한 자로서 5년간(2019년 7월 1일~2024년 6월 30일)의 실적을 평가함.

※ 해동상에 제출하는 논문은 마감일 기준으로 게재된 논문만 인정하며, 아래 조건 중 하나 이상에 해당하는 자는 후보가 될 수 없음.

- 타 기관 해동상 수상자
- 동일한 연구개발 업적으로 본 학회 또는 타 기관에서 수상한 자
- 10년 이내 대한전자공학회에서 수여하는 다른 상(해동상 타 부문, IEIE/IEEE Joint Award 등) 수상자

2. 추천권자

가. 소속기관장 (연구소, 대학, 기업체, 행정기관 등)　　나. 개인(본인 포함)

3. 제출서류

* 제출서류 작성양식은 학회 홈페이지(http://www.theieie.org)를 참조하기 바람.

4. 서류 또는 이메일 접수

가. 접수마감: 2024년 7월 16일(화)

나. 접 수 처: 서울 강남구 테헤란로 7길 22(역삼동) 과학기술회관 1관 907호

대한전자공학회 사무국 이안순 부장 (02-553-0255(내선 6번), ieie@theieie.org)

5. 수상자 발표 및 시상

가. 발 표: 2024년 11월 초순

나. 시 상: 2024년 12월 13일(금) (장소: 추후공지)

The Institute of Electronics and
Information Engineers
Summer Annual Conterence of IEIE, 2024

2024 하계종합 학술대회

6. 26(수)~28(금)
롯데호텔 제주(중문)

논문모집
홈페이지 http://conf.theieie.org

최우수/일반 논문

논문제출 : 2024년 5월 10일(금)
심사통보 : 2024년 5월 24일(금)
사전등록 : 2024년 6월 7일(금)

발표분야 (학회 6개 소사이어티)

소사이어티	연구회
통신(Communication)	통신, 지능형네트워크, 마이크로파 및 전파전파, ITS, 군사전자, 무선PAN/BAN
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 및 퍼지시스템, M2M/IoT, 휴먼ICT, 인공지능 및 보안, 증강휴먼, AI응용, 블록체인
인공지능 신호처리(AI Signal Processing)	영상신호처리, 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, 디지털통신시스템
New Emerging Area	의료, 에너지, Software, 기타

행사문의

학회 사무국 전화 02)553-0255(내선 2번) FAX 02)552-6093
홈페이지 http://conf.theieie.org e-mail conf@theieie.org

대한전자공학회
IEIE The Institute of Electronics and Information Engineers

IEEE
IEIE

ICCE-Asia 2024

The 9th International Conference on Consumer Electronics (ICCE) Asia

11. 3(Sun) - 11. 6(Wed), 2024
Sheraton Grand Danang, Vietnam

Presentation Guidelines

The conference will be held with face-to-face presentations of papers at the conference site at Da Nang, Vietnam.

Organized by the IEEE Consumer Technology Society and the Institute of Electronics and Information Engineers, ICCE-Asia 2024 which will be held in the Da Nang, Vietnam is an event open to researchers and engineers from industry, research centres, and academia to exchange information and results related to consumer electronics (CE). The conference will feature outstanding keynote speakers, high quality tutorials, special sessions and peer-reviewed papers. It hopes to attract a global audience from industry and academia. It is a perfect opportunity to promote affiliated company/organization to an audience of world-class researchers in the CE industry.

TOPICS OF IEEE/IEIE ICCE-ASIA 2024

- Artificial Intelligence and Machine Learning for CE Applications (AIM)
- Robotics, Drones, Automation Technologies and Interfaces (RDA)
- Security and Privacy of CE Hardware and Software Systems (SPC)
- Energy Management of CE Hardware and Software Systems (EMC)
- Application-Specific CE for Smart Cities (SMC)
- RF, Wireless, and Network Technologies (WNT)
- Internet of Things and Internet of Everywhere (IoT)
- Entertainment, Gaming, and Virtual and Augmented Reality (EGV)
- AV Systems, Image and Video, and Cameras and Acquisition (AVS)
- Automotive CE Applications (CEA)
- CE Sensors and MEMS (CSM)
- Consumer Healthcare Systems (CHS)
- Enabling and HCI Technologies (HCI)
- Smartphone and Mobile Device Technologies (MDT)
- Semiconductor Devices for Consumer Electronics (SCE)
- Other Technologies Related with CE (MIS)

SPECIAL SESSIONS

Special session proposals are invited to IEEE/IEIE ICCE-Asia 2024, and inquiries regarding submission should be directed to the Special Session Chair.

BEST PAPER AWARDS

The authors of the best papers will be presented Gold, Silver, and Bronze awards.
Selected top quality papers will be recommended to be published in the Journal of Semiconductor Technology and Science (JSTS) or a special issue of IEIE Transactions on Smart Processing and Computing.

PAPER SUBMISSION

Prospective authors can submit their papers by following the guidelines posted on the conference webpage (http://www.icce-asia2024.org).
Accepted papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore's scope and quality requirements.

AUTHOR'S SCHEDULE

- Submission of Paper : August 16th, 2024
- Notification of Acceptance : September 5th, 2024
- Submission of Final Paper : September 13th, 2024

CONTACT POINT

- Secretariat : inter@theieie.org

IEEE IEEE CTSoc CONSUMER TECHNOLOGY SOCIETY IEIE

전자공학회논문지
단편논문 신설 안내

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

논문투고시스템 투고구분 선택

☞ **정규논문(기존)** : 긴급 / 일반 중 택일

▶ 단편논문(신설) : 특급

단편논문 양식

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

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

* 심사본 : 3쪽 이내, 최종본: 4쪽 이내

단편논문 심사비 : 10만원/편당

단편논문 게재료 : 게재논문 면당 5만원, 최대 4쪽 이내

☞ 지원 문구 추가 시 10만원 추가

☞ 교신저자가 비회원인 경우 산정된 게재료의 150% 부과

단편논문 심사 기간 : 2주 이내 1차 심사를 원칙으로 함

시 행 : 2022년 10월 이후

IEIE | 대한전자공학회

The Institute of Electronics and Information Engineers

IEEE DiscoveryPoint Communications(IDPC)

IEEE
DiscoveryPoint
for Communications

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 컨텐츠 제공	설계 문제에 대한 솔루션 제공하여 실무 엔지니어에게 최적화

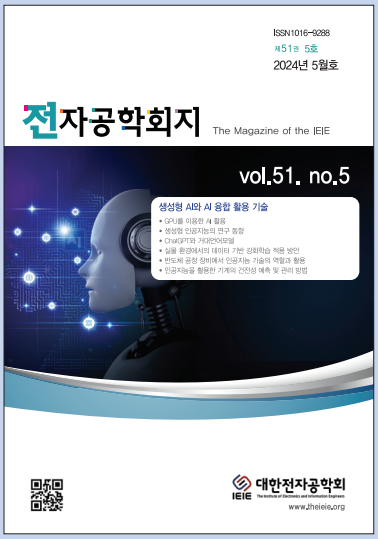
Trial가능,
문의 02-3474-5290
이희진 과장 hjlee@kitis.co.kr

IEEE | Authorized Dealer

키티스 産學研情報(株)
KITIS Info. & Co., Ltd.
www.kitis.co.kr / T. 02 3474 5290

CONTENTS

제51권 5호 (2024년 5월)



※ 학회지 5월호 표지 (vol 51. No 5)

대한전자공학회
IEIE The Institute of Electronics and Information Engineers

2024년도 임원 및 각 위원회 위원

회 장	이 충 용 (연세대학교 교수)	
수석부회장	백 광 현 (중앙대학교 교수)	
고 문	권 오 경 (한양대학교 석좌교수)	김 기 남 (삼성전자 고문)
	김 영 재 (해동과학문화재단 이사장)	박 성 욱 (차세대지능형반도체사업단 이사장)
	방 승 찬 (한국전자통신연구원 원장)	신 희 동 (한국전자기술연구원 원장)
	안 승 권 (연암공과대학교 총장)	전 영 현 (삼성전자 부회장)
	천 경 준 (씨젠 회장)	최 창 식 (DB하이텍 부회장)
감 사	이 재 훈 (유정시스템 대표이사) – 감사	전 병 우 (성균관대학교 교수) – 감사
부 회 장	강 문 식 (강릉원주대학교 교수) – 교육/표준화	구 본 태 (한국전자통신연구원 본부장) – 산학연
	권 호 열 (강원대학교 교수) – 대외협력/정책자문	김 종 옥 (고려대학교 교수) – 사업총괄/대외협력
	김 진 상 (경희대학교 교수) – 운영자문/대외협력	노 원 우 (연세대학교 교수) – 국제총괄/ICCE Asia/영문논문/신기술
	노 태 문 (한국전자통신연구원 연구전문위원) – 산학연	류 수 정 (사피온코리아 대표이사) – 회원 및 여성
	송 병 철 (인하대학교 교수) – 학술총괄/SPC/하계/추계/논문편집	이 승 호 (한밭대학교 교수) – 대외협력/지부
	정 영 모 (한성대학교 교수) – 홍보 및 정보화 총괄/회지편집	
소사이어티 회장	고 병 철 (계명대학교 교수) – 인공지능신호처리	고 정 환 (인하공업전문대학 교수) – 산업전자
	신 오 순 (숭실대학교 교수) – 통신	장 성 진 (삼성전자 상근고문) – 반도체
	주 영 복 (한국기술교육대학교 교수) – 시스템 및 제어	최 용 수 (신한대학교 교수) – 컴퓨터
협동부회장	강 성 원 (한국전자통신연구원 부원장)	김 강 태 (삼성전자 부사장)
	김 상 태 (전남테크노파크 본부장)	김 영 한 (UC San Diego / 가우스랩스 대표이사)
	김 은 원 (대림대학교 교수)	김 형 준 (한국과학기술연구원 소장)
	노 승 원 (LG이노텍 CTO)	송 문 섭 (심텍 회장)
	심 동 규 (광운대학교 교수)	엄 낙 웅 (한국전자통신연구원 책임연구원)
	오 윤 제 (정보통신기획평가원 PM)	원 제 형 (도쿄일렉트론코리아 대표이사)
	유 창 동 (KAIST 교수)	이 광 엽 (서경대학교 교수)
	이 동 규 (카카오모빌리티 부사장)	이 병 선 (김포대학교 교수)
	이 상 훈 (웨이브피아 대표이사)	이 서 규 (한국팹리스산업협회 회장)
	이 장 규 (텔레칩스 대표이사)	이 재 관 (한국자동차연구원 소장)
	인 치 호 (세명대학교 교수)	전 선 익 (파이낸셜뉴스 부회장)
	정 은 승 (삼성전자 고문)	정 준 (쏠리드 대표이사)
상 임 이 사	강 명 곤 (한국교통대학교 교수) – 기획/신기술	강 석 주 (서강대학교 교수) – 총무
	고 형 호 (충남대학교 교수) – 추계	권 혁 인 (중앙대학교 교수) – 사업 총괄2
	김 동 현 (연세대학교 교수) – 바이오메디컬연구회 위원장	김 영 로 (명지전문대학 교수) – 회지편집
	김 영 민 (서울대학교 교수) – 하계	김 용 신 (고려대학교 교수) – 대외협력/산학연
	김 원 종 (한국전자통신연구원 책임연구원) – 표준화	김 윤 (서울시립대학교 교수) – 사업
	김 익 균 (한국전자통신연구원 본부장) – 하계	김 재 준 (서울대학교 교수) – JSTS
	김 태 욱 (연세대학교 교수) – ICCE Asia 총괄/국제협력	김 현 (서울과학기술대학교 교수) – ICEIC/기획
	김 훈 (인천대학교 교수) – 하계 총괄	동 성 수 (용인예술과학대학교 교수) – 교육
	박 철 수 (광운대학교 교수) – SPC	변 대 석 (삼성전자 마스터) – 교육
	서 병 석 (상지대학교 교수) – 홍보 및 정보화	선 우 경 (서울대학교 교수) – 회원 및 여성
	손 교 민 (삼성전자 마스터) – 사업	송 진 호 (연세대학교 교수) – ICCE Asia
	연 규 봉 (한국자동차연구원 팀장) – 사업	유 찬 세 (한국전자기술연구원 수석연구원) – 사업 총괄1
	이 강 윤 (성균관대학교 교수) – 추계 총괄	이 덕 우 (계명대학교 교수) – 논문 편집
	이 형 민 (고려대학교 교수) – ICCE Asia	이 후 진 (한성대학교 교수) – ICCE Asia
	정 승 원 (고려대학교 교수) – 하계	정 일 권 (한국전자통신연구원 본부장) – 사업
	조 성 재 (이화여자대학교 교수) – ICEIC/재무	조 성 현 (한양대학교 교수) – 논문편집 총괄
	한 재 호 (고려대학교 교수) – 신기술	황 원 준 (아주대학교 교수) – 회지편집
산업체이사	강 석 판 (LG전자 상무)	고 용 남 (하나마이크론 전무)
	김 동 현 (ICTK 대표이사)	김 현 수 (삼성전자 상무)
	배 순 민 (케이티 Lab장)	서 경 열 (뷰웍스 부사장)
	오 의 열 (LG디스플레이 연구위원)	우 정 호 (비전넥스트 대표이사)
	윤 영 권 (삼성전자 마스터)	이 도 훈 (국가보안기술연구소 수석연구원)
	이 수 민 (한국센서연구소 대표이사)	이 수 인 (텔레칩스 상무)
	조 영 민 (SkyMirr CEO)	천 이 우 (넥스트칩 연구소장)
	최 성 민 (해치텍 대표이사)	한 은 혜 (에스에스앤씨 대표이사)
	함 철 희 (삼성전자 마스터)	홍 국 태 (LX세미콘 연구위원)
이 사	공 규 열 (한성대학교 교수) – 홍보/정보화	궁 재 하 (고려대학교 교수) – 대외협력

권 구 덕 (강원대학교 교수) – 기획
권 태 수 (서울과학기술대학교 교수) – 정보화
김 선 욱 (고려대학교 교수) – 교육
김 소 영 (성균관대학교 교수) – JSTS
김 재 욱 (한국과학기술연구원 선임연구원) – 학술(하계)
김 혁 (서울시립대학교 교수) – 교육
문 상 미 (나사렛대학교 교수) – 여성/회원
배 현 철 (한국전자통신연구원 책임연구원) – 학술(추계)
변 영 재 (UNIST 교수) – ICCE Asia
안 상 철 (한국과학기술연구원 책임연구원) – 신기술
안 호 균 (한국전자통신연구원 실장) – 사업
윤 상 민 (국민대학교 교수) – 신기술
이 구 순 (파이낸셜뉴스 이사) – 홍보
이 영 택 (ASML 전무) – 교육
이 철 (동국대학교 교수) – 학회지
임 동 구 (전북대학교 교수) – 정보화
장 지 원 (연세대학교 교수) – 학회지
정 해 준 (경희대학교 교수) – 국문논문
진 성 훈 (인천대학교 교수) – 학술(하계)
최 광 성 (한국전자통신연구원 실장) – 대외협력
최 영 규 (인하대학교 교수) – 학술(하계)
최 재 혁 (서울대학교 교수) – ICCE Asia
하 정 우 (네이버 AI연구소장) – 신기술
한 동 국 (국민대학교 교수) – 사업
한 태 희 (성균관대학교 교수) – 국문논문
허 재 두 (한국전자통신연구원 연구전문위원) – 사업
황 태 호 (한국전자기술연구원 본부장) – 학술(하계)

협 동 이 사

구 민 석 (인천대학교 교수) – 기획
권 경 하 (KAIST 교수) – 사업
김 건 우 (목포대학교 교수) – 회원
김 기 현 (전북대학교 교수) – 회원
김 대 영 (순천향대학교 교수) – 회원
김 사 혁 (KISDI 책임연구원) – 대외협력
김 정 석 (가천대학교/셀리코 교수/대표이사) – 대외협력
김 형 진 (인하대학교 교수) – 회원/여성
민 경 식 (국민대학교 교수) – JSTS
배 준 성 (강원대학교 교수) – 정보화
서 민 재 (가천대학교 교수) – 홍보/정보화
서 종 열 (LG전자 그룹장) – 산학연
송 준 영 (인천대학교 교수) – 홍보
신 세 운 (POSTECH 교수) – 사업
오 윤 호 (고려대학교 교수) – 학술(추계)
유 동 훈 (디사일로 연구소장) – 산학연
윤 명 국 (이화여자대학교 교수) – 신기술
윤 희 인 (UNIST 교수) – 홍보
이 재 규 (삼성전자 마스터) – 산학연
이 정 원 (서울대학교 교수) – 회원/여성
장 지 수 (삼성전자 PE) – 사업
정 방 철 (충남대학교 교수) – 정보화
정 완 영 (KAIST 교수) – 홍보
차 철 웅 (한국전자기술연구원 센터장) – 표준화
최 강 선 (한국기술교육대학교 교수) – SPC
추 민 성 (한양대학교 교수) – 기획
홍 성 완 (서강대학교 교수) – 대외협력
권 기 룡 (부경대학교 교수) – 학술(하계)
김 민 규 (LG이노텍 상무) – 산학연
김 성 우 (서울대학교 교수) – 대외협력
김 유 철 (LG AI연구원 부문장) – 신기술
김 중 헌 (고려대학교 교수) – 사업
류 은 석 (성균관대학교 교수) – 사업
박 성 정 (건국대학교 교수) – 국제협력
백 종 덕 (연세대학교 교수) – 신기술
안 광 호 (한국전자기술연구원 본부장) – 대외협력
안 성 수 (명지전문대학 교수) – 학회지
원 용 욱 (명지대학교 교수) – 신기술
윤 종 윤 (파두 사장) – 교육
이 상 만 (고려대학교 교수) – 산학연
이 창 우 (가톨릭대학교 교수) – 국제협력
이 태 동 (국제대학교 교수) – 홍보/정보화
장 성 욱 (카카오모빌리티 부사장) – 신기술
전 동 석 (서울대학교 교수) – 학술(하계)
제 민 규 (KAIST 교수) – ICCE Asia
채 관 엽 (삼성전자 마스터) – 대외협력
최 병 호 (한국전자기술연구원 연구소장) – 산학연
최 재 용 (가천대학교 교수) – 국문논문
최 정 욱 (한양대학교 교수) – 학회지
하 태 준 (광운대학교 교수) – 학술(추계)
한 진 호 (한국전자통신연구원 책임연구원) – 국제협력
함 범 섭 (연세대학교 교수) – 학술(하계)
황 진 영 (한국항공대학교 교수) – 홍보

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

지부장 명단

강 원 지 부	강 문 식 (강릉원주대학교 교수)
대구·경북지부	이 찬 수 (영남대학교 교수)
부산·경남·울산지부	고 진 환 (경상국립대학교 교수)
제 주 지 부	고 석 준 (제주대학교 교수)
호 서 지 부	강 윤 희 (백석대학교 교수)
미 국	최 명 준 (텔레다인 박사)
광주·전남지부	최 수 일 (전남대학교 교수)
대전·충남지부	김 철 영 (충남대학교 교수)
전 북 지 부	이 주 연 (전주비전대학교 교수)
충 북 지 부	조 문 규 (한국교통대학교 교수)
일 본	강 유 선 (Tokyo Polytechnic University 교수)
러시아지부	Prof. Edis B. TEN (National University of Science and Technology)

위원회 명단

자문위원회

위 원 장	박 항 구 (소암시스텔 회장)		
부 위 원 장	박 성 한 (한양대학교 명예교수)		
위 원	고 성 제 (고려대학교 명예교수)	공 준 진 (삼성전자공과대학교 교수)	구 용 서 (단국대학교 석좌교수)
	김 성 대 (KAIST 명예교수)	김 수 중 (경북대학교 명예교수)	김 영 권 (건국대학교 명예교수)
	김 재 희 (연세대학교 명예교수)	나 정 웅 (KAIST 명예교수)	문 영 식 (한양대학교 명예교수)
	박 규 태 (연세대학교 명예교수)	박 진 옥 (육군사관학교 명예교수)	백 준 기 (중앙대학교 교수)
	서 승 우 (서울대학교 교수)	성 굉 모 (서울대학교 명예교수)	윤 종 용 (한국공학교육인증원 이사장)
	이 문 기 (연세대학교 명예교수)	이 상 설 (한양대학교 명예교수)	이 재 홍 (서울대학교 명예교수)
	이 진 구 (동국대학교 명예교수)	이 혁 재 (서울대학교 교수)	임 제 탁 (한양대학교 명예교수)
	임 혜 숙 (이화여자대학교 교수)	전 국 진 (서울대학교 명예교수)	전 홍 태 (중앙대학교 명예교수)
	정 정 화 (한양대학교 명예교수)	천 경 준 (씨젠 회장)	홍 대 식 (연세대학교 교수)
	홍 승 홍 (인하대학교 명예교수)		

기획위원회

위 원 장	강 명 곤 (한국교통대학교 교수)		
부 위 원 장	김 현 (서울과학기술대학교 교수)		
위 원	구 민 석 (인천대학교 교수)	권 구 덕 (강원대학교 교수)	권 민 우 (강릉원주대학교 교수)
	배 종 호 (국민대학교 교수)	송 익 현 (한양대학교 교수)	심 원 보 (서울과학기술대학교 교수)
	조 문 규 (한국교통대학교 교수)	채 주 형 (광운대학교 교수)	추 민 성 (한양대학교 교수)

학술연구위원회 – 하계

위 원 장	김 훈 (인천대학교 교수)		
위 원	강 문 식 (강릉원주대학교 교수)	김 근 영 (한국전자통신연구원 연구원)	김 도 현 (제주대학교 교수)
	김 영 민 (서울대학교 교수)	김 익 균 (한국전자통신연구원 본부장)	김 혁 (서울시립대학교 교수)
	이 권 형 (LG전자 팀장)	인 치 호 (세명대학교 교수)	정 성 엽 (차세대융합기술연구원 실장)
	정 승 원 (고려대학교 교수)	조 성 현 (한양대학교 교수)	진 성 훈 (인천대학교 교수)
	한 상 욱 (삼성전자 연구원)	한 재 호 (고려대학교 교수)	

학술연구위원회 – 추계

위 원 장	이 강 윤 (성균관대학교 교수)		
위 원	고 형 호 (충남대학교 교수)	권 종 원 (한국산업기술시험원 책임연구원)	김 승 환 (한국전자통신연구원 책임)
	박 관 서 (연세대학교 교수)	배 현 철 (한국전자통신연구원 책임연구원)	오 윤 호 (고려대학교 교수)
	하 태 준 (광운대학교 교수)		

논문편집위원회

위 원 장	조 성 현 (한양대학교 교수)		
위 원	강 성 복 (한국생산기술연구원 수석연구원)	강 제 원 (이화여자대학교 교수)	권 구 덕 (강원대학교 교수)
	권 종 원 (한국산업기술시험원 책임연구원)	김 기 연 (한국산업기술시험원 선임연구원)	김 명 선 (한성대학교 교수)
	김 상 범 (한국폴리텍대학 교수)	김 선 용 (건국대학교 교수)	김 영 로 (명진전문대학 교수)
	김 영 선 (대림대학교 교수)	김 학 구 (중앙대학교 교수)	심 규 성 (한경국립대학교 교수)
	어 영 정 (연세대학교 교수)	유 재 준 (UNIST 교수)	윤 중 현 (조선이공대학교 교수)
	이 덕 우 (계명대학교 교수)	이 종 호 (숭실대학교 교수)	임 민 중 (동국대학교 교수)
	조 성 인 (동국대학교 교수)	최 재 용 (가천대학교 교수)	추 민 성 (한밭대학교 교수)

국제협력위원회

위 원 장	노 원 우 (연세대학교 교수)		
위 원	김 태 욱 (연세대학교 교수)	박 성 정 (건국대학교 교수)	이 창 우 (가톨릭대학교 교수)
	한 진 호 (한국전자통신연구원 책임연구원)		

대외협력위원회

위 원 장	김 용 신 (고려대학교 교수)		
위 원	궁 재 하 (고려대학교 교수)	김 사 혁 (KISDI 책임연구원)	김 성 우 (서울대학교 교수)
	김 정 석 (가천대학교/셀리코 교수/대표이사)	안 광 호 (한국전자기술연구원 본부장)	채 관 엽 (삼성전자 마스터)

	최 광 성 (한국전자통신연구원 실장)	홍 성 완 (서강대학교 교수)	

회원관리위원회

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

회지편집위원회

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

사업위원회

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

교육연구위원회

위 원 장	변 대 석 (삼성전자 마스터)	동 성 수 (용인예술과학대학교 교수)	
위 원	김 선 욱 (고려대학교 교수)	김 혁 (서울시립대학교 교수)	류 현 석 (서울대학교 교수)
	박 영 우 (TEL 부사장)	변 영 재 (UNIST 교수)	윤 종 윤 (파두 대표이사)
	이 영 택 (ASML 전무)	이 후 진 (한성대학교 교수)	정 성 엽 (차세대융합기술연구원 실장)
	최 병 수 (한국전자통신연구원 박사)		

홍보/정보화위원회

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

표준화위원회

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

신기술위원회

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

지부담당위원회

위 원 장	이 승 호 (한밭대학교 교수)		
위 원	강 문 식 (강릉원주대학교 교수)	강 윤 희 (백석대학교 교수)	고 석 준 (제주대학교 교수)
	고 진 환 (경상국립대학교 교수)	김 철 영 (충남대학교 교수)	이 주 연 (전주비전대학교 교수)

통신소사이어티

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

반도체소사이어티

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

	최 재 혁 (서울대학교 교수)	최 재 혁 (성균관대학교 교수)	추 민 성 (한양대학교 교수)
	한 정 환 (충남대학교 교수)	황 태 호 (한국전자기술연구원 박사)	

컴퓨터소사이어티

회 장	최 용 수 (신한대학교 교수)		
명 예 회 장	강 문 식 (강릉원주대학교 교수)	김 승 천 (한성대학교 교수)	김 형 중 (고려대학교 교수)
	박 인 정 (단국대학교 명예교수)	박 준 명 (한국교통대학교 교수)	신 인 철 (단국대학교 명예교수)
	안 병 구 (홍익대학교 교수)	안 현 식 (동명대학교 교수)	이 규 대 (공주대학교 교수)
	정 용 규 (을지대학교 교수)	허 영 (스마트의료기기산업진흥재단 부이사장)	홍 유 식 (상지대학교 교수)
	황 성 운 (가천대학교 교수)		
자 문 위 원	권 호 열 (강원대학교 교수)	남 상 엽 (국제대학교 교수)	정 교 일 (한국전자통신연구원 책임연구원)
감 사	박 수 현 (국민대학교 교수)	변 영 재 (UNIST 교수)	
부 회 장	김 도 현 (제주대학교 교수)	심 정 연 (강남대학교 교수)	윤 은 준 (경일대학교 교수)
	조 민 호 (고려대학교 교수)		
협 동 부 회 장	강 상 욱 (상명대학교 교수)	김 병 서 (홍익대학교 교수)	김 영 학 (산업기술평가관리원 본부장)
	박 성 욱 (강릉원주대학교 교수)	우 운 택 (KAIST 교수)	유 성 철 (쌍용정보통신주식회사 본부장)
	이 기 영 (인천대학교 교수)	이 후 진 (한성대학교 교수)	정 은 성 (홍익대학교 교수)
	황 인 정 (명지병원 책임연구원)		
총 무 이 사	김 대 홍 (을지대학교 교수)		
재 무 이 사	박 영 훈 (숙명여자대학교 교수)		
홍 보 이 사	오 희 석 (한성대학교 교수)		
편 집 이 사	강 병 권 (순천향대학교 교수)	기 장 근 (공주대학교 교수)	김 선 욱 (고려대학교 교수)
	김 정 욱 (경희대학교 교수)	김 진 홍 (배재대학교 교수)	김 천 식 (세종대학교 교수)
	김 한 울 (서울과학기술대학교 교수)	노 소 영 (월송출판 교수)	심 규 성 (한경국립대학교 교수)
	윤 상 훈 (한국전자기술연구원 책임연구원)	이 문 구 (김포대학교 교수)	이 민 호 (경북대학교 교수)
	이 세 호 (전북대학교 교수)	이 주 형 (가천대학교 교수)	이 찬 수 (영남대학교 교수)
	정 혜 명 (김포대학교 교수)	한 규 필 (금오공과대학교 교수)	
학 술 이 사	고 한 얼 (고려대학교 교수)	김 명 선 (한성대학교 교수)	김 종 윤 (경동대학교 교수)
	김 형 진 (전북대학교 교수)	김 홍 균 (다스파워 이사)	백 광 현 (중앙대학교 교수)
	서 민 석 (고려대학교 교수)	이 덕 기 (연암공과대학교 교수)	이 정 선 (을지대학교 교수)
	이 충 규 (조선대학교 교수)	임 경 원 (대림대학교 교수)	임 재 균 (명지병원 소장)
	진 훈 (경기대학교 겸임교수)	최 한 용 (신한대학교 교수)	한 상 민 (순천향대학교 교수)
	한 영 선 (부경대학교 교수)	한 태 화 (연세대의료원 팀장)	황 석 중 (SK Telecom 박사)
	황 재 정 (군산대학교 교수)		
사 업 이 사	박 승 창 (유오씨 대표이사)	차 시 호 (청운대학교 교수)	황 진 영 (한국항공대학교 교수)
산 학 이 사	권 태 윤 (네이버시스템즈 대표이사)	김 대 휘 (한국정보통신 대표이사)	김 성 길 (우주텔레콤 이사)
	김 용 수 (비온시이노베이터 대표이사)	김 은 영 (투와이시스템즈 이사)	김 진 수 (지엔소프트 대표이사)
	김 효 선 (연세세브란스병원 연구원)	서 봉 상 (올포랜드 이사)	송 치 봉 (웨이버스 이사)
	신 동 희 (대보정보통신 부장)	오 승 훈 (주얼린 대표이사)	이 명 옥 (진우ATS 팀장)
	이 재 홍 (유비벨록스모바일 대표이사)	이 학 준 (이노지에스코리아 연구소장)	임 준 섭 (대신정보통신 차장)
	전 한 수 (세림TSG 전무)	조 병 순 (CNCinstrument 사장)	조 병 영 (태진인포텍 전무)

연구회위원장	한 태 화 (연세대의료원 팀장) – 멀티미디어	진 훈 (경기대학교 겸임교수) – 휴먼ICT
	윤 은 준 (경일대학교 교수) – 융합컴퓨팅	이 민 호 (경북대학교 교수) – 인공지능/신경망/퍼지
	김 도 현 (제주대학교 교수) – M2M/IoT	우 운 택 (KAIST 교수) – 증강휴먼
	황 성 운 (가천대학교 교수) – 인공지능 및 보안	정 은 성 (홍익대학교 교수) – 블록체인
	김 명 선 (한성대학교 교수) – AI 응용	

인공지능 신호처리소사이어티

회 장	고 병 철 (계명대학교 교수)		
자 문 위 원	김 정 태 (이화여자대학교 교수)	김 종 옥 (고려대학교 교수)	김 창 익 (KAIST 교수)
	김 홍 국 (광주과학기술원 교수)	박 종 일 (한양대학교 교수)	송 병 철 (인하대학교 교수)
	심 동 규 (광운대학교 교수)	이 영 렬 (세종대학교 교수)	전 병 우 (성균관대학교 교수)
	조 남 익 (서울대학교 교수)	홍 민 철 (숭실대학교 교수)	
부 회 장	강 석 주 (서강대학교 교수)	황 원 준 (아주대학교 교수)	김 승 룡 (고려대학교 교수)
감 사	임 성 훈 (DGIST 교수)	조 성 인 (동국대학교 교수)	
총 무	정 승 원 (고려대학교 교수)		
협 동 이 사	강 현 수 (충북대학교 교수)	곽 노 준 (서울대학교 교수)	곽 진 태 (고려대학교 교수)
	구 형 일 (아주대학교 교수)	권 구 락 (조선대학교 교수)	권 준 석 (중앙대학교 교수)
	김 광 주 (한국전자통신연구원 연구원)	김 기 백 (숭실대학교 교수)	김 동 현 (연세대학교 교수)
	김 상 효 (성균관대학교 교수)	김 선 주 (연세대학교 교수)	김 용 환 (한국전자기술연구원 수석연구원)
	김 원 하 (경희대학교 교수)	김 재 곤 (한국항공대학교 교수)	김 정 진 (계명대학교 교수)
	김 종 민 (강원대학교 교수)	김 중 욱 (항공우주연구원 연구원)	김 진 우 (경성대학교 교수)
	김 해 광 (세종대학교 교수)	남 승 우 (한국전자통신연구원 연구원)	박 구 만 (서울과학기술대학교 교수)
	박 구 만 (서울과학기술대학교 교수)	박 상 윤 (명지대학교 교수)	박 성 홍 (KAIST 교수)
	박 현 진 (성균관대학교 교수)	박 호 종 (광운대학교 교수)	배 성 호 (경희대학교 교수)
	서 영 호 (광운대학교 교수)	서 정 일 (한국전자통신연구원 교수)	서 진 근 (연세대학교 교수)
	손 광 훈 (연세대학교 교수)	송 진 호 (연세대학교 교수)	신 재 섭 (픽스트리 대표이사)
	신 지 태 (성균관대학교 교수)	심 학 준 (캐논메디칼시스템즈코리아 박사)	양 현 종 (UNIST 교수)
	어 영 정 (연세대학교 교수)	엄 일 규 (부산대학교 교수)	오 세 홍 (한국외국어대학교 교수)
	오 태 현 (POSTECH 교수)	유 명 호 (인텔리빅스 대표이사)	유 양 모 (서강대학교 교수)
	윤 국 진 (KAIST 교수)	윤 일 동 (한국외국어대학교 교수)	이 기 승 (건국대학교 교수)
	이 상 근 (중앙대학교 교수)	이 상 윤 (연세대학교 교수)	이 상 헌 (DGIST 교수)
	이 상 훈 (연세대학교 교수)	이 승 용 (POSTECH 교수)	이 연 정 (경북대학교 교수)
	이 의 진 (서울과학기술대학교 교수)	이 장 원 (한국항공대학교 교수)	이 재 성 (서울대학교 교수)
	이 종 석 (연세대학교 교수)	이 종 설 (한국전자기술연구원 책임연구원)	이 종 하 (계명대학교 교수)
	이 준 재 (계명대학교 교수)	이 준 호 (성균관대학교 교수)	이 창 우 (카톨릭대학교 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 기 완 (국가수리과학연구소 박사)	전 해 곤 (광주과학기술원 교수)	정 미 라 (대구과학관 교수)
	정 원 기 (고려대학교 교수)	정 찬 호 (한밭대학교 교수)	정 호 기 (한국교통대학교 교수)
	정 호 열 (영남대학교 교수)	정 홍 (HDXWILL 박사)	조 승 룡 (KAIST 교수)
	조 윤 지 (네이버 연구원)	진 현 성 (제주대학교 교수)	최 승 호 (서울과학기술대학교 교수)
	최 윤 식 (연세대학교 교수)	최 장 환 (이화여자대학교 교수)	최 준 원 (한양대학교 교수)
	최 현 철 (영남대학교 교수)	한 종 기 (세종대학교 교수)	허 용 석 (아주대학교 교수)
	홍 성 훈 (전남대학교 교수)	황 도 식 (연세대학교 교수)	황 성 주 (KAIST 교수)
	황 영 배 (충북대학교 교수)	강 경 진 (LG전자 연구위원)	권 기 룡 (부경대학교 교수)
	김 남 수 (서울대학교 교수)	김 진 웅 (한국전자통신연구원 그룹장)	김 창 수 (고려대학교 교수)
	백 준 기 (중앙대학교 교수)	예 종 철 (KAIST 교수)	윤 재 웅 (LG전자 연구위원)

	이병욱 (이화여자대학교 교수)	이찬수 (영남대학교 교수)	지인호 (홍익대학교 교수)
	최강선 (한국기술교육대학교 교수)	최병호 (한국전자기술연구원 연구소장)	하정우 (네이버 AI연구소장)
	한재준 (삼성전자 마스터)		
이 사	민동보 (이화여자대학교 교수)	구본학 (LG전자 연구원)	김동국 (전남대학교 교수)
	박상현 (DGIST 교수)	유윤규 (두산인프라코어 연구원)	유현우 (UNIST 교수)
	윤정인 (한화디펜스 연구원)	이금하 (KT 연구원)	정교일 (한국전자통신연구원 연구전문위원)
	차영수 (고려대학교 교수)	최성준 (고려대학교 교수)	강제원 (이화여자대학교 교수)
	고영준 (충남대학교 교수)	고종환 (성균관대학교 교수)	고현석 (한양대학교 교수)
	곽노준 (서울대학교 교수)	곽수영 (한밭대학교 교수)	곽수하 (POSTECH 교수)
	김준모 (KAIST 교수)	김진규 (고려대학교 교수)	김태현 (한양대학교 교수)
	김학구 (중앙대학교 교수)	김한울 (서울과학기술대학교 교수)	김휘용 (경희대학교 교수)
	남우정 (경북대학교 교수)	박영경 (이화여자대학교 교수)	박인규 (인하대학교 교수)
	박철수 (광운대학교 교수)	백종덕 (연세대학교 교수)	신종원 (광주과학기술원 교수)
	심재영 (UNIST 교수)	심현정 (KAIST 교수)	오병태 (한국항공대학교 교수)
	우성민 (한국기술교육대학교 교수)	우운택 (KAIST 교수)	유재준 (UNIST 교수)
	유종빈 (아주대학교 교수)	이덕우 (계명대학교 교수)	이범식 (조선대학교 교수)
	이상철 (인하대학교 교수)	이윤구 (광운대학교 교수)	이채은 (한양대학교 교수)
	이 철 (동국대학교 교수)	장준혁 (한양대학교 교수)	전세영 (서울대학교 교수)
	정영주 (숙명여자대학교 교수)	정희철 (경북대학교 교수)	조동현 (충남대학교 교수)
	조성현 (POSTECH 교수)	최동걸 (한밭대학교 교수)	최 욱 (인천대학교 교수)
	최종원 (중앙대학교 교수)	최종현 (연세대학교 교수)	최해철 (한밭대학교 교수)
	한보형 (서울대학교 교수)	한재호 (고려대학교 교수)	함범섭 (연세대학교 교수)
	홍성은 (인하대학교 교수)	황효석 (경희대학교 교수)	강수명 (계명대학교 박사)
	손진희 (광주과학기술원 교수)	박진선 (부산대학교 교수)	한재호 (고려대학교 교수)
연구회위원장	강석주 (서강대학교 교수) – 영상처리		김원준 (건국대학교 교수) – 영상이해
	이종호 (서울대학교 교수) – 바이오영상신호처리		장길진 (경북대학교 교수) – 음향 및 음성신호처리
	황원준 (아주대학교 교수) – 딥러닝		김성우 (서울대학교 교수) – 로봇지능

시스템 및 제어소사이어티

회 장	주영복 (한국기술교육대학교 교수)		
명예회장	김덕원 (연세대학교 교수)	김희식 (서울시립대학교 교수)	박종국 (경희대학교 교수)
	서일홍 (한양대학교 교수)	오상록 (KIST 분원장)	오승록 (단국대학교 교수)
	오창현 (고려대학교 교수)	유정봉 (공주대학교 교수)	정길도 (전북대학교 교수)
	허경무 (단국대학교 교수)		
부 회 장	권종원 (한국산업기술시험원 책임연구원)	이덕진 (전북대학교 교수)	
감 사	김영철 (군산대학교 교수)	김영진 (한국생산기술연구원 수석연구원)	
총무이사 (겸)	이덕진 (전북대학교 교수)		
편집/학술이사	강성복 (한국생산기술연구원 수석연구원)	김기연 (한국산업기술시험원 선임연구원)	
홍보이사	최현택 (한국해양과학기술원 책임연구원)		
산학연이사	김광식 (소올아이티 전무)	문태주 (부일하우징 대표)	신대현 (대연씨앤아이 대표이사)
	유병길 (현대선기 대표이사)		
회원이사	권오민 (충북대학교 교수)	김수찬 (한경국립대학교 교수)	김종만 (전남도립대학교 교수)
	김지홍 (전주비전대학교 교수)	김현갑 (B&P Instrument 연구소 책임연구원)	김호철 (을지대학교 교수)
	남기창 (동국대학교 교수)	류지형 (한국전자통신연구원 박사)	문정호 (강릉원주대학교 교수)
	박명진 (경희대학교 교수)	박재병 (전북대학교 교수)	변영재 (UNIST 교수)
	서영석 (영남대학교 교수)	송철규 (전북대학교 교수)	여희주 (대진대학교 교수)
	유재현 (한경국립대학교 교수)	이상준 (선문대학교 교수)	이수열 (경희대학교 교수)
	이용귀 (한국전자통신연구원 책임연구원)	이태희 (전북대학교 교수)	이학성 (세종대학교 교수)
	정재훈 (동국대학교 교수)	최수범 (KISTI 연구원)	최우영 (전북대학교 교수)
	한 아 (한국산업기술시험원 선임연구원)		
연구회위원장	김규식 (서울시립대학교 교수) – 전력전자		한수희 (POSTECH 교수) – 제어계측
	남기창 (동국대학교 교수) – 의용전자 및 생체공학		정재훈 (동국대학교 교수) – 지능로봇
	이성준 (한양대학교 교수) – 회로 및 시스템		이석재 (대구보건대학교 교수) – 국방정보및제어
	연규봉 (한국자동차연구원 센터장) – 자동차전자		오창현 (고려대학교 교수) – 의료영상시스템
	권종원 (한국산업기술시험원 책임연구원) – 스마트팩토리		정범진 (서울과학기술대학교 교수) – 스마트미터링

산업전자소사이어티

회 장	고정환 (인하공업전문대학 교수)		
명예회장	김은원 (대림대학교 교수)		
자문위원	김대휘 (한국정보기술 대표이사)	김용민 (충청대학교 교수)	원영진 (부천대학교 교수)
	윤기방 (인천대학교 명예교수)	이상회 (동서울대학교 교수)	이원석 (동양미래대학교 명예교수)
	조규남 (로봇신문사 대표이사)	한성준 (대흥정보 고문)	
수석부회장	동성수 (용인예술과학대학교 교수)		
상임이사	김남섭 (서일대학교 교수)	김상범 (한국폴리텍대학 교수)	김영로 (명지전문대학 교수)
	김 현 (부천대학교 교수)	박병섭 (인하공업전문대학 교수)	서병석 (상지대학교 교수)
	서춘원 (K-My지능정보기술 대표이사)	안성수 (명지전문대학 교수)	엄우용 (인하공업전문대학 교수)
	원우연 (한국폴리텍대학 교수)	윤중현 (조선이공대학교 교수)	이시현 (동서울대학교 교수)
	이정석 (인하공업전문대학 교수)	이태동 (국제대학교 교수)	조도현 (인하공업전문대학 교수)
	최의선 (한국폴리텍대학 교수)		
협동상임이사	강현석 (로보웰코리아 대표이사)	권오병 (진인프라 전무이사)	김세종 (SJ정보통신 이사)
	김윤철 (트라콤 상무이사)	김정석 (오디에이테크놀로지 대표이사)	김젬마 (엠티데이타 상무이사)
	김진선 (청파이엠티 대표이사)	박현영 (진우에이티에스 전무이사)	서봉상 (올포랜드 이사)
	서승현 (글로벌텔레콤 대표이사)	성재용 (오픈링크시스템 대표이사)	송관식 (아이씨티웨이 상무이사)
	송광헌 (복두출판사 대표이사)	송치봉 (웨이버스 대표이사)	유성철 (쌍용정보통신 상무이사)
	이영준 (투비콤 대표이사)	장대현 (LG U+ 상무이사)	전한수 (세림TSG 전무이사)
	조병영 (태진티엔에스 대표이사)	조한일 (하이제이컨설팅 대표이사)	최석우 (대보정보통신 상무이사)
이 사	강희훈 (여주대학교 교수)	구자일 (인하공업전문대학 교수)	권오상 (경기과학기술대학교 교수)
	권윤중 (세명대학교 교수)	김대순 (전주비전대학교 교수)	김백기 (강릉원주대학교 교수)
	김태원 (상지대학교 교수)	김필중 (조선이공대학교 교수)	성홍석 (부천대학교 교수)
	안태원 (동양미래대학교 교수)	우찬일 (서일대학교 교수)	이문구 (김포대학교 교수)
	이상철 (인천재능대학교 교수)	이성재 (대림대학교 교수)	이용구 (한림성심대학교 교수)
	이주연 (전주비전대학교 교수)	장기동 (동양미래대학교 교수)	
협동이사	고강일 (이지테크 대표이사)	곽정희 (한국정보기술 이사)	김응연 (인터그래텍 대표이사)
	김철룡 (쌍용정보통신 책임연구원)	변용필 (진인프라 전무이사)	신동희 (대보정보통신 부장)
	신우현 (시티랩스 이사)	오재곤 (한국정보기술 상무이사)	이경원 (동해종합기술공사 상무이사)
	이성대 (엔클라우드 상무이사)	이승민 (튠시스템 대표이사)	이승태 (하나텍시스템 이사)
	이용우 (쌍용정보통신 상무이사)	이지학 (송암시스콤 대표이사)	임준섭 (대신정보통신 차장)
	장기웅 (나날에스엠아이 부장)	장 철 (우송정보기술 대표이사)	정민우 (대보정보통신 상무이사)
	정용훈 (롯데정보통신 선임)	한찬석 (동해종합기술공사 부사장)	
감 사	김동식 (인하공업전문대학 교수)	김영선 (대림대학교 교수)	

제24대 평의원 명단

강 명 곤 (한국교통대학교 교수)
강 석 주 (서강대학교 교수)
강 성 원 (한국전자통신연구원 소장)
고 병 철 (계명대학교 교수)
고 정 환 (인하공업전문대학 교수)
공 배 선 (성균관대학교 교수)
구 민 석 (인천대학교 교수)
권 구 덕 (강원대학교 교수)
권 오 경 (한양대학교 석좌교수)
권 종 원 (한국산업기술시험원 책임연구원)
권 호 열 (강원대학교 교수)
김 규 식 (서울시립대학교 교수)
김 도 현 (제주대학교 교수)
김 동 식 (인하공업전문대학 교수)
김 명 준 ((전)한국전자통신연구원 원장)
김 봉 태 (한국전자통신연구원 소장)
김 선 용 (건국대학교 교수)
김 성 우 (서울대학교 교수)
김 수 연 (동국대학교 교수)
김 수 환 (서울대학교 교수)
김 시 호 (연세대학교 교수)
김 영 민 (홍익대학교 교수)
김 영 진 (한국항공대학교 교수)
김 영 한 (UC San Diego/가우스랩스 교수/대표이사)
김 원 종 (한국전자통신연구원 책임연구원)
김 윤 (서울시립대학교 교수)
김 재 준 (서울대학교 교수)
김 정 범 (강원대학교 교수)
김 종 선 (홍익대학교 교수)
김 준 모 (KAIST 교수)
김 진 상 (경희대학교 교수)
김 창 수 (고려대학교 교수)
김 철 우 (고려대학교 교수)
김 한 구 (삼성전자공과대학교 교수)
김 현 수 (삼성전자 상무)
김 형 탁 (홍익대학교 교수)
나 정 웅 (KAIST 명예교수)
남 상 욱 (서울대학교 교수)
노 정 진 (한양대학교 교수)
류 수 정 (사피온코리아 대표이사)
문 용 (숭실대학교 교수)
박 관 서 (연세대학교 교수)
박 성 욱 (차세대지능형반도체사업단 이사장)
박 성 한 (한양대학교 명예교수)
박 영 훈 (숙명여자대학교 교수)
박 종 일 (한양대학교 교수)
배 준 성 (강원대학교 교수)
백 광 현 (중앙대학교 교수)
백 준 기 (중앙대학교 교수)
변 영 재 (울산과학기술원 교수)
서 창 호 (KAIST 교수)
성 굉 모 (서울대학교 명예교수)
손 보 익
송 병 철 (인하대학교 교수)
송 준 영 (인천대학교 교수)
신 오 순 (숭실대학교 교수)
심 동 규 (광운대학교 교수)
안 상 철 (한국과학기술연구원 책임연구원)
안 호 균 (한국전자통신연구원 실장)
여 희 주 (대진대학교 교수)
오 상 록 (한국과학기술연구원 PM)
오 정 훈 (삼성전자 마스터)
우 정 호 (비젼넥스트 대표이사)
유 명 식 (숭실대학교 교수)
유 창 동 (KAIST 교수)
강 문 식 (강릉원주대학교 교수)
강 석 판 (LG전자 상무)
강 윤 희 (백석대학교 교수)
고 석 준 (제주대학교 교수)
고 진 환 (경상국립대학교 교수)
공 정 택 (성균관대학교 교수)
구 본 태 (한국전자통신연구원 본부장)
권 기 룡 (부경대학교 교수)
권 오 규 (인하대학교 교수)
권 태 수 (서울과학기술대학교 교수)
김 경 기 (대구대학교 교수)
김 달 수 (티엘아이 대표이사)
김 동 규 (한양대학교 교수)
김 동 현 (ICTK 대표이사)
김 민 휘 (중앙대학교 교수)
김 부 균 (숭실대학교 교수)
김 선 욱 (고려대학교 교수)
김 성 철 (서울대학교 교수)
김 수 중 (경북대학교 명예교수)
김 승 천 (한성대학교 교수)
김 영 권 (후레대학교 명예총장)
김 영 선 (대림대학교 교수)
김 영 진 (한국생산기술연구원 수석연구원)
김 용 석 (성균관대학교 교수)
김 원 준 (건국대학교 교수)
김 은 원 (대림대학교 교수)
김 재 현 (아주대학교 교수)
김 정 석 (가천대학교 교수)
김 종 옥 (고려대학교 교수)
김 중 헌 (고려대학교 교수)
김 진 영 (광운대학교 교수)
김 창 익 (KAIST 교수)
김 태 욱 (연세대학교 교수)
김 현 (서울과학기술대학교 교수)
김 형 준 (한국과학기술연구원 소장)
김 호 식 (SK하이닉스 담당)
남 궁 선 (유니트론텍 부회장)
남 일 구 (부산대학교 교수)
노 태 문 (한국전자통신연구원 연구전문위원)
류 현 석 (서울대학교 교수)
민 경 식 (국민대학교 교수)
박 규 태 (연세대학교 명예교수)
박 성 욱 (강릉원주대학교 교수)
박 수 현 (국민대학교 교수)
박 인 규 (인하대학교 교수)
박 진 옥 (육군사관학교 명예교수)
배 준 호 (가천대학교 교수)
백 상 헌 (고려대학교 교수)
범 진 욱 (서강대학교 교수)
서 민 재 (가천대학교 교수)
선 우 경 (서울대학교 교수)
성 원 진 (서강대학교 교수)
송 문 섭 (심텍 회장)
송 상 헌 (중앙대학교 교수)
송 철 (DGIST 교수)
신 요 안 (숭실대학교 교수)
심 용 (중앙대학교 교수)
안 승 권 (연암공과대학교 총장)
양 준 성 (연세대학교 교수)
연 규 봉 (한국자동차연구원 팀장/수석연구원)
오 윤 호 (고려대학교 교수)
우 성 민 (한국기술교육대학교 교수)
원 제 형 (도쿄일렉트론코리아 대표이사)
유 정 봉 (공주대학교 교수)
유 창 식 (삼성전자 부사장)
강 민 석 (LG이노텍 부사장(CTO))
강 석 형 (POSTECH 교수)
강 진 구 (인하대학교 교수)
고 성 제 (고려대학교 명예교수)
고 형 호 (충남대학교 교수)
공 준 진 (삼성전자공과대학교 교수)
구 용 서 (단국대학교 석좌교수)
권 영 수 (한국전자통신연구원 책임연구원)
권 종 기 (한국전자통신연구원 연구전문위원)
권 혁 인 (중앙대학교 교수)
김 광 순 (연세대학교 교수)
김 도 현 (국민대학교 명예교수)
김 동 순 (세종대학교 교수)
김 명 선 (한성대학교 교수)
김 병 서 (홍익대학교 교수)
김 상 태 (전남테크노파크 본부장)
김 성 대 (KAIST 명예교수)
김 소 영 (성균관대학교 교수)
김 수 찬 (한경국립대학교 교수)
김 승 환 (한국전자통신연구원 책임연구원)
김 영 로 (명지전문대학 교수)
김 영 재 (해동과학문화재단 이사장)
김 영 철 (군산대학교 교수)
김 용 신 (고려대학교 교수)
김 유 철 (LG AI연구원 부문장)
김 익 균 (한국전자통신연구원 본부장)
김 재 희 (연세대학교 명예교수)
김 정 호 (이화여자대학교 교수)
김 주 성 (한밭대학교 교수)
김 지 훈 (이화여자대학교 교수)
김 진 태 (건국대학교 교수)
김 철 영 (충남대학교 교수)
김 태 진 (더즈텍 대표이사)
김 현 (부천대학교 교수)
김 형 진 (인하대학교 교수)
김 훈 (인천대학교 교수)
남 기 창 (동국대학교 교수)
노 원 우 (연세대학교 교수)
동 성 수 (용인예술과학대학교 교수)
문 영 식 (한양대학교 명예교수)
민 동 보 (이화여자대학교 교수)
박 부 견 (POSTECH 교수)
박 성 정 (건국대학교 교수)
박 영 기 (싸인텔레콤 대표이사)
박 종 선 (고려대학교 교수)
박 항 구 (소암시스템 회장)
배 현 철 (한국전자통신연구원 책임연구원)
백 종 덕 (연세대학교 교수)
변 대 석 (삼성전자 마스터)
서 승 우 (서울대학교 교수)
선우명훈 (아주대학교 교수)
손 교 민 (삼성전자 마스터)
송 민 규 (동국대학교 교수)
송 익 현 (한양대학교 교수)
신 세 운 (UNIST 교수)
신 현 철 (광운대학교 교수)
심 정 연 (강남대학교 교수)
안 현 식 (동명대학교 교수)
엄 낙 웅 (한국전자통신연구원 책임연구원)
예 종 철 (KAIST 교수)
오 의 열 (LG디스플레이 연구위원)
우 운 택 (KAIST 교수)
유 동 훈 (디사일로 연구소장)
유 찬 세 (한국전자기술연구원 수석연구원)
유 호 영 (충남대학교 교수)

윤 명 국 (이화여자대학교 교수)
윤 석 현 (단국대학교 교수)
윤 종 용 (한국공학교육인증원 이사장)
이 경 중 (연세대학교 교수)
이 규 복 (한국전자기술연구원 부원장)
이 덕 진 (전북대학교 교수)
이 병 선 (김포대학교 교수)
이 상 훈 (웨이브피아 대표이사)
이 수 민 (한국센서연구소 대표이사)
이 승 호 (한밭대학교 교수)
이 왕 상 (경상국립대학교 교수)
이 인 규 (고려대학교 교수)
이 재 진 (숭실대학교 교수)
이 정 우 (중앙대학교 교수)
이 종 호 (서울대학교 교수)
이 찬 수 (영남대학교 교수)
이 천 희 ((전) 청주대학교 교수)
이 태 동 (국제대학교 교수)
이 형 민 (고려대학교 교수)
인 치 호 (세명대학교 교수)
임 제 탁 (한양대학교 명예교수)
장 석 호 (건국대학교 교수)
장 익 준 (경희대학교 교수)
전 병 우 (성균관대학교 교수)
전 영 현 (삼성전자 부회장)
정 교 일 (조선대학교 교수)
정 민 채 (세종대학교 교수)
정 성 엽 (차세대융합기술연구원 실장)
정 완 영 (KAIST 교수)
정 윤 호 (한국항공대학교 교수)
정 일 권 (한국전자통신연구원 본부장)
정 진 곤 (중앙대학교 교수)
조 남 익 (서울대학교 교수)
조 성 래 (중앙대학교 교수)
조 영 민 (SkyMirr CEO)
차 혁 규 (서울과학기술대학교 교수)
채 주 형 (광운대학교 교수)
최 강 선 (한국기술교육대학교 교수)
최 기 창 (서울대학교 교수)
최 성 민 (해치텍 대표이사)
최 승 종 (웰랑 대표이사)
최 우 영 (서울대학교 교수)
최 재 혁 (서울대학교 교수)
최 진 성 (도이치텔레콤 부사장)
최 현 택 (한국해양과학기술원 책임연구원)
하 태 준 (광운대학교 교수)
한 영 선 (부경대학교 교수)
한 정 환 (충남대학교 교수)
함 범 섭 (연세대학교 교수)
허 준 (고려대학교 교수)
홍 민 철 (숭실대학교 교수)
홍 인 기 (경희대학교 교수)
황 승 훈 (동국대학교 교수)
황 인 철 (강원대학교 교수)

윤 상 훈 (한국전자기술연구원 책임연구원)
윤 성 로 (서울대학교 교수)
윤 종 윤 (파두 사장)
이 광 엽 (서경대학교 교수)
이 남 윤 (고려대학교 교수)
이 동 규 (카카오모빌리티 부사장)
이 상 만 (고려대학교 교수)
이 석 희 (솔리다임 의장)
이 승 아 (연세대학교 교수)
이 영 렬 (세종대학교 교수)
이 용 욱 (한화시스템 부사장)
이 재 관 (한국자동차연구원 소장)
이 재 홍 (서울대학교 명예교수)
이 정 원 (서울대학교 교수)
이 주 연 (전주비전대학교 교수)
이 창 한 (한국반도체산업협회 상근부회장)
이 철 (동국대학교 교수)
이 한 림 (중앙대학교 교수)
이 흥 노 (광주과학기술원 교수)
임 매 순 (한국과학기술연구원 책임연구원)
임 혜 숙 (이화여자대학교 교수)
장 성 준 (한국전자기술연구원 팀장)
전 국 진 (서울대학교 명예교수)
전 선 익 (파이낸셜뉴스 사장)
전 정 훈 (성균관대학교 교수)
정 길 도 (전북대학교 교수)
정 방 철 (충남대학교 교수)
정 승 원 (고려대학교 교수)
정 용 규 (을지대학교 교수)
정 은 성 (홍익대학교 교수)
정 정 화 (한양대학교 석좌교수)
정 해 준 (경희대학교 교수)
조 도 현 (인하공업전문대학 교수)
조 성 재 (이화여자대학교 교수)
조 혜 정 (삼성물산 그룹장)
채 관 엽 (삼성전자 마스터)
채 찬 병 (연세대학교 교수)
최 광 성 (한국전자통신연구원 실장)
최 병 수 (부경대학교 교수)
최 수 일 (전남대학교 교수)
최 영 규 (인하대학교 교수)
최 웅 (숙명여자대학교 교수)
최 중 호 (서울시립대학교 교수)
최 창 식 (DB하이텍 부회장)
추 민 성 (한양대학교 교수)
한 동 국 (국민대학교 교수)
한 은 혜 (에스에스앤씨 대표이사)
한 진 호 (한국전자통신연구원 책임연구원)
함 철 희 (삼성전자 마스터)
홍 국 태 (엘엑스세미콘 연구위원)
홍 승 홍 (인하대학교 명예교수)
홍 철 호 (중앙대학교 교수)
황 원 준 (아주대학교 교수)
황 인 태 (전남대학교 교수)

윤 석 진 (한국과학기술연구원 원장)
윤 영 권 (삼성전자 마스터)
이 강 윤 (성균관대학교 교수)
이 규 대 (공주대학교 교수)
이 덕 기 (연암공과대학교 교수)
이 문 기 (연세대학교 명예교수)
이 상 설 (한양대학교 명예교수)
이 성 준 (한양대학교 교수)
이 승 은 (서울과학기술대학교 교수)
이 영 택 (ASML 전무)
이 윤 식 (UNIST 교수)
이 재 성 (고려대학교 교수)
이 재 훈 (유정시스템 대표이사)
이 종 호 (서울대학교 교수)
이 진 구 (동국대학교 석좌교수)
이 채 은 (한양대학교 교수)
이 충 용 (연세대학교 교수)
이 혁 재 (서울대학교 교수)
이 희 덕 (충남대학교 교수)
임 승 찬 (한경국립대학교 교수)
장 길 진 (경북대학교 교수)
장 성 진 (삼성전자 상근고문)
전 동 석 (서울대학교 교수)
전 세 영 (서울대학교 교수)
전 홍 태 (중앙대학교 명예교수)
정 민 수 (라온텍 부사장)
정 범 진 (서울과학기술대학교 교수)
정 영 모 (한성대학교 교수)
정 원 영 (강운공업 본부장)
정 은 승 (삼성전자 고문)
정 준 (쏠리드 대표이사)
제 민 규 (KAIST 교수)
조 문 규 (한국교통대학교 교수)
조 성 현 (한양대학교 교수)
진 훈 (경기대학교 교수)
채 영 철 (연세대학교 교수)
천 경 준 (씨젠 회장)
최 광 표 (삼성전자 마스터)
최 병 호 (한국전자기술연구원 연구소장)
최 승 범 (삼성전자 부사장)
최 용 수 (신한대학교 교수)
최 인 수 (삼성전자 상무)
최 지 웅 (DGIST 교수)
최 천 원 (단국대학교 교수)
하 정 우 (네이버 AI연구소장)
한 동 석 (경북대학교 교수)
한 재 호 (고려대학교 교수)
한 태 희 (성균관대학교 교수)
허 재 두 (한국전자통신연구원 연구전문위원)
홍 대 식 (연세대학교 교수)
홍 유 식 (상지대학교 교수)
황 성 운 (가천대학교 교수)
황 인 정 (명지병원 수석연구원)
황 진 영 (한국항공대학교 교수)

사무국 직원 명단

송기원 국장 – 산학연관 협력, 신규 사업, 자문/서울IT포럼, 지부, 인사, 규정, 회장단 관련, 총회 등 사무국 총괄
이안순 부장 – 학회지, 추계학술대회, 이사회/평의원회, 총무업무(선거, 공문, 임원관련, 송년회, 포상 및 Wiset 등), 산업전자소사이어티
배지영 부장 – 국제학술대회 총괄(ITC-CSCC, ICEIC, ICCE-Asia), 하계종합학술대회, 신기술(총괄), 시스템 및 제어소사이어티
배기동 부장 – 사업행사(기술워크숍 등), 국문논문, 표준화, 용역업무, 인공지능신호처리소사이어티
변은정 부장 – 본회/소사이어티/연구회 재무, 회원관리(개인회원 및 특별회원), 기획, 통신소사이어티
이소진 서기 – 국제학술대회 담당(ITC-CSCC, ICEIC, ICCE-Asia), 외국 기관과 국제협력(Joint Award 등), JSTS 및 SPC 발간, 컴퓨터소사이어티
김예빈 서기 – 정보화(홈페이지 관리 및 디지털 업무지원, 장비관리), 교육, 홍보, 신기술(담당), 기타 지원업무, 반도체소사이어티

제3차 상임이사회 회의록

제 3차 상임이사회가 4월 12일(금) 17시 학회 회의실(한국과학기술회관 1관 907호)에서 개최되었으며, 이번 회의 결과는 다음과 같다.

– 다 음 –

1. 성원 보고

– 제3차 상임이사회는 41명의 상임이사 중 26명의 참여로 성원되었음.

2. 본 학회(각 위원회) 및 소사이어티 보고

– 본 회(각 위원회)/ 각 소사이어티별 사업 및 활동 계획 · 추진경과 보고

3. 심의사항 의결

– 신규 개인회원 가입 승인에 대해 원안대로 승인함.

– 25대 직선 평의원 예비 후보 승인과 관련하여, 후보자 선정 기준에 대해 전체 상임이사에게 공유하기로 하였으며, 일부 후보(김영재, 천경준)에 대해서는 회장단 추천 간선 평의원으로 선출하기로 하고 후보 대상에서는 제외, 나머지 후보에 대해 승인함.

4. 기타

2024 정보 및 제어 심포지엄

시스템 및 제어 소사이어티(주영복 회장, 한국기술교육대) 대한전기학회와 공동 주최로 4월 22일(월) ~ 24일(수), 부여 롯데리조트에서 양일간에 걸쳐 "정보 및 제어 심포지엄"을 진행하였다. 3명의 초청강연으로 "차세대 저전력 PIM 기반 On-Device AI칩을 이용한 지능형 모빌리티 시스템 개발" 배학열 교수(전북대), "제조 패러다임 전환과 자율 제조" 박준희 박사(ETRI), "피할 수 없는 미래, 인공지능과 생성형 AI" 김준하 교수(광주과학기술원) 강연 및 논문발표에 약 50여명이 오프라인으로 참석한 가운데 성황리에 개최되었다.

개회사

기조강연

단체 기념촬영

보안 SoC(System on Chip) 설계 기술 워크숍

본 학회 사업위원회(프로그램위원장: 김익균 본부장(ETRI)) 개최로 "보안 SoC(System on Chip) 설계 기술 워크숍"이 5월 2일(목) 과학기술회관 중회의실 3에서 하이브리드 형태로 개최되었다.

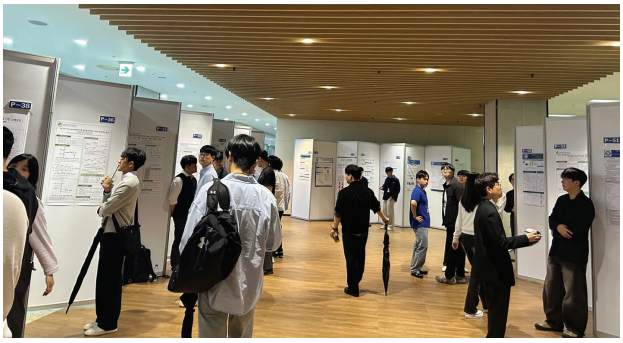
본 워크숍에서는 최신 SoC 기술 분야에서 보안의 고려사항과 고신뢰 SoC 설계 기술 주제와 더불어 RF 무선 보안 영역으로 세부 프로그램을 구성하여, 전자공학 전문가들이 정보보호 기술을 어떻게 고려하는지에 대하여 의미 있는 논의의 장이 되었다. 참석은 약 120여 명 이었다.

보안 SoC(System on Chip) 설계 기술 워크숍 (5.2)

2024 SoC 학술대회

SoC설계연구회(위원장: 이강윤 교수(성균관대))에서는 5월 11일(토) 가천대학교에서 "2024 SoC 학술대회"를 개최하였다. 이번 학술대회는 100여편의 논문이 발표되었으며 학술발표와 더불어 신진 교수 25편의 초청발표를 같이 구성하여 최신 연구 동향을 공유할 수 있는 프로그램을 구성하여 진행하였다. 이런 프로그램들이 국내 SoC 관련 연구 인력이 모여 정보 및 의견을 교환할 수 있는 장이 되었으며, SoC 연구 분야의 최신 연구 동향 및 향후 설계 연구에 도움이 되는 학술대회가 되었다. 그리고 약 300여명이 참석하였다.

2024 SoC 학술대회 행사모습



논문발표장

단체 기념촬영

신규회원 가입현황

기간 : 2024년 3월 1일 – 3월 31일

〈정회원〉

최재용(가천대학교). 김준화(건양대학교), 장민영(두산에너빌리티), 주성준(부산대학교), 김연준(성균관대학교), 서연호(전북대학교), 김춘원(한국항공우주연구원)

이상 7명

〈평생회원〉

임성훈(대구경북과학기술원), 현종일(데브릭스(주)), 안성수(명지전문대학), 서정목(연세대학교), 전재훈(인하대학교), 유광동(한경대학교), 박건식(한국전자통신연구원), 이봉민(Kyocera Korea)

이상 8명

〈학생회원〉

오주원(강남대학교), 김건회, 김윤관, 김종한, 김혜진, 송연섭, 송현민, 신유섭, 우지훈, 원지은, 유승완, 이성재, 이영규, 이영준, 이원빈, 임승환, 임원형, 장병옥, 장지성, 정채은, 조정규, 최우석, 최준혁, BASIM MUHAMMAD, Ghayur Hussain(성균관대학교), 김민우(연세대학교), 박동이, 이예지(이화여자대학교), 이진성(인천대학교), 이소혜, 이지영(한국항공대학교),

이상 31명

학회일지

www.theieie.org

THE INSTITUTE OF
ELECTRONICS AND INFORMATION
ENGINEERS

(2024년 4월 15일 ~ 5월 14일)

1. 회의 개최

회의 명칭	일시	장소	주요 안건
제4차 ICCE-Asia 조직위원회 회의	4.19 (17:00)	학회 회의실	– 학술대회 운영 및 프로그램 구성 등
제1차 자문위원회 회의	4.26 (10:30)	학회 회의실	– 전임회장 평가 및 학회 자문 등
제4차 하계 조직위원회 회의	5.10 (15:00)	학회 회의실	– 학술대회 운영 및 프로그램 구성 등

2. 행사 개최

구분	행사명	기간	장소
시스템및제어소사이어티	2024 정보 및 제어 심포지엄	4.22~24	롯데리조트 부여
산업전자소사이어티	2024 춘계학술대회 및 산업기술 교육 세미나	4.26~27	서울가든호텔
사업위원회	보안 SoC(System on Chip) 설계 기술 워크숍	5.2	한국과학기술회관 (온라인 병행)
컴퓨터소사이어티	디지털 의료 영상 기술 워크숍	5.3~4	을지인력개발원 (강원도 양양)
반도체소사이어티	2024 SoC학술대회	5.11	가천대학교

특 · 집 · 편 · 집 · 기

생성형 AI와 AI 융합 활용 기술

김 상 범 편집위원
(한국폴리텍대학)

최근 OpenAI의 ChatGPT를 포함한 생성형 AI 기반 기술을 바탕으로 대화, 번역, 요약에서부터 이미지 합성, 코드생성, 음성합성 등 다양한 분야에서 인간과 유사한 수준의 직무를 수행할 수 있게 되었다. 빅테크 회사는 경쟁적으로 생성형 AI기술을 활용한 서비스들을 B2B 비즈니스로 확장하여 모든 영역에서 인간을 대체하고 있다. 이와 아울러 하이엔드 AI 가속기 칩 발전과 AI 생태계를 통해 인공지능 개발 환경이 더욱더 용이해져서 AI를 융합한 기반 기술 활용이 더욱더 가속화될 것이다. 본 특집호는 GPU 기술, 생성형 AI 그리고 AI 융합 활용 기술에 관련된 학계 및 관련 산업계 전문가들의 기고문 6편으로 구성되었다.

첫째, "GPU를 이용한 AI 활용(김상범)"에서는 GPU 소개와 NVIDIA GPU 아키텍처 구조, CUDA 플랫폼 활용, 그리고 NVIDIA AI 생태계 활용 기술을 소개하였다. 둘째, "생성형 인공지능의 연구 동향(권혁준)"은 생성형 AI 기술 전반에 대해 소개하였으며, 특히 미드저니(Midjourney)와 같은 프롬프트 기반 이미지를 생성하는 생성적 적대 신경망인 GAN(Generative Adversarial Network)에 대한 다양한 기술을 소개하고 있다. 셋째, "ChatGPT와 거대 언어 모델(김동옥)"에서는 거대언어모델((large language model, LLM)과 Transformer 아키텍처 기반 GPT(Generative Pre-trained Transformer)언어모델을 소개하였다. 그리고 ChatGPT 서비스에서 프롬프트 활용 미세 조정 기술에 대해 살펴보았다. 넷째, "실물 환경에서의 데이터 기반 강화학습 적용 방안(강찬석)"에서는 온라인 및 오프라인 강화학습 기술을 소개하고 오프라인 강화학습에서 제한된 데이터세트 활용 한계를 해결하기 위한 온라인 미세 조정(fine-tuning) 기술에 대해 다루었다. 실제 세탁기에 Offline-to-Online 강화학습을 적용하여 세탁기 전문가 모델보다 훨씬 높은 개선효과를 얻었다. 다섯째, "반도체 공정장비에서 인공지능 기술의 역할과 활용(김홍연)"에서는 반도체 공정장비에 대해 고장진단 및 예지, 웨이퍼 품질 추정, 자동공정제어, 그리고 제조 스케줄링 최적화에 적용된 AI 융합 기술들을 소개하였다. 여섯째, "인공지능을 활용한 기계의 건전성 예측 및 관리 방법 (김관필 외1)"에서는 인공지능을 활용한 PHM 기술에 대한 연구동향에 대해 소개하였다.

바쁜 일정 중에 본 특집호를 위하여 원고를 보내주신 집필진 여러분께 감사드리며, 본 특집호가 생성형 AI와 AI융합 활용 기술 소개를 통해서 우리나라 AI산업의 발전과 경쟁력 강화에 기여할 수 있기를 기원한다.

GPU를 이용한 AI 활용

Ⅰ. 서 론

최근 OpenAI의 ChatGPT를 포함한 생성형 AI 기술의 발전으로 인간의 창조적 능력을 대신할 수 있을 만큼 풍부하고 다양한 독창적인 콘텐츠를 창작할 수 있는 새로운 패러다임을 등장하게 하였다. 이러한 생성형 AI는 거대언어모델 (LLM)로서 트랜스포머(Transformer)를 기반으로 방대한 양의 텍스트 데이터를 사전 학습하여 대화, 번역, 요약을 할 수 있게 되었다. 또한 코드생성, 이미지생성, 음성합성 등 다양한 분야에 맞게 미세조정(fine-tuning)을 통해서 정교하게 적용할 수 있게 되었다. 이러한 기술 발전의 원동력은 AI 가속기인 NVIDIA GPU의 고성능 컴퓨팅 기술 때문에 가능하게 된 것이다. 현재 NVIDIA의 가속 컴퓨팅 기술을 이용하여 데이터 사이언스, 생성형 AI, 의료, 헬스, 로보틱스, 자율주행 분야까지 AI 기술과 산업 발전에 큰 영향을 미치고 있다. 이는 NVIDIA CUDA 플랫폼기반으로 하는 병렬 컴퓨팅 성능을 통해서 딥러닝, 데이터센터의 비지니스 모델, 연구, 슈퍼컴퓨팅, 게이밍분야 발전에 핵심적인 기술을 제공하고 있다. 따라서 이번 특집기는 NVIDIA GPU소개 및 관련 아키텍처와 CUDA 통합 플랫폼 활용 그리고 GPU 기반 AI 활용에 대해 알아본다.

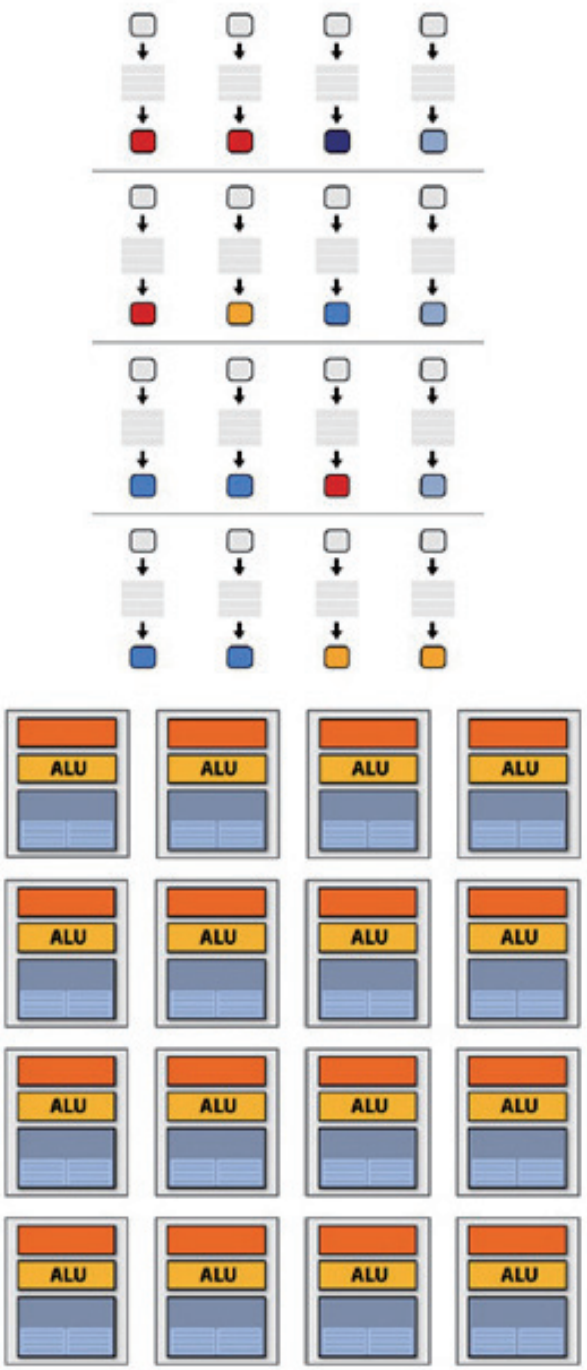
Ⅱ. GPU 소개

김상범
한국폴리텍대학

1. GPU 개요

GPU는 그래픽 처리 장치(Grapic Processing Unit)로서 컴퓨터시스템에서 컴퓨터그래픽과 영상처리 연산을 빠르게 처리하여 출력장치에 출력하는 연산장치이다. 초기에는 게임을 위한 3D 그래픽 렌더링 연산을 목적으로 만들어졌지만, 최근에는 빠른 병렬연산처리를 활용

한 GPU의 범용화(General Purpose GPU)로서 빅데이터, 인공지능, 엣지 컴퓨터, 자율주행 자동차 구현 등에 활용이 되고 있다. 최근에는 CPU와 GPU를 통합한 임베디드 시스템을 활용하여 게임기, 모바일 단말기에서부터 임베디드 IoT, 로보틱스, 그리고 소형 AI 시스템에 이용되고 있다. 일반적으로 CPU는 코어 수가 적지만 개별적인 작업과 신속한 작업 처리 작업에 최적화 되어 있어 연속적인 작업이나 특히 대기 시간이나 코어당 성능이 중요한 워크로드에 적합하다. 반면에 GPU는 3D 렌더링 작업 속도를 단축하기 위해 여러 개의 코어를 구성하여 병렬처리 할 수 있게 설계하였다. 그리고 GPU 코어를 활용해서 범용적인 병렬처리 프로그램을 할 수 있게 발전이 되어 딥러닝 및 AI 학습 응용 프로그램을 처리 할 수 있게 되었다. 최근에는 거대언어모델 (LLM)에 대한 딥러닝 학습에 최적화된 전용 AI 가속기 프레임워크를 통해 성능이 크게 향상되어 학습시간이 획기적으로 단축되었다. GPU는 그래픽처리를 빠르게 하기 위해 부동소수점명령과 픽셀 셰이더, 버텍스 셰이더, 슈퍼샘플링, 색공간 변환 등의 랜더링 전용 명령을 지원한다. 그리고 빠른 부동소수점 연산과 행렬, 벡터연산속도를 활용하여 3D 그래픽 이외의 컴퓨팅분야에 활용할 수 있다. 대표적인 GPU 개발 회사는 NVIDIA, AMD가 있다. 특히 NVIDIA는 3D 렌드링 기술이 뛰어난 회사로 GTX, GT계열의 그래픽카드 제품에서부터 AI 가속기 전용 GPU인 A100, H100, B200 계열을 제공하고 있다. 그리고 NVIDIA는 AI 가속기를 상호연결하여 고성능 컴퓨터, 클라우드, 데이터센터, 그리고 슈퍼컴퓨터 환경을 제공하면서 NVIDIA AI Enterprise라는 AI 지원 통합 플랫폼을 제공하고 있다. 이를 통해 GPU에서부터 AI 플랫폼까지 독보적인 생태계를 구축하고 있다.

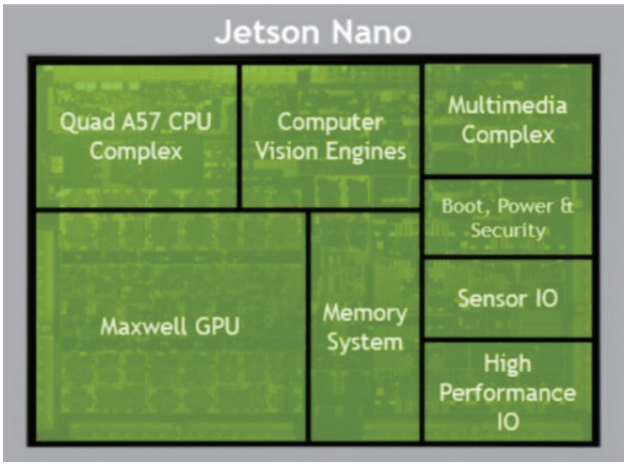


〈그림 1〉 GPU 병렬연산 처리

2. GPGPU

GPGPU(general purpose computing on GPU)는 컴퓨터 그래픽 계산 전용 GPU를 일반적인 응용프로그램에 활용하는 기술을 의미한다. 일반 CPU에 비해 GPU는 그래픽 전용 파이프라인 구현으로 많은 코어 수를 가지고 있기 때문에 병렬처리에 적합하다. 그래픽 처리 전용 코어는 부동소수점 연산, 행렬, 그리고 벡터 연산을 빠르게 처리할 수 있다. 그리고 GPU는 CPU와 달리 ALU 연산 Core가 여러 개 가지고 있어 동시에 병렬처리 가능하다. 다음은 16개 코어가 동시에 병렬 연산처리 수행하여 CPU보다 빠른 연산 속도를 구현한다.

GPU를 활용해서 응용 프로그램을 실행하기 위해서는 CUDA, OpenCL과 같은 라이브러리가 필요하다. CUDA는 NVIDIA GPU를 범용 컴퓨팅 목적으로 이용하게 하는 병렬 컴퓨터 플랫폼이다. 반면에 OpenCL은 CPU(IntelArm), GPU(AMD, NVIDA), DSP와 같은 다양한 프로세스에서 같이 사용할 수 있는 라이브러리이다. 이들 GPU 전용 라이브러리의 역할은 실행 데이터를 GPU 메모리에 복사를 하고. GPU에 실행 할 단일 커널(함수) 프로그램을 생성한다. 그리고 단일 커널 프로그램을 여러 코어에서 독립적으로 병렬 실행을 할 수 있게 처리한다.



〈그림 2〉 Jeston nano AI computer module

3. GPU 임베디드 시스템

최근에는 서로 다른 프로세스를 통합해서 컴퓨팅 파워를 같이 이용하여 처리 속도를 높이는 이기종(heterogeneous computing) 컴퓨팅 시스템을 많이 활용하고 있다. 특히 모바일 환경의 발전과 AI 활용 앱이 많아짐으로 해서 CPU와 GPU를 통합한 GPU 임베디드 시스템이 다양하게 개발되어지고 활용되고 있다. 최근 삼성에서 출시한 Exynos 2400은 데카코어(10코어)와 NPU(Neural Processing Unit)으로 구성되어 있으며, 이전 Exynos 2200에 비해 AI 처리가 14.7배 향상되어 온디바이스 AI를 실행한다. 그밖에 GPU 임베디드 칩으로 NVIDIA Jetson 시리즈가 있다. Jetson 계열은 GPU, CPU, 메모리, 전력 관리, 고속 인터페이스 등이 포함된 시스템 온 칩(System on chip)으로 Jetson AGX Orin, Jetson Orin NX, Jetson AGX Xavier, Jetson Xavier NX, Jetson TX2, 그리고 Jetson Nano가 있다. Jetson Nano는 강력한 소형 임베디드 컴퓨터로, 보급형 엣지 AI 애플리케이션 및 디바이스를 지원하도록 설계되었다. Jetson Nano는 Maxwell GPU 아키텍처 기반의 Tegra X1 모바일 프로세서로 구성되어 있다. Maxwell GPU는 그래픽 처리 클러스터(GPC), 스트리밍 멀티프로세서(SM), 메모리 컨트롤러로 구성되어 있다. SM은 128개의 CUDA 텐서 코어를 가지고 있으며, 16비트 부동소수점 계산을 지원하기 때문에 컴퓨터 비전 기반 자동차 및 임베디드 애플리케이션에 적은 비용으로 고성능 AI 컴퓨팅을 얻을 수 있다.

그리고 딥 러닝, 컴퓨터 비전, 그래픽, 멀티미디어 등을 위한 가속 라이브러리가 포함되어 있어 Jetson Nano 개발자 키트를 사용하여 빠른 프로토타입 제작을 가능하게 한다. Jetson Nano 개발자 키트는 이미지 분류, 객체 감지, 분할 및 음성 처리와 같은 애플리케이션에서 딥러닝을 병렬처리하는 AI 임베디드 시스템이다. Jetson Nano는 NVIDIA GPU 시리즈에 사용되는 것과 동일한 CUDA-X™ 소프트웨어를 포함하는 NVIDIA JetPack™ 플랫폼을 지원한다. JetPack은 애플리케이션 개발 및 최적화를 위한 최신 NVIDIA 도구를 포함하고 있다.

Ⅲ. NVIDIA GPU 아키텍처

1. Ampere(A100) 아키텍처

Ampere 아키텍처는 단일 및 다중 GPU 워크스테이션, 서버, 클러스터, 클라우드 데이터 센터, 에지 시스템, 슈퍼컴퓨터에서 실행되는 GPU 컴퓨팅 및 딥러닝 애플리케이션을 위해 수십, 수백, 또는 수천개의 GPU로 확장할 수 있는 매우 강력한 확장성을 제공한다. Ampere 아키텍처 기반 GA100 GPU로 구동되는 A100은 이전 세대 기반 기술(Tesla V100 GPU)을 기반으로 해서 새로운 기능을 추가하였으며, HPC, AI 및 데이터 분석 업무에 훨씬 더 빠른 성능을 제공한다. TSMC의 7nm N7 제조

공정에서 제작된 A100은 다이 크기가 826mm2인 542억 개의 고밀도 트랜지스터가 포함되어 있으며, 이전 세대 GPU보다 훨씬 뛰어난 와트당 향상된 성능 및 더 나은 전력 효율성을 제공한다. 그리고 A100 GPU는 탄력적이고 다재다능하며 처리량이 높은 데이터 센터를 구축할 수 있다. A100 GPU는 다중 인스턴스 GPU(MIG) 가상화 및 GPU 파티셔닝 기능을 제공하여, 최대 7배의 GPU 인스턴스를 생성할 수 있으며, GPU를 안전하고 결함 없이 분할할 수 있다. A100 GPU는 NVLink, NVSwitch 및 InfiniBand를 사용하여 수천개 이상의 GPU로 확장하거나 MIG를 사용하여 여러 독립 사용자를 지원하도록 확장하여 뛰어난 성능과 높은 GPU 사용 효율을 제공한다. 그리고 A100은 이전 세대 보다 처리량을 높인 3세대 Tensor Core를 추가하였고, 딥러닝 및 HPC 처리 전용 데이터 타입 및 처리량을 두 배로 늘리는 희소성처리 기능을 추가하였다. 여기서 희소성은 데이터의 대부분이 0으로 구성된 데이터를 의미한다. 딥러닝 처리 시, 희소성 데이터가 많이 존재하기 때문에 데이터 차원을 줄여 처리하게 되면 메모리를 줄일 수 있고 처리속도를 높일 수 있다. A100 GPU는 다중 GPU 처리 클러스터(GPC), 텍스처 처리 클러스터(TPC), 스트리밍 멀티프로세서(SM) 및 HBM2 메모리 컨트롤러로 구성되어 있다. 여기서 GPC는 7개, TPC/GPC는 8개, SM/TPC는 2개, SM/GPC는 16개, 그리고 SM은 108개로 구성되어 있다. 스트리밍 멀티프로세서(SM)는 CUDA 커널을 실행하는 프로세서이다. SM은 Control logic과 CUDA core를 가지고 있으며, 실행 스레드 간에 공유하는 특수 레지스터, 캐시, 공유메모리를 포함한다. SM은 4개의 서브코어로 나뉘어져 있으며, 각각의 서브코어는 16개의 INT32, FP32 코어, 8개의 FP64 코어, 그리고 1개의 Tensor 코어로 구성되어 있다. 그리고 SM내의 L1 cache의 크기는 192KB로 설계되어 있다. 따라서 SM은 64개의 FP32 CUDA 코어를 가지고 있으므로 GPU는 총 6912개의 FP32 CUDA 코어를 가진다. 그리고 SM은 4개의 3세대 Tensor 코어를 가지고 있기 때문에 GPU당 432개의 3세대 Tensor 코어를 가진다. 그리고 5개의 HBM2 스택과 10개의 512비트 메모리 컨트롤러로 구성되어있다

〈그림 3〉 GA100 Streaming Multiprocessor (SM)

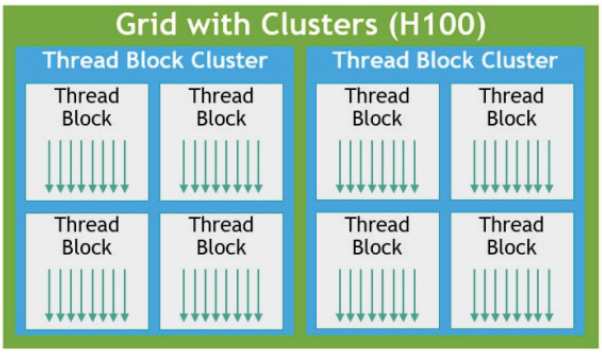
여기서 Tensor 코어는 AI 및 HPC 애플리케이션에 획기적인 성능을 제공하는 행렬 수학 연산을 위한 특수 고성능 컴퓨팅 코어이다. 텐서 코어는 MMA(행렬 곱셈 및 누적) 계산을 수행한다. MMA 연산은 신경망 훈련 및 추론에 활용되며, 딥러닝 계층에서 입력 데이터 및 가중치의 행렬을 곱하는 데 사용된다. A100의 새로운 3세대 Tensor Core 아키텍처는 V100에 비해 SM당 Tensor 처리량이 두 배 높으며, FP16, BF16, TF32, FP64, INT8, INT4 및 Binary를 포함한 모든 데이터 유형을 가속화한다. 특히 희소 행렬 계산을 위한 가속화 연산을 통해 2배 이상의 속도 향상을 제공한다. TF32는 딥러닝 프레임워크 및 HPC에서 FP32 입력/출력 데이터를 가속화 하여, 이전 세대보다 10배 이상 빠르게 실행되며, 희소성이 있는 경우 20배 더 빠르게 실행한다.

2. Hopper(H100) 아키텍처

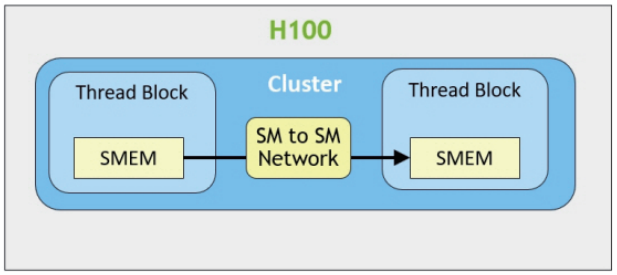
Hopper GPU 아키텍처를 기반으로 하는 NVIDIA H100 은 새로운 4세대 Tensor 코어를 채용하여, A100에 대해 2배 빠른 연산 계산을 수행할 수 있으며, FP8 도입으로 부동 소수점 연산 성능은 A100 보다 4배 빠르다.

새로운 트랜스포머 엔진을 통해 H100은 이전 세대 A100에 비해 거대언어 모델(LLM)에서 최대 9배 더 빠른 AI 훈련과 최대 30배 더 빠른 AI 추론 속도를 제공한다. 새로운 NVLink 네트워크 인터커넥트는 외부 컴퓨팅 노드에서 최대 256개의 GPU 간 통신이 가능하다. 그리고 보안 MIG는 GPU를 독립적인 다수개의 워크로드 크기로 격리된 인스턴스로 분할해서 실행할 수 있다. Hopper의 새로운 DPX 명령어는 유전체 및 단백질 서열 분석을 위한 Smith-Waterman 알고리즘 처리를 최대 7배 빠르게 할 수 있다. 4세대 텐서 코어, 텐서 메모리 가속기 및 SM과 H100 아키텍처 개선 사항을 통해 최대 3배 더 빠른 HPC 및 AI 성능을 제공한다. NVIDIA H100 GPU는 8개의 GPC와 72개의 TPC (9 TPC/GPC) 그리고 2개의 SM/TPC로 구성되어 있어 전체 GPU당 144개의 SM을 가지고 있다. SM은 128개의 FP32 CUDA 코어를 가지고 있어 전체 GPU당 18432개의 FP32 CUDA 코어를 가진다. SM은 4개의 4세대 텐서 코어를 가지고 있기 때문에 전체 GPU당 576개의 텐서 코어를 가진다. H100은 6개의 HBM3 또는 HBM2e 스택으로 구성되어 있고, 12개의 512비트 메모리 컨트롤러와 60MB의 L2 캐시, 그리고 4세대 NVLink 및 PCIe Gen 5 유닛이 장착되어 있다. NVIDIA의 H100은 최초로 완전 비동기 GPU로서 비동기 전송을 모든 주소 공간으로 확장하고, 텐서 메모리 액세스 패턴을 지원한다. 새로운 텐서 메모리 가속기를 사용하여 H100의 전체 메모리 대역폭을 관리하기 때문에 CUDA 스레드는 데이터 처리에만 집중할 수 있다.

H100은 Thread Block Cluster을 이용해서 CUDA 스레드 그룹 계층 구조를 확장하여, 동시에 SM들을 통해



〈그림 4〉 클러스터 계층구조 그리드(H100)



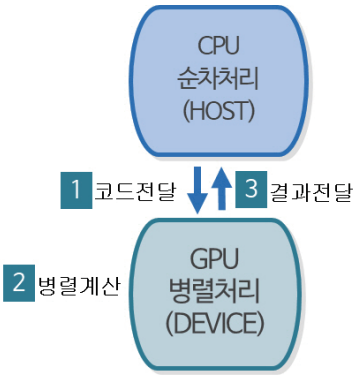
〈그림 5〉 스레드 블록 간 데이터 교환(H100)

스레드 간의 효율적인 협력과 데이터 공유를 가능하게 한다. 클러스터는 또한 텐서 메모리 가속기 및 텐서 코어를 비동기적으로 구동할 수 있다. 그리고 NVIDIA의 비동기 트랜잭션 장벽은 클러스터 내의 CUDA 스레드 및 칩 내 가속기가 효율적으로 동기화되도록 가능하게 해서 응용 프로그램이 언제든지 모든 유닛에서 실행 할 수 있게 한다. 클러스터는 모든 스레드가 다른 SM의 공유 메모리에 직접적으로 로드, 저장 및 원자 연산을 수행할 수 있다. 이 기능은 분산 공유 메모리 (DSMEM)라고 한다. 이는 공유 메모리의 가상 주소 공간이 클러스터 내의 모든 블록에 논리적으로 분배되어 있기 때문이다. 따라서 DSMEM은 SM 간의 더 효율적인 데이터 교환을 가능하게 한다. 반면에 A100은 데이터를 전달하기 위해 데이터를 글로벌 메모리에 쓰고 읽어야 했다. 글로벌 메모리 사용 대비 DSMEM은 스레드 블록 간 데이터 교환을 약 7배 정도 빠르게 한다.

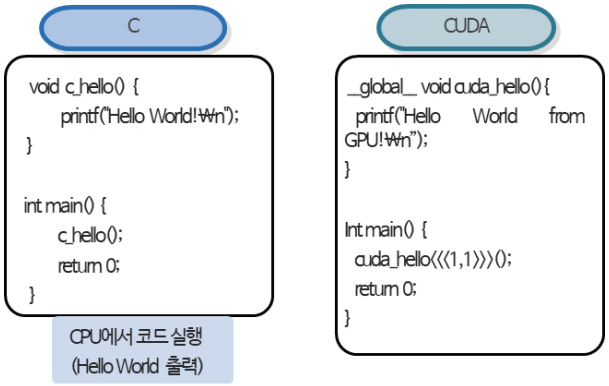
Ⅳ. CUDA 플랫폼

1. CUDA 소개

CUDA(Compute Unified Device Architecture)는 NVIDIA에서 개발한 그래픽 처리 장치(GPU)를 이용해서 병렬처리 프로그램을 할 수 있게 개발한 소프트웨어 아키텍처이다. CPU는 연산속도가 빠르지만 CPU가 가지고 있는 코어 수는 많지 않다. 반면 GPU는 계층적으로 구성된 수천 개의 연산 코어를 가지고 있다. CUDA를 활용해서 단순 집약 연산인 경우는 수천 개의 GPU 코어에서 병렬처리하고, 재귀연산과 같은 직렬 연산처리는 빠른 성능을 가지는 CPU에게 처리할 수 있게 한다. CUDA는 GPU



〈그림 6〉 CUDA 프로그래밍 프로세스



〈그림 7〉 CUDA 프로그래밍 예

에서 수행하는 병렬 처리 알고리즘을 C/C++, 파이썬, 매트랩 등을 이용해 프로그램 할 수 있게 한다. 그리고 CUDA 기반 딥러닝 라이브러리인 cuDNN, Convolution 연산을 더 빠르게 만들어주는 cuFFT, 딥러닝 추론 모듈인 tensorRT, 그리고 선형대수 모듈인 cuBLAS가 구현되어 있다. 따라서 그래픽 API를 알지 못하는 개발자들도 GPU를 활용할 수 있게 되었다. 텐스플로우, Pytorch, Mathlab 등 딥러닝 플랫폼들이 CUDA를 지원하고 있기 때문에 GPU 기반 딥러닝 프로그램이 가능하게 되었다. CUDA 프로그램은 호스트 컴퓨터 메모리에 있는 실행 코드를 DEVICE(GPU) 메모리에 전달 한 후 GPU 코어에서 스레드 단위로 병렬처리 연산을 수행한다. 그리고 DEVICE 메모리에 있는 실행결과를 HOST 메모리에 전달을 한다.

2. CUDA 프로그램

일반적인 C언어 코딩 방식은 메인함수에서 서브함수들을 호출해서 실행을 한다. 실행파일은 호스트 컴퓨터 램에 적재가 된다. 그리고 CPU 코어가 실행코드를 읽어서 연산을 하고 실행 결과를 출력한다. 반면에 CUDA 프로그램은 GPU 코어가 실행코드를 읽어서 연산을 하고 실행결과를 다시 호스트 CPU와 램에 전달한다. 따라서 CUDA 프로그램은 데이터와 실행코드를 GPU에 전달하는 과정과 GPU에서 연산한 결과 값을 호스트 PC 메모리에 전달하는 과정이 필요하다. 다음은 일반 C언어 코딩 형식과 CUDA 프로그램 형식을 비교한 예이다. 일반 C 언어에서는 메인함수에서 c_hello() 서브함수를 호출하여 호스트 PC에서 실행한다. CUDA 프로그램은 GPU에서 실행할 함수 코드를 __global__ 키워드를 사용해서 정의한다. 이 함수는 GPU에서 실행되고 전역적으로 호출될 수 있음을 나타낸다. 그리고 메인함수에서 GPU에 실행할 함수를 다음과 같은 형식으로 cuda_hello<<<1, 1>>>(); 호출한다. 여기서 cuda_hello는 앞에서 정의한 __global__ void cuda_hello() GPU 실행 함수이다. GPU에서 실행하는 함수를 커널이라고 부른다. 그리고 <<<1,1>>>는 커널을 실행할 GPU 코어에서 실행할 스레드 개수를 정의하는 문장이다. 첫 번째 매개변수는 GPU 코어에서 실행할 스레드 계층을 지정하는 값으로 스레드 블록(그룹)의 수를 정의한다. 두 번째는 블록에서 실행할 스레드 수를 지정한다. 여기서는 1개의 스레드 블록에 있는 1개의 스레드를 이용해 커널을 실행한다는 의미이다. 일반적으로 CPU에서 실행되는 코드를 호스트 코드라고 하고 GPU에서 실행되는 코드를 장치 코드라고 한다.

GPU에서 실행되는 커널은 비동기적으로 실행된다. 따라서 호스트(CPU) 실행 프로그램이 커널 실행이 완료될 때까지 기다리지 않고 계속 실행이 된다. 이때 CUDA 런타임에서 제공하는 cudaDeviceSynchronize를 호출하여 호스트(CPU) 실행프로그램을 장치(GPU) 커널 프로그램이 완료될 때까지 기다렸다가 실행 할 수 있게 한다.

가) CUDA 스레드 계층 구조

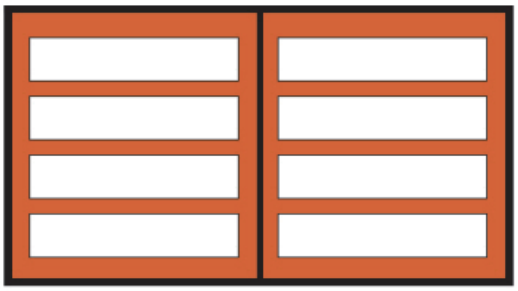
GPU에서 실행되는 커널은 GPU 코어에서 스레드에 할당되어 병렬처리 실행된다. 실행 구성을 통해서 스레드 실행 계층 구조를 설계한다. 스레드 계층 구조의 상위는

그리드(grid)이다. 그리드 하위는 스레드의 그룹인 블록으로 구성되어 있다. 블록의 하위는 스레드로 구성되어 있다. 실행설정은 아래 형식과 같이 numBlocks에 블록수를 설정하고, threadfPerBlock에 블록별 스레드 개수를 설정한다. 그리고 그리드 안에 있는 구성된 블록들은 같은 수의 스레드를 가진다. 커널 프로그램은 실행 설정에 의해서 구성된 블록들의 스레드에 할당되어 병렬 처리된다.

```
<<< numBlocks, threadfPerBlock>>>
```

메인함수에서 cuda_hello<<<2, 4>>>(); 함수를 호출하게 되면, 실행 설정 정의대로 블록수 2개, 블록별 스레드 4개가 그리드에 구성이 된다. 그리고 커널은 2개의 블록에 각각 할당된 4개의 스레드(총8개)에서 동시에 실행이 된다.

다음 테이블은 실행 설정 정의별 블록수, 블록당 스레드 수, 그리고 커널 실행 횟수를 비교하였다. <<<1, 1>>>() 실행 설정은 한 개 블록에 한 개 스레드가 실행되도록 구성되어 있으므로 커널이 한 번만 실행이 된다. <<<1, 5>>>() 실행 설정은 한 개 블록에 5개 스레드가 실행되도록 구성되어 있으므로 커널이 5번 실행이 된다. <<<5, 1>>>() 실행 설정은 단일 스레드를 가지는 블록이 5개 실행되도록 구성되어 있으므로 커널이 5번 실행이 된다. <<<5, 5>>>() 실행 설정은 5개 스레드를 가지는 블록이 5개 실행되도록 구성되어 있으므로 커널이 25번 실행이 된다.



<그림 8>cuda_hello<<<2, 4>>>()

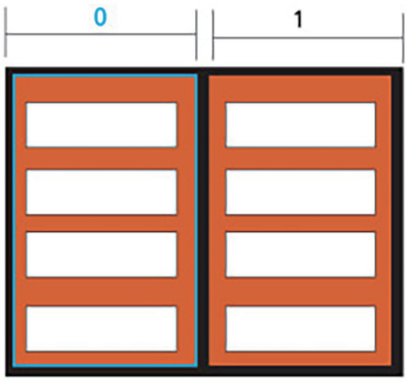
<표 1> 실행 설정별 블록수, 블록당 스레드 수, 그리고 커널 실행 횟수 비교

실행 설정	블록수	블록당 스레드 수	커널 실행 횟수
<<<1, 1>>>();	1	1	1
<<<1, 5>>>();	1	5	5
<<<5, 1>>>();	5	1	5
<<<5, 5>>>();	5	5	25

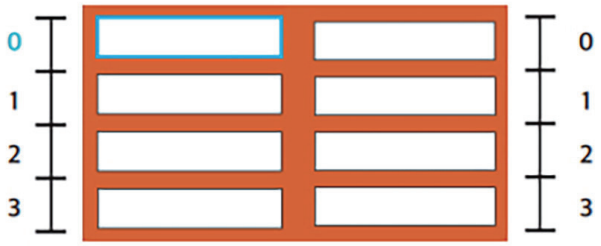
나) CUDA 스레드 계층 구조 변수

CUDA는 스레드 계층 상위인 그리드에 포함되어 있는 블록들에게 인덱스를 부여한다. 인덱스 번호는 0부터 시작한다. 마찬가지로 블록에 포함된 스레드는 0부터 시작하는 인덱스를 자동 부여받는다. 그리고 CUDA는 스레드의 인덱스 번호를 참고할 수 있는 특별한 변수를 제공한다. 이는 스레드 계층 구조변수로서 threadIdx.x 와 blockIdx.x가 있다. CUDA 커널은 커널을 실행하는 스레드의 블록을 참고하기 위해 blockIdx.x 변수를 이용한다. 그리고 각각의 블록에 구성되어있는 스레드를 참고하기 위해 threadIdx.x 변수를 이용한다. cuda_hello<<<2, 4>>>() 함수 실행에서 그리드는 2개의 블록을 가진다. 이때 각각의 블록을 참고하기 위해 인덱스를 부여한다. 인덱스 값은 0부터 시작이 된다. 여기서는 블록이 2개이므로 0과 1로 인덱스가 부여된다. CUDA 커널은 커널을 실행하는 스레드의 블록을 접근하기 위해 blockIdx.x 변수를 이용한다.

그리고 각각의 블록에 구성되어있는 스레드를 참고하기 위해 인덱스를 부여한다. 인덱스 값은 0부터 시작이 된다. 여기서는 각 블록별 4개의 스레드로 구성



<그림 9>blockIdx.x(0,1)



〈그림 10〉threadIdx.x(0,1,2,3)

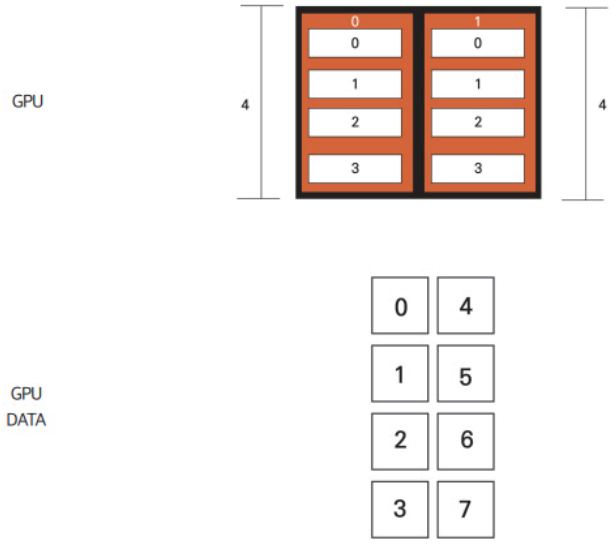
되어 있다. 따라서 인덱스는 0,1,2,3으로 부여된다. CUDA 커널은 커널을 실행하는 블록의 스레드를 참고하기 위해 threadIdx.x 변수를 이용한다. 다음은 printThreadConfig<<<2, 4>>>() 실행 설정을 통해서 4개의 스레드를 가지는 2개 스레드 블록을 구성한다.

GPU 병렬처리 개수를 늘이게 되면 실행설정을 통해서 블록과 스레드를 병렬처리 개수만큼 구성을 해야 한다. 이때 블록 당 실행 할 수 있는 스레드 수는 최대 1024개로 한정되어 있다. 따라서 실행 커널의 병렬 처리 개수를 늘릴 때 최대 1024 스레드 개수제한을 감안해서 블록 개수를 설정해야 한다. CUDA는 블록에 구성된 총 스레드 개수를 참고할 수 있는 특수 변수를 제공한다. 이 특수 변수는 blockDim.x이다. 그리고 앞에서 언급한 blockIdx.x 변수와 hreadIdx.x를 함께 사용한 아래 식을 이용해서 그리드에 할당된 블록의 스레드에 대해 인덱스 번호를 부여해서 참고 할 수 있다.

스레드 인덱스 : threadIdx.x + blockIdx.x * blockDim.x

다음은 cuda_hello<<<2, 4>>>() 실행 설정을 통해서 4개의 스레드를 가지는 2개 스레드 블록을 구성하였다. 따라서 그리드는 4개의 스레드로 구성된 2개의 블록을 포함해서 총 8개의 스레드를 가진다. 8개의 스레드는 스레드 인덱스 수식을 통해서 인덱스 번호를 부여받고 참고할 수 있다. 여기서 블록의 스레드 개수 변수인 blockDim.x는 4가 된다. 왜냐하면 실행설정에서 스레드 수를 4개로 구성하였다.

실행설정에서 구성한 그리드의 스레드보다 데이터의 크기가 많을 경우는 동시에 병렬처리를 할 수가 없다. 이 경우 커널실행은 그리드에서 구성한 스레드만큼 데이터



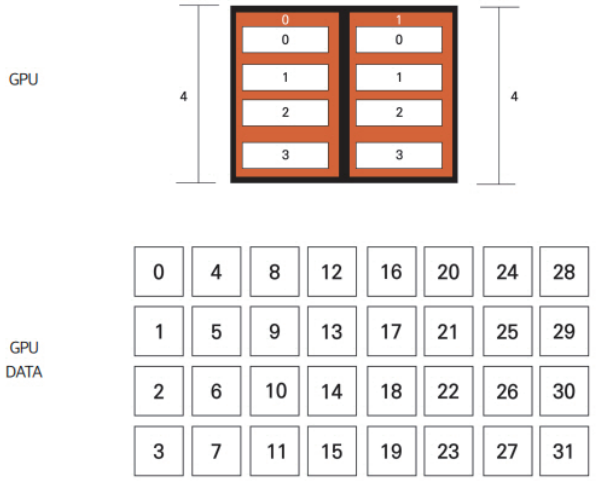
〈그림 11〉병렬스레드 인덱스: threadIdx.x + blockIdx.x * blockDim.x

〈표 2〉 병렬 커널 처리 스레드 인덱스

threadIdx.x	blockIdx.x	blockDim.x	스레드 인덱스
0	0	4	0
1	0	4	1
2	0	4	2
3	0	4	3
0	1	4	4
1	1	4	5
2	1	4	6
3	1	4	7

를 읽어서 처리한다. 이어서 그리드의 스레드 변위만큼 증가한 다음 데이터를 읽어서 스레드를 실행한다. 반복해서 데이터를 다 읽을 때까지 반복 수행한다. 이때 그리드의 스레드 변위만큼 증가해서 데이터를 읽을 수 있게 반복 처리하는 반복문을 grid-stride 반복문이라 한다. 그리드의 스레드 변위를 계산하기 위해 CUDA는 그리드의 블록 수를 제공하는 특수 변수를 제공한다. 이 특수변수는 gridDim.x이다. 그리드의 총 스레드 수를 gridDim.x * blockDim.x 수식으로 계산할 수 있으며, 다음 데이터를 읽을 그리드의 스레드 변위(stride)가 된다.

그리드의 총 스레드 수: gridDim.x * blockDim.x



〈그림 12〉 32개 병렬 커널 처리 스레드 인덱스:
threadIdx.x(4) + blockIdx.x(2) * blockDim.x(4)

다음은 32개의 데이터를 처리하기 위해 〈〈〈2, 4〉〉〉() 실행 설정을 구성한 내용이다. 실행 구성은 4개의 스레드를 가지는 2개 스레드 블록으로 그리드는 총 8개의 스레드를 가진다. 이때 그리드의 8개 스레드는 32개의 데이터를 동시에 처리할 수 없기 때문에 8개씩 데이터를 나누어 처리한다. 따라서 그리드 내의 각 스레드는 총 4번 실행이 된다.

Ⅴ. NVIDIA AI 생태계 활용

1. AI 활용분야

NVIDIA의 가속 컴퓨팅 기술을 이용하여 데이터 사이언스, 의료, 헬스, 로보틱스, 자율주행 분야까지 AI 기술과 산업 발전에 큰 영향을 미치고 있다. 이는 NVIDIA CUDA 플랫폼기반으로 하는 병렬 컴퓨팅 성능을 통해서 딥러닝, 데이터센터의 비지니스 모델, 연구, 슈퍼컴퓨팅, 게이밍분야 발전에 핵심적인 기술을 제공하고 있다.

2. 인공지능과 데이터 사이언스

AI는 전 세계의 모든 산업에서 변화를 주도하고 있다. 기업들은 빅데이터 분석에 대해 AI 기술을 활용해서 인사이트를 도출하고 의사결정을 시작하고 있다. 또한 AI 기술을 이용해서 음성 인식 및 추천 시스템 활용에서부터 의료영상, 공급망 관리, 자율주행 시스템까지 활용되고 있다. NVIDIA는 AI 기술을 구현하는데 필요한 고성능 GPU와 AI 플랫폼을 제공하고 있다. 이에 기업, 과학자 및 일반 사용자는 AI와 데이터 사이언스 분석에 필요한 컴퓨팅 리소스와 툴을 쉽게 제공 받게 되었다. NVIDIA에서 지원하는 데이터 사이언스 분야 기술로는 데이터분석, 머신러닝, 딥러닝, 그리고 대화형 AI기술이 있다.

가) 데이터 분석

일반적인 데이터 분석 처리는 CPU 컴퓨팅을 활용해서 데이터 수집, 전처리, 학습, 그리고 분석 및 시각화를 진행하였다. 이는 속도가 느리고, 프로세스 처리가 번거로웠다. 반면에 가속화된 GPU 컴퓨팅을 이용하게 됨으로서 데이터와 인프라를 최대한 활용할 수 있게 되어 적은 비용에 빠른 데이터 분석이 가능하게 되었다. 고속 GPU 컴퓨팅 기술은 데이터 변환에서부터 시각화까지 모든 데이터 분석 처리 단계를 향상시키고 있다. NVIDIA는 데이터분석 과정의 전체 프로세스를 GPU기반에서 모두 처리 할 수 있도록 최적화 된 솔루션인 RAPIDS와 NVIDIA CUDA를 제공한다. RAPIDS와 NVIDIA CUDA를 통해 NVIDIA GPU의 병렬처리 분석을 가속화하여 데이터 로딩, 처리, 학습과 같은 데이터 분석 작업 시간을 며칠에서 몇 분으로 단축할 수 있게 되었다. 이는 CPU 기반 데이터 분산저장 및 처리하는 프레임워크인 하둡(Hadoop)이나 Apache Spark를 이용하여 데이터 처리하는 것보다 GPU 기반 데이터 분석처리가 월등히 빠름을 알 수 있다. 여기서 RAPIDS는 데이터 분석 병렬 처리를 완전히 GPU에서 실행하기 위한 오픈 소스 소프트웨어 라이브러리 및 API이다. RAPIDS는 머신 러닝, 딥 러닝, 그래픽, 고성능 컴퓨팅(HPC) 등의 기능이 통합되어 있다.

나) 머신러닝

기업은 머신러닝을 활용하여 데이터를 분석하고 사용자의 행동 패턴을 예측하여 제품이나 서비스를 개선할 수 있다. 이러한 머신러닝은 인간에게 혁신적인 가치를 제공하지만 대규모 데이터 세트를 사용하여 반복적인 학습과정을 거치기 때문에 많은 컴퓨팅 리소스가 필요 되어 진다. 고성능 GPU 기반 플랫폼 활용은 머신러닝 학습 속도

를 향상시켜 많은 양의 데이터를 빠른 시간에 학습 할 수 있게 되어 정확한 예측 결과를 얻을 수 있게 되었다. 고성능 GPU 기반 머신러닝은 RAPIDS와 NVIDIA CUDA를 통해 병렬처리를 가속화 할 수 있으며, Python이나 Java 기반 언어를 통해 CUDA의 성능을 활용하여 손쉽게 가속 머신 러닝을 활용할 수 있다.

다) 딥러닝

GPU 기반 딥러닝은 컴퓨터비전, 음성인식, 자연어 처리, 추천엔진 분야에 획기적인 발전을 가져다주고 있다. 특히 GPU 병렬컴퓨팅 기술과 다양한 오픈 데이터 세트를 활용해서 다양한 딥러닝 학습모델이 소개되고 활용되고 있다. 컴퓨터 비전분야는 수백만 개의 이미지를 분류할 수 있으며, 콘텐츠 분석 및 탐지에서부터 이미지 레이블링, 얼굴인식 등의 고급기능을 구현할 수 있다. 음성인식분야는 통화내용, 동영상 대화 내용 등을 인식하고 키워드를 빠르게 추출할 수 있다. 음성 인식 기술을 활용하여 미디어 콘텐츠 검색 및 분석을 할 수 있으며, 오디오 및 비디오를 아카이브로 자동 변환해서 검색할 수 있게 한다. 자연어 처리는 컴퓨터가 언어를 알아들을 수 있는 기술로서 번역 및 맞춤법 검사 등에 활용되고 있으며, AI 스피커의 자연언어 처리를 통한 대화시스템과 챗봇 서비스 등을 제공하고 있다. SNS 데이터에 대한 자연어 처리를 통해 사용자 분석을 하고 기업체의 마케팅에 활용하고 있다. 특히 채팅/음성/콜센터에서 자연어 처리를 통해 사기탐지 서비스를 제공하고 있다. 이는 고성능 GPU가 머신 러닝(ML) 및 딥 러닝(DL)을 학습하는 데 필요한 컴퓨팅 성능을 제공하여 짧은 시간 동안에 문제점을 파악하고 개선할 수 있게 되었으며, 동시에 다채널에서 이상탐지를 감지할 수 있게 되었다. NVIDIA는 웹기반 대화형 AI로 자동음성인식(Automatic Speech Recognition)과 개체명 인식(Named-Entity Recognition)을 처리하는 NVIDIA Riva 프로그램을 제공한다. Riva는 음성을 실시간으로 스트림해서 문자로 필사하고, 자연언어 처리를 통해 핵심 구문 내용을 추출해서 표시한다. 추천엔진은 넷플릭스와 같이 사용자에게 선호하는 다양한 영화를 제안하고, 아마존과 같이 고객에 맞는 상품을 추천하는 기술들을 의미한다. 추천시스템 데이터 분석은 일반적으로 머신러닝을 활용한다. NVIDIA에서는 고성능 추천 시스템을 구축하기 위한 오픈 소스 프레임워크인 NVIDIA Merlin™을 제공한다. Merlin은 수백 테라바이트의 데이터를 처리할 수 있고, 복수개의 GPU를 분산해서 병렬 처리하는 딥러닝 기술과 추론기술을 가지고 있다. NVIDIA는 딥 러닝 기술과 솔루션을 쉽게 액세스 할 수 있게 NGC(NVIDIA GPU CLOUD)를 제공한다. NGC는 딥 러닝뿐만 아니라 머신 러닝 및 HPC용 GPU 가속 소프트웨어를 제공한다.

3. 클라우드와 데이터센터

AI와 데이터 분석은 많은 컴퓨팅 리소스가 필요하기 때문에 고성능 컴퓨팅(HPC;high-performance computing)을 제공할 수 있는 데이터센터를 활용한다. 최근에는 빅데이터 분석과 복잡한 AI 학습으로 인해 점점 더 높은 연산 성능이 요구되어 지고 있다. 이러한 수요에 대응하기 위해 NVIDIA는 하드웨어와 소프트웨어가 최적으로 통합된 가속 GPU 컴퓨팅 플랫폼을 제공하여 개발에서 배포까지의 구현을 지원하는 강력한 보안 인프라를 활용할 수 있게 하였다. NVIDIA는 데이터센터용 GPU와 클라우드 기반 GPU 솔루션을 제공한다. GPU 가속 데이터센터는 필요한 만큼 서버를 설정한 후 컴퓨팅 및 그래픽 작업에 혁신적인 성능을 제공하여 더 빠르게 정보를 얻고 비용을 획기적으로 절감하게 만든다. 또한 기밀 데이터를 저장, 처리, 분석하면서도 운영 보안을 굳건히 유지할 수 있다. 클라우드 기반 GPU 솔루션을 통해 고성능 컴퓨팅 리소스와 가상 워크스테이션에서 언제 어디서든 물리적 데이터센터를 구축할 필요도 없이 원하는 만큼 고성능 GPU 컴퓨팅 파워를 제공받아 데이터 분석, 학습, 추론, 고성능 컴퓨팅, 렌더링 및 가상화 업무를 빠르고 적은 비용으로 구현할 수 있다. NVIDIA EGX™ 플랫폼은 클라우드에서 AI와 데이터 분석, 제품 혁신을 위한 고급 시각화에 이르기까지 가속 컴퓨팅을 위한 통합된 인프라를 제공한다. NVIDIA EGX™ 플랫폼은 인프라 전반에 걸쳐 컴퓨팅을 가속화하는 최적화된 소프트웨어인 NVIDIA AI Enterprise을 제공한다. NVIDIA

AI Enterprise는 VMware vSphere에서 실행되는 클라우드 기반 AI 및 데이터 분석 소프트웨어로서 AI 업무를 빠르게 배포, 관리 및 확장을 할 수 있다. NVIDIA EGX 플랫폼은 또한 IT 공급업체 및 소프트웨어 개발도구와 호환되는 IT 인프라를 제공한다. NVIDIA EGX 플랫폼은 데이터센터 AI, 에지 AI, 데이터분석, 전문적인 시각화 및 원격협업 작업을 지원한다. NVIDIA Omniverse Enterprise는 실시간 성능으로 원격 협업을 수행한다. 사용자는 고성능 GPU기반 클라우드에서 제공하는 NGC를 이용해서 딥 러닝, 머신 러닝, 그리고 HPC용 GPU 최적화 소프트웨어를 무료로 이용할 수 있으며, 클라우드의 GPU 기반 병렬 프로세싱을 활용하여 복잡한 대규모 애플리케이션 프로그램을 빠르고 효율적으로 실행할 수 있게 되었다. 그리고 GPU 가속 그래픽을 위한 NVIDIA RTX 가상 워크스테이션(vWS)을 통해 가상 머신에서 NVIDIA RTX™ 기술을 사용한 최신 실시간 레이 트레이싱을 비롯하여 가장 앞선 3D 그래픽 플랫폼을 활용할 수 있다. 그리고 최근에는 NVIDIA DGX 플랫폼을 제공하여 통합 솔루션을 제공하고 있다.

4. 에지 컴퓨팅

에지 컴퓨팅(Edge Computing)은 중앙 집중식 서버나 클라우드가 아닌 데이터가 수집되고 분석되는 물리적 위치 근처에서 컴퓨팅이 이루어지는 분산 컴퓨팅 모델이다. 즉 네트워크의 에지(edge)에서 처리하므로 지연시간과 대역폭 요구사항이 최소화되어 오토노머스(autonomous) 머신들이 데이터 전달과 의사 결정을 실시간으로 진행할 수 있게 된다. 최근 개인, 산업, 공공기관에서 수십억 개의 IoT 센서를 통해서 매일 엄청난 양의 데이터를 생성하고 있으며, 이 데이터들을 실시간으로 분석해서 인사이트를 추출하고 의사결정을 할 수 있기 위해서는 데이터가 있는 네트워크의 에지에 AI 컴퓨팅을 설치해야 한다. NVIDIA는 에지 컴퓨팅을 위한 NVIDIA Jetson 플랫폼을 제공한다. NVIDIA Jetson 플랫폼은 자율주행, 로보틱스, 인텔리전트 비디오 분석, 그리고 오토노머스 머신을 개발하고 배포하는 도구를 제공한다. NVIDIA Jetson 플랫폼은 Jetson AGX Orin 시리즈, Jetson Orin NX 시리즈, Jetson Nano, Jetson AGX Xavier 시리즈, Jetson Xavier NX 시리즈, 그리고 Jetson TX2 시리즈가 있다.

5. 자율주행 자동차

자율주행기술은 자동차가 정보통신기술(ICT)과 융합해 스스로 운행하는 첨단 기술을 말한다. 자율주행기술은 운전자가 있는 자율주행차를 위주로 차로유지, 차선변경, 고속도로 자율주행, 도심 자율주행, 자동주차, 완전 자율주행 등으로 영역을 넓혀 가고 있다. 자율주행 핵심 개발기술은 무인자동차 시스템 구현기술, 비전과 다양한 센서(카메라, 라이다, 레이더, 초음파 센서 등)를 이용하여 시각정보를 입력받고 처리하는 기술, 자동운행기술, 그리고 클라우드 기반 통합관제 시스템과 운행감시 고장진단 체계 기술 등이 있다. 일반적인 자율주행차는 카메라, 라이다, 레이더, 기타 센서와 HDMap을 장착한다. 센서로부터 획득한 데이터는 클라우딩 컴퓨팅 환경에 수집이 되어 뉴럴 네트워크 학습을 통해서 학습모델을 생성한다. 이는 자율주행 차량을 네트워크로 연결해서 실제 운행 데이터베이스를 축적한다. 그리고 인공신경망을 적용하여 인식과 운전방법을 딥러닝으로 학습한다. 축적된 데이터는 각각의 인공신경망이 분석하여, 의미적 분할, 객체 인식, 거리의 예측을 수행한다. 모든 이미지 데이터를 통합하여 도로의 모양과 정적인 구조물들, 3차원 객체를 분석해 내며, 자동차로부터 수집되는 데이터를 실시간으로 분석하여 최적의 자율주행 운행정보를 제공하며, 각각의 운행 상황으로부터 학습을 한다. NVIDIA는 고성능 GPU 컴퓨팅, 이미징 및 AI 분야에서 수십 년간 쌓은 전문 지식을 활용하여 NVIDIA DRIVE AGX 플랫폼을 구축하였다. NVIDIA DRIVE AGX 플랫폼은 자율주행 자동차가 대량의 센서 데이터를 처리하고 실시간 주행 결정을 내릴 수 있게 지원한다. NVIDIA DRIVE AGX 플랫폼 하드웨어는 NVIDIA DRIVE Orin, NVIDIA DRIVE AGX Pegasus, NVIDIA DRIVE AGX Xavier, 그리고 NVIDIA DRIVE Atlan이 있다. NVIDIA DRIVE는 완전 자율 주행에서부터 제한된 주행까지 모든 기능을 포함한 운전자 보조 시스템 솔루션으로서 자율주행 딥러

닝 학습, 인식, 매핑, 계획 그리고 운전자 모니터링 기능을 구축할 수 있게 한다. NVIDIA DRIVE는 DRIVE OS, DriveWorks, DRIVE AV, DRIVE Chauffeur, DRIVE IX, DRIVE Concierge, 그리고 NVIDIA Map으로 구성되어 있다. DRIVE OS는 자동차 전용 가속 컴퓨팅용 운영 체제로서 센서 입력 처리를 위한 NvMedia, 효율적인 병렬 컴퓨팅 구현을 위한 NVIDIA CUDA 라이브러리, 실시간 AI 추론을 위한 NVIDIA TensorRT, 하드웨어 엔진 액세스를 위한 기타 개발자 도구 및 모듈을 포함한다. NVIDIA DriveWorks는 자율주행 자동차 개발용 오픈 프레임워크로서 센서 추상화 레이어(SAL), 센서 플러그인, 데이터 레코더, 자동차 I/O 지원, 딥 뉴럴 네트워크(DNN) 프레임워크로 구성되어 있다. DRIVE AV 소프트웨어는 인식, 매핑, 계획 레이어 뿐 아니라 고품질 실제 주행 데이터로 학습된 다양한 DNN도 포함되어 있다. DRIVE AV를 활용하여 자율주행과 매핑에 사용할 수 있으며, 계획 및 제어 레이어는 차량 손상과 위험한 상황을 방지한다. NVIDIA DRIVE Chauffeur는 NVIDIA DRIVE AV SDK 기반의 AI 보조 주행 플랫폼으로 고속도로 및 도시 교통 상황에서 안전 운행할 수 있게 한다. DRIVE IX는 AI 운전석 솔루션을 위한 인텔리전서 사용자 경험을 제공하는 오픈 소프트웨어 플랫폼으로 운전자 및 탑승자 동작 모니터링, AR/VR 시각화, 자연어 이해를 통한 음성 제어 기능 및 DNN을 활용한 인식 애플리케이션 개발을 제공한다. NVIDIA DRIVE Concierge는 NVIDIA DRIVE IX 및 Omniverse Avatar를 기반으로 제작된 실시간 대화형 AI 서비스로 운전자에게 지능형 서비스를 제공한다. 운전자에게 디지털 비서 역할을 수행하여 추천, 예약, 전화통화에서부터 자동차를 주차하고 호출하는 발렛 서비스를 제공한다. NVIDIA Map은 지상 실측 맵과 클라우드에서 제공하는 정보를 통해 최신 맵으로 유지되며, DRIVE Hyperion에서 카메라, 레이더, 라이더와 같은 다양한 센서 정보를 사용하여 교차로 세부 정보, 신호등, 주차 공간, 도로 경계를 확인하고 안전한 주행 경로를 판단할 수 있게 한다. NVIDIA는 Orin 임베디드 시스템 기반에서 자율주행 자동차를 설계하기 위한 개발 플랫폼이자 레퍼런스 아키텍처인 NVIDIA DRIVE Hyperion을 제공한다. DRIVE Hyperion은 자율주행을 위한 전체 소프트웨어 스택과 운전자 모니터링 및 시각화(DRIVE IX) 기능을 가지고 있으며, 자체 무선 업데이트 기능이 있어 새로운 기능을 추가할 수 있다.

Ⅵ. 전망과 결론

최근 OpenAI의 ChatGPT를 포함한 생성형 AI 기술의 발전으로 빅테크 회사는 사활을 걸고 관련 서비스를 경쟁적으로 출시하고 있다. 이러한 비지니스 트렌드로 인해 NVIDIA GPU의 AI 가속기 수요가 폭발적으로 늘어나고 있으며, 특히 H100 가속기 칩의 수요와 가격이 매년 큰 폭으로 증가하고 있다. 또한 AI 칩 시장에서 엔비디아가 AI 생태계 표준으로 자리 잡고 있어서 80%이상의 점유율을 차지하고 있다. AI 가속기의 발전으로 삶의 질과 생산성이 높아지고 있지만, AI 칩이 소비하는 연간 소비전력량이 소규모 국가가 소비하는 양을 넘어서고 있다. 따라서 최근 추론 영역에서는 NVIDIA GPU 기반이 아니라 추론 전용 AI 칩을 사용하고자 하는 움직임이 나타나고 있다. MS, 아마존, 메타 등의 빅테크 기업에서 신경망처리장치(NPU) 등을 활용한 자체 추론용 AI 가속기를 개발해서 사용하고 있다. 이와 같이 AI 가속칩은 성능이 뛰어나면서 전력 효율성이 있는 방향으로 개발이 진행되고 있다. 대표적으로 Grog 회사에서는 LPU 추론칩에 HBM을 사용하지 않고 SRAM을 사용하여 빠른 데이터 전송 속도를 활용해서 실시간 ChatGPT를 운용할 수 있다고 한다. 그리고 최근 삼성전자는 그래픽처리장치(GPU)와 메모리반도체 사이의 병목현상을 8분의 1로 줄일 수 있는 파격적인 구조와 기존 AI 칩에서 쓰였던 고대역폭메모리(HBM) 없이 저전력(LP) 메모리만으로도 활용할 수 있게 한 FPGA형 AI 가속기 '마하-1'을 개발 중에 있다. 이러한 하이엔드 AI 가속기 칩 발전과 아울러 AI 생태계 표준 환경을 선점해서 AI 시장의 주도권을 독점하기 위한 싸움이 치열할 것으로 예상된다. 향후 GPU기반 생성형 AI를 활용하여 사용자 제품추천, OTT 콘텐츠 추천, 기업 고객 응대서비스 등이 더욱 정교해지고, 관련 서비스들이 B2B 비즈니스로 확장되어 모든 영역에서 AI 서비스를 활용하

게 될 것이다. 이와 함께 생성형 AI에서 발생될 수 있는 할루시네이션 폐해 방지와 AI 발전으로 인한 국민안전과 기본권을 보호하는 AI 규제법이 마련되어야 할 것이다.

참고문헌

[1] Nvidia-ampere-architecture-whitepaper

[2] Nvidia-hopper-architecture-whitepaper

[3] Ofer Rosenberg, Introduction to GPU Architecture

[4] Jason sanders, CUDA by example, nvidia

[5] CUDA C++ Programming Guide, nvidia

[6] GPU Architecture & CUDA Programming, CMU

[7] Accelerating Application with CUDA C/C++, nvidia

[8] https://www.nvidia.com/

[9] https://developer.nvidia.com/embedded/jetson-nano - developer-kit

[10] https://github.com/NVIDIA-AI-IOT/jetracer

[11] https://www.tensorflow.org/

[12] 김상범 외1인, "Nvidia GPU 보드를 이용한 AI활용" 한국폴리텍대학 2022

[13] Alec Radford et al. "Improving language understanding by generative pre-training. OpenAI" Preprint 2018

[14] Ashish Vaswani et al. "Attention is All you Need", Advances in Neural Information Processing Systems 30 (NIPS 2017)

[15] Alec Radford, et al. "Language models are few-shot learners", Advances in neural information processing systems 2020

[16] J Choquette et al. "Nvidia a100 tensor core gpu: Performance and innovation", IEEE Micro, 2021

[17] J Achiam et al. "Gpt-4 et al. technical report", arXiv preprint arXiv:2303.08774

김상범

- 1993년 동아대학교 전자공학과 학사
- 1995년 동아대학교 전자공학과 석사
- 1999년 동아대학교 전자공학과 박사
- 현 한국폴리텍대학 스마트소프트웨어과 교수

〈관심 분야〉
IOT, 비전처리, 빅데이터, 인공지능

생성형 인공지능의 연구 동향

Ⅰ. 서 론

바야흐로 인공지능의 시대이다. 초등학생으로부터 일반인에 이르기까지 대중적으로 인공지능을 사용하는 세대에 이르렀다. 대형언어모델(LLM)을 기반으로 한 ChatGPT(Generative Pre-trained Transformer)는 출시된 지 2달 만인 지난 2023년 1월 1억 명을 돌파했고, 그해 5월에 18억 명의 사용자를 기록하였다. 미국 인구의 63.4%인 약 2.1억 명이 사용하고 있다고 한다.

생성형 인공지능은 2018년 트랜스포머(Transformer)출현 이전과 이후로 나눌 수 있다. 그전엔 Computer Vision 분야는 CNN이 Natural Language 분야는 RNN+LSTM 등으로 각 분야별 모델들이 훌륭한 모델들이 적용돼서 각기 발전해 왔다. 하지만, 2017년 NIPS에 구글팀이 내놓은 논문 "Attention is All You Need" 이후 언어모델에 있어서 압도적인 성능 우위를 보이게 되며 기초모델(foundation model)로 자리잡게 되었다.[5]

이 논문은 Attention 메커니즘에 의존적인 것이 가장 큰 특징으로 기존의 RNN이나 CNN보다 훨씬 효율적이다. 특히, Self Attention을 제안함으로써 시퀀스를 구성하는 각 요소가 재귀적 패턴 없이 시퀀스 내의 모든 다른 요소와의 상호작용을 모델링함으로써 시퀀스를 표현하였다. 게다가, Multi-Head Attention의 개념을 도입하여 모델이 다양한 위치에서 다양한 표현공간의 정보를 동시에 주시할 수 있도록 함으로써 데이터의 맥락적 관계를 더 잘 포착할 수 있게 했으며, 기존의 단순한 시퀀스 정렬에 기반한 아키텍처보다 훨씬 더 효율적으로 작업처리가 가능한 병렬처리를 도입하여 빠른 학습이 가능하도록 설계되었다.

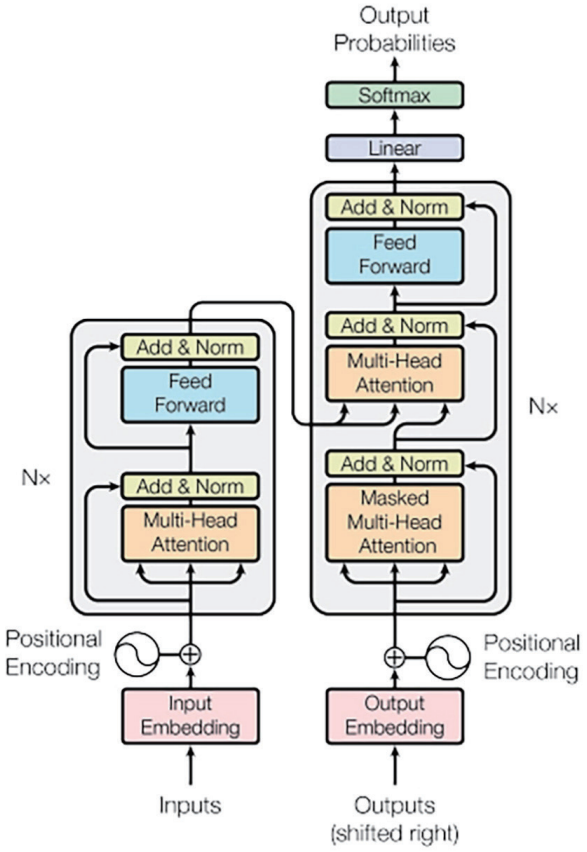
순환연결이 없는 트랜스포머(Transformer)는 계산 부담을 줄일 수 있기에 하드웨어의 부담을 상대적으로 줄일 수 있으며, 번역 작업에서

권혁준
한국폴리텍대학

의 뛰어난 성능뿐만 아니라 문장의 구분 분석과 같은 영역에서도 유용한 결과를 보여주며, 다양한 분야에 적용 가능성을 보여주며 그를 입증하였다. Transformer에서의 Attention은 'Scaled-Dot Attention'으로 불리며 그 표현은 다음과 같다.[5]

$$Attention(Q,K,V) = softmax(\frac{QK^T}{\sqrt{d_k}})V$$

위의 수식에서 Q는 query의 집합, K는 key의 집합, V는 Value의 집합을 의미하며, d_k는 key 벡터의 차원으로 스케일링을 통해 softmax 함수의 입력값 범위를 조절하여 학습의 안정성을 높인다. 즉, Attention 메커니즘은 주어진 query에 대해 모든 keys와의 유사도를 계산하고, 이 유사도 점수를 사용하여 values를 가중 평균하여 출력을 계산한다. Query와 key의 dot product 수행을 통해 두 벡터 간의 유사도를 계산한다. 그 값이 크면 클수록 query와 해당 key가 더 유사하다는 것을 의미한다. $\sqrt{d_k}$로 나누는 스케일링은 학습 과정에서의 안정성을 도모한다. Softmax 함수를 통해 모든 key에 대한 유사도 점수를 확률로 변환하여 각 value에 대한 가중치로 사용한다. 이 가중치는 모든 key-value쌍에 대해 계산되어 최종적으로 value들의 가중합으로 출력이 결정되고, 결국 각 쿼리(query)에 대해 계산된 가중합은 다음 레이어에 전달되며 출력이 된다. 이러한 Attention 매커니즘을 통해 주어진 문제에 대한 중요한 정보에 집중하고 관련 없는 정보를 무시할 수 있게 되는 결과를 가져온다. 이는 곧 의미론적 가중치를 선별할 수 있는 기준을 제공하는 셈이된다. 또한 이 과정은 병렬처리가 가능하기에 대규모 데이터에 대한 처리가 기존 모델에 비해 효율적으로 이루어질 수 있다.

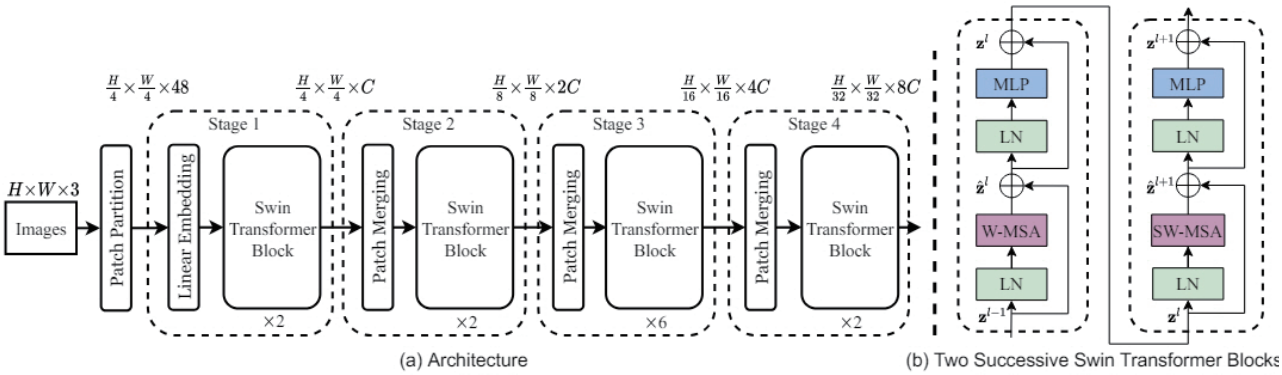


〈그림 1〉 Transformer Model Architecture[5]

Ⅱ. 생성형 인공지능의 역사

1. 트랜스포머(Transformer) 이전과 이후

생성형 인공지능 역사의 전부를 다루는 것은 지면 관계상 쉽지 않은 일이다. 따라서, 생성인공지능의 이정표가 되는 부분을 골라서 언급하고자 한다. 우선, 트랜스포머(Transformer) 이전에는 어떠했는지 생성형 인공지능의 역사에 대해서 간략히 살펴보고자 한다. 생성형 모델은 인공지능분야에서는 상대적으로 꽤나 긴 역사를 가지고 있으며, 1950년대로 거슬러 올라간다. 은닉마르코브모형(HMM, Hidden Markov Model)과 가우시안 혼합모형(GMM, Gaussian Mixture Model)은 음성 및 시계열과 같은 순차데이터를 생성했다. 자연어 처리(NLP) 분야는 N-gram 언어모형을 사용하여 단어 분포를 학습한 다음 최적의 시퀀스를 찾는 방법이 사용되었다. 하지만, 이 방법은 긴 문장을 효과적으로 처리하지 못하기 때문에 이후 이를 해결하기 위해 순환신경망(RNN)이 언어모델링작업에 도입되어 이전에 비해 상대적으로 긴 의존성을 모델링할 수 있게 되었다. 그 이후 장단기메모리(LSTM)와 게이트 순환 유닛(GRU)이 개발되었으며, 훈련 중에 메모리를 제어하는 게이팅 매커니즘을 활용하였다. 이 방법들은 샘플에서 약 200개의 토큰에 주의를 기울일 수 있게 하여 N-gram 언어모형에 비해 상당한 개선을 이루었다.[6]



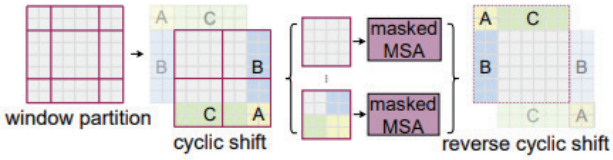
〈그림 2〉 (a) Swin Traansforme(Swin-T)r의 구조와 (b) 2개의 Swin-T를 연결한 모습[12]

컴퓨터 비전(CV) 분야는 딥러닝 등장 이전엔 텍스처 합성과 텍스처 매핑과 같은 전통적인 이미지 생성 알고리즘에 사용되었으며, 이러한 알고리즘은 수작업으로 설계된 특징에 기반을 두고 있으며, 복잡하고 다양한 이미지를 생성하는 능력이 제한적이었다. 하지만, 2014년 생성적 적대 신경망(GAN)이 처음 제안되었고, 이는 컴퓨터 비전 분야에서 중요한 분기점이 되었다. 변분 오토 인코더(VAE)와 확산 생성 모델과 같은 다른 방법들도 이미지 생성 과정을 더 세밀하게 제어하고 고품질 이미지를 생성할 수 있도록 개발되었다.[6]

이렇듯 다양한 생성형 인공지능 분야에서 각기 다른 형태의 발전 형태를 가지고 있었지만, 앞서 설명한 트랜스포머(Transformer)의 출현으로 인해 새로운 교차점으로 컴퓨터 비전 분야에서는 비전 트랜스포머(ViT)와 스윈 트랜스포머(Swin Transformer)[12]가 등장하였다. 자연어처리 분야는 BERT와 GPT와 같은 거대언어모형(LLM)이 탄생하였고, 그밖에 Music Transformer[11] 등이 출현하며 현재와 같은 기초 모델(Foundation Model)로 자리 잡게 되었다. 특히, 컴퓨터 비전 분야를 좀 더 살펴보자면 트랜스포머가 기존의 컴퓨터 비전의 CNN처럼 범용적인 모델로서 사용하기 위해선 두 가지 문제를 해결해야 했다. 첫째, 기존의 트랜스포머 모델의 토큰은 모두 고정된 스케일이라는 것과 둘째, 자연어처리에서의 텍스트에 비해 이미지의 픽셀 해상도가 높다는 문제점을 안고 있었다. 따라서, 이를 해결하기 위해 feature map을 구성하여 계층적 구조로서 처리하고자 하였다. 즉, 기존의 트랜스포머에서는 전역 self-attention을 수행하며, 예를 들어, N개의 시퀀스 개수만큼 같은 수의 토큰을 가지기 때문에 N^2의 복잡도를 갖게 되는데, 그에 반해 스윈 트랜스포머(Swin Transformer)는 그 계산의 복잡성을 줄이기 위해 기존의 2차원(N^2)인 구조를 다층구조로 분할하여 전역이 아닌 로컬 윈도우만을 계산하는 방법을 채택하여 한차원(N) 낮췄다. 각 윈도우 내에서의 7x7 패치, 49개의 패치에 대해서만 수행하기 때문에 선형적 차원에서 수행할 수 있으므로 고해상도 이미지 처리의 계산 부담을 줄이게 해준다.[12] 이를 통해 로컬 윈도우 내에서는 효율적인 계산이 가능하지만, 여전히 문제는 남아있다. 각 윈도우 간의 연결이 부족하여 모델링 능력이 제한되기 때문이다. 따라서, 이를 해결하기위해 스윈 트랜스포머(Swin Transformer) 블록에서 두 개의 파티션을 구성하여 번갈아서 처리하는 구조를 가진다.(Shifted Windows Partitioning in Successive Blocks)[12]

Multi-head Self Attention과 Shifted window Multi-head Self Attention의 계산은 다음과 같다.[12]

$$MSA = 4hwC^2 + 2(hw)^2C$$
$$(S)W\text{-}MSA = 4hwC^2 + 2M^2hwC$$

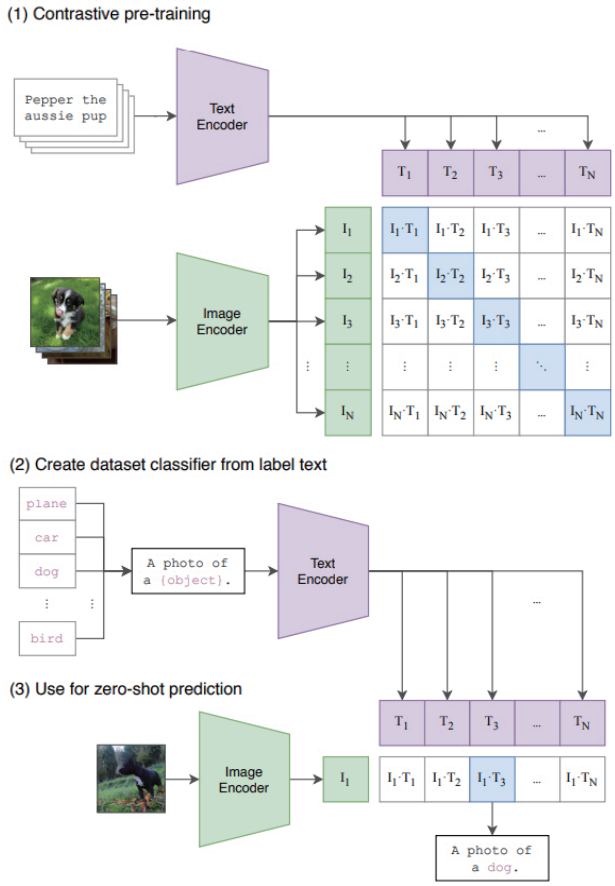


〈그림 3〉 Shifted Widows artitioning의 효율적인 배치계산 도식화[12]

2. 다중모달(Multi-Modal)로의 진화

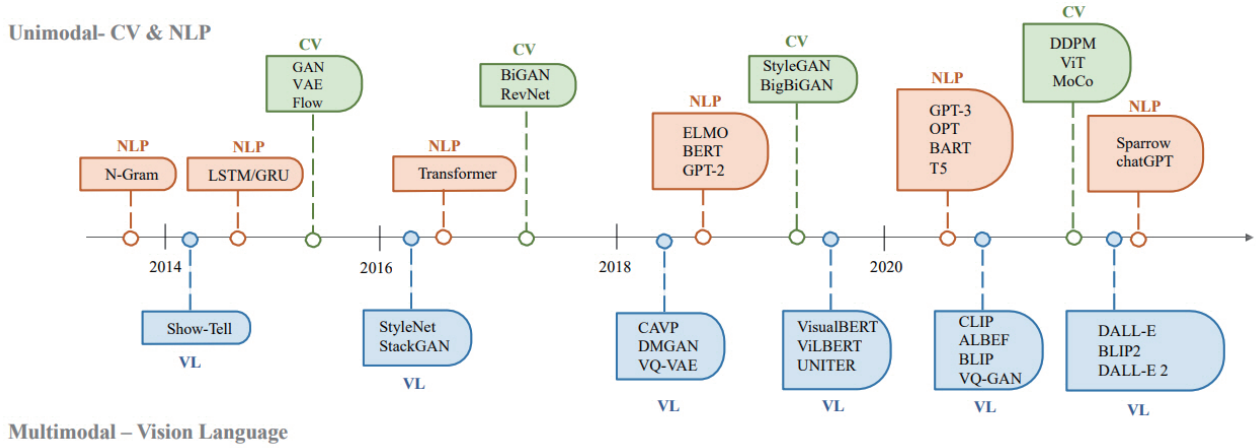
단일모달(Uni-modal)로 시작한 AGI의 역사는 2010년대 중반으로 들어서 CNN과 GAN의 발전으로 다중모달(Multi-Modal) 형태로 진화하기 시작했다. 2021년 다중 모달 모델에 대한 기념비적 논문이 나오게 되는데, OpenAI의 Alec Radford를 포함한 12명의 공저자가 쓴 'Learning Transferable Visual Models From Natural Language Supervision', 일명 CLIP이라고 부르는 논문이다.[13] CLIP은 'Contrastive Language-Image Pre-training'의 약자로 자연어 감독을 통해 이미지 간의 상호작용을 통해 학습하는 비전언어(VL)모델이다.

CLIP은 자연어 감독학습을 통해 이미지와 함께 학습하는 것인데 그전에 Zhang et al. (2020), Gomez et al. (2017), Joulin et al. (2016) 및 Desai & Johnson (2020) 다양한 시도들이 있었으며[13], 대규모 학습을 통해 그 효과를 극대화할 수 있다. CLIP은 훈련된 텍스트 인코더를 사용하여 이미지 분류를 수행한다. 테스트 시에는 텍스트 인코더가 대상 데이터 셋의 클래스 이름이나 설명을 임베딩하여 이미지에 대한 분류를 하며, 이러한 방식으로, CLIP은 훈련 과정에서 본 적 없는 새로운 클래스에 대해도 분류를 수행할 수 있게 된다. 이를 제로샷 분류(Zero-shot Classification)이라고 부른다.



〈그림 4〉 CLIP의 동작방식[13]

〈그림 6〉 다양한 모델의 파라미터수와 훈련속도



〈그림 5〉 단일모달(Uni-modal)과 다중모달(Multi-modal)의 발전 과정[12]

III. 생성적 적대 신경망의 종류

앞서 간략히 살펴본 생성인공지능 중에서 이미지 생성 등에 많이 쓰이는 생성적 적대 신경망 GANs(Generative Adversarial Nets)에 대해서 좀 더 살펴보고자 한다.

1. 바닐라 GAN(Vanilla GAN)

인공지능 분야의 거장 중 한 명인 이안 굿 펠로우의 대표 논문 중 하나로 'Generative Adversarial Nets'에 발표한 모델을 흔히 바닐라 GAN이라고 부른다. 그 이유는 아이스크림의 기본 맛이 바닐라이듯 GAN의 기본이 되는 모델이기에 바닐라 GAN으로 불린다. 바닐라 GAN은 생성기(Generator)와 판별기(Discriminator)로 구성된 가장 간단한 유형의 GAN이다.

여기서 이미지의 분류 및 생성은 다층 퍼셉트론을 사용하여 내부적으로 생성기와 판별기에 의해 수행된다. 생성자는 데이터 분포를 캡처하고 판별자는 입력이 특정 클래스에 속할 확률을 찾으려고 시도하며 마지막으로 손실 함수를 계산한 후 생성자와 판별자 모두에 피드백이 전송되므로 손실을 최소화하려는 노력이 이루어진다. 그 핵심 방법론은 적대적 모델링 프레임워크에서 생성기(Generator) G와 판별기(Discriminator) D가 멀티퍼셉트론(MLPs)일 때 데이터 x에 대한 생성기의 분포 PG를 학습하는 것으로 입력 노이즈변수 z에 대한 사전확률 분포 pz(z)가 정의된다. 생성기(Generator)는 $G(z;\theta_g)$로 미분가능한 함수이며, 파라미터 θ_g를 가지며, 노이즈 변수 z를 데이터 공간 $x = G(z;\theta_g)$로 매핑한다. 판별기는 $D(x;\theta_d)$로 표현하며, 입력 x를 받으며, θ_g를 파라미터를 갖고 있다. 이는 x가 실제 데이터에서 왔는지 아니면 생성기의 분포 pg에서 왔는지를 나타내는 확률이다. 이러한 학습 과정 속에서 두 가지의 최적화 목표를 갖는데 첫째는 판별기 D를 학습하여 실제 데이터와 생성기 G에서 생성된 샘플에 대해 정확한 레이블을 부여할 확률을 최대화하는 것이며, 둘째는 생성기 G를 학습하여 $\log(1 - D(G(z)))$를 최소화함으로써 판별기를 속여 생성 샘플이 실제 데이터인 것처럼 만드는 것이다. 이러한 두 참가자 사이의 미니맥스게임(최소최대게임)과정으로 이어지며 V(G,D)는 가치함수로서 이를 최적화하는 것이다.[1]

2. 조건부 GAN(Conditional GAN)

조건부 GAN(Conditional GAN)은 바닐라 GAN을 간단히 변형한 것으로 생성기 G와 판별기 D 모두에 추가 정보 y를 조건화(Conditioning)하는 것으로 y는 클래스 레이블이나 다른 형식의 데이터와 같은 다양한 종류의 보조 정보가 될 수 있다. 이 조건화는 생성기와 판별기 모두에게 y라는 추가 입력층을 공급하여 수행된다.

생성기에서는 사전 입력 노이즈 pz(z)와 y가 결합되어 공동 은닉 표현(Joint Hidden Representation)을 형성하고, 적대적 학습 프레임워크를 통해 이 은닉 표현을 구성하는 방식에 상당한 유연성이 허용 된다.[2]

이러한 조건부 GAN의 가치함수는 〈수식 2〉와 같다.

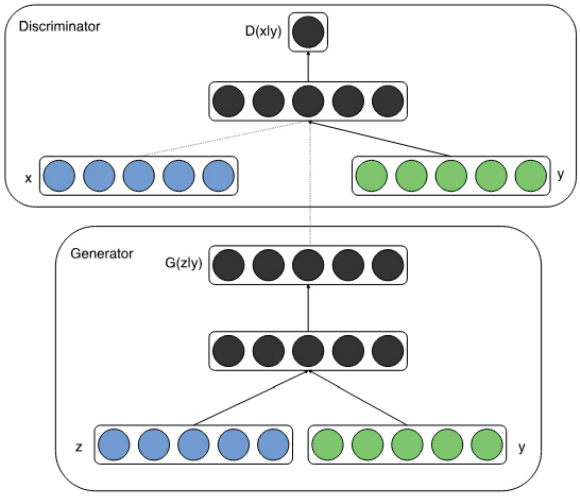
2014년에 발표한 Mehi Mirza, Simon Osindero의 논문 'Conditional Generative Adversarial Nets'에 따르면 〈그림 8〉과 같이 MNIST 데이터 셋을 이용한 조건부 GAN의 실험을 통해 특정 클래스 레이블에 따라 숫자 이미지를 생성할 수 있음을 보여주었다. 그뿐만 아니라, 다중모달(Multi-modal)의 가능성을 보여주며, 유명한 이미지 사이트인 Flickr의 수많은 레이블 처리된 이미지들 즉, 사용자 생성 메타데이터(User Generated Metadata)를 이용하여 각 이미지들에 대해 일반적인 레이블 방식이 아닌 사용자가 설명하는 자연어적 레이블을 사용한 MIR Flickr 25,000 데이터 셋을 프레임워크에 적용하였다. 생성기는 100차원의 가우시안 노이즈와 4096차원 이미지 특징을 각각 500차원, 2000차원 ReLU 은닉층으로 매

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x)] + E_{z\sim p_z(z)}[\log(1 - D(z)))]$$

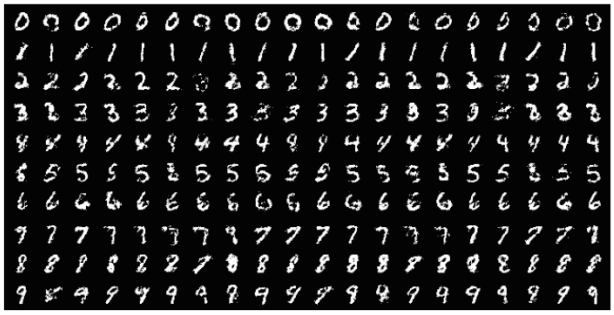
〈수식 1〉 바닐라 GAN의 가치함수 V(G,D)

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}(x)}[\log D(x \mid y)] + E_{z\sim p_z(z)}[\log(1 - D(z \mid y)))]$$

〈수식 2〉 조건부 GAN(Conditional GAN)의 가치함수 V(G,D)



〈그림 7〉 조건부 GAN의 간략한 구조[2]



〈그림 8〉 생성된 MNIST 숫자로 각 행은 하나의 레이블에 따라 조건 지정됨[2]

핑한 후 200차원 선형층에서 결합하여 단어 벡터를 생성한다. 판별기는 단어 벡터(500차원)와 이미지 특징(1200차원)을 ReLU 은닉층으로 매핑하고, 1000 Unit 3조각 맥스아웃층에서 결합한 후 시그모이드(Sigmoid) Unit으로 출력한다.

모델은 미니배치 100, 초기학습률 0.1, 지수감소, 모멘텀, 드롭아웃 0.5 등의 하이퍼파라미터로 훈련된다. 이런 하이퍼파라미터와 아키텍처는 교차검증, 랜덤 그리드 탐색 등으로 선정되어 생성된 샘플들 중 가장 일반적인 단어 벡터들이 태그로 선택된다.

〈그림 9〉는 조건부 GAN을 통해 이미지에 대한 태그 분포를 생성하고, 가장 빈번한 태그를 다중 레이브 예측으로 출력한 결과를 보여준다. 참고로, 조건부 GAN은 마르코프 체인이 필요 없으며, 기울기를 얻기 위해 오로지 역전파(back propagation)만을 사용 한다.[2]

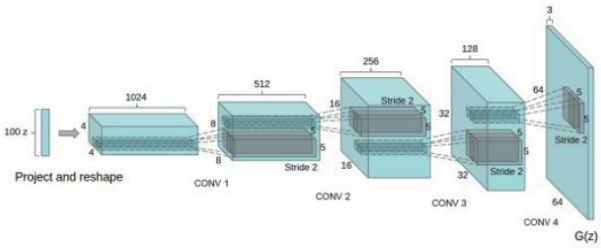
	User tags + annotations	Generated tags
	montanha, trem, inverno, frio, people, male, plant life, tree, structures, transport, car	taxi, passenger, line, transportation, railway station, passengers, railways, signals, rail, rails
	food, raspberry, delicious, homemade	chicken, fattening, cooked, peanut, cream, cookie, house made, bread, biscuit, bakes
	water, river	creek, lake, along, near, river, rocky, treeline, valley, woods, waters
	people, portrait, female, baby, indoor	love, people, posing, girl, young, strangers, pretty, women, happy, life

〈그림 9〉 생성된 태그들의 샘플[2]

3. DCGAN(Deep Convolution GAN)[21]

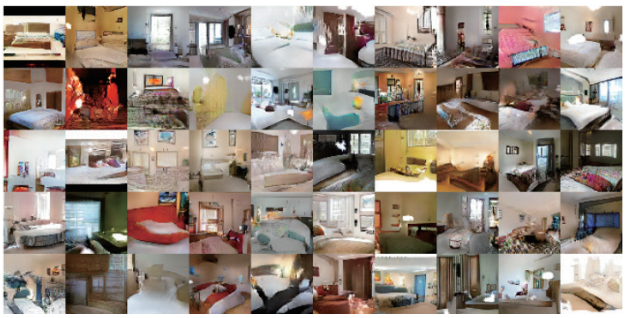
DCGAN은 생성기가 깊은 컨볼루션 네트워크를 사용한 최초의 GAN으로 고해상도 이미지를 생성한다. 기존 GAN의 훈련이 불안정하여 무의미한 출력을 종종 생성하기에 때문에 이를 이해하고 시각화하려는 연구의 동기로 시작되었다.

Indico Research의 Alec Radford와 Luke Metz, 페이스 북(현. 메타)의 Soumith Chintala는 복잡한 이미지 분포의 비지도 학습을 안정적으로 수행하기 위한 아키텍처를 목표로 하여 기존 GAN 아키텍처에 많은 변화를 주었다. 풀링 레이어를 스트라이드 합성곱으로 대체하고, 학습을 안정화하기 위해 배치 정규화를 사용하며, 합성곱 기능 위의 완전 연결 레이어를 제거하였다. 적대적 훈련은 미니 배치와 Adam 최적화기를 사용하여 튜닝된 하이퍼파라미터를 사용하였으며, LSUN(Large-scale Scene



〈그림 10〉 LSUN 장면 모델링에 사용된 DCGAN생성기

〈그림 11〉 데이터 셋을 통한 한 번의 훈련 이후 생성된 침실 이미지[21]



〈그림 12〉 5번의 Epoch 훈련 후 생성된 침실 이미지[21]

Model	Accuracy	Accuracy (400 per class)	max # of features units
1 Layer K-means	80.6%	63.7% (±0.7%)	4800
3 Layer K-means Learned RF	82.0%	70.7% (±0.7%)	3200
View Invariant K-means	81.9%	72.6% (±0.7%)	6400
Exemplar CNN	84.3%	77.4% (±0.2%)	1024
DCGAN (ours) + L2-SVM	82.8%	73.8% (±0.4%)	512

〈그림 13〉 사전 훈련된 모델을 사용한 CIFAR-10 분류 결과[21]

Understanding), Imagenet-1k 및 새로운 얼굴 데이터 셋을 사용하여 훈련하였으며, 선형 SVM 분류기에 대한 특성을 추출하여 CIFAR-10에서 82.8%의 정확도를 달성하였다.[21]

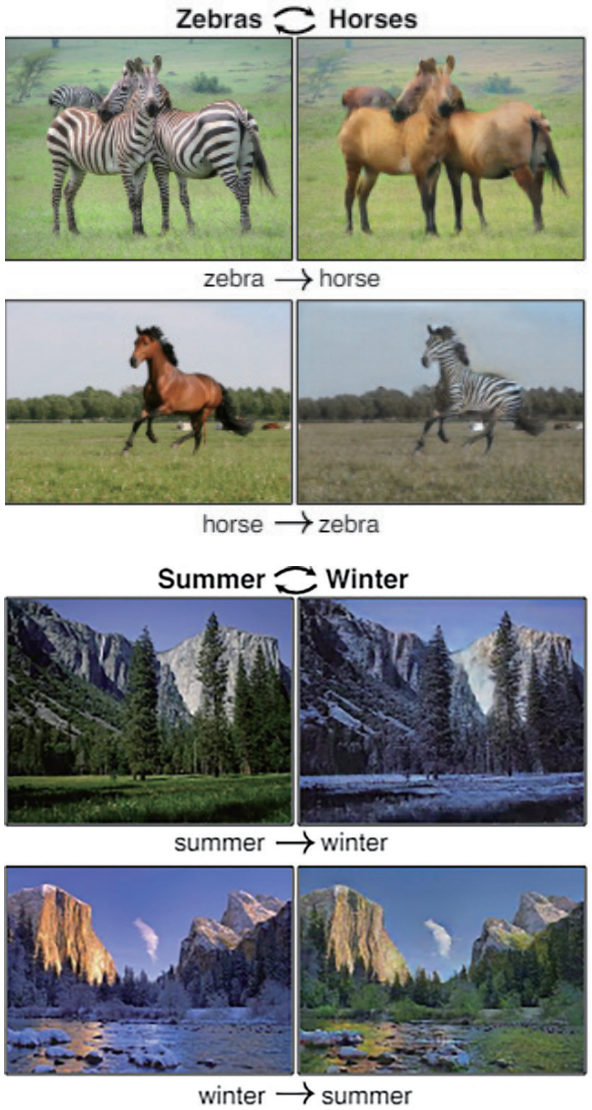
〈그림 10〉은 DCGAN에 대한 구조를 나타낸다. 100차원의 균일 분포 Z가 많은 feature map을 가진 작은 공간 영역의 컨볼루션 표현으로 투영되며, 이 고수준 표현은 4개의 분수 스트라이드 컨볼루션을 거쳐 64x64 픽셀이미지로 변환된다. 주목할 점은 완전연결층이나 풀링층이 전혀 사용되지 않는다는 것이다. 정리하면, 랜덤 노이즈 벡터가 초기 컨볼루션 표현으로 매핑되고, 여기에 스트라이드 컨볼루션을 적용하여 점진적으로 업샘플링하여 최종 이미지를 생성하는 구조로서 이때 완전연결층이나 풀링층 없이 오직 컨볼루션 연산만을 사용하는 특징을 보인다.

Model	error rate
KNN	77.93%
TSVM	66.55%
M1+KNN	65.63%
M1+TSVM	54.33%
M1+M2	36.02%
SWWAE without dropout	27.83%
SWWAE with dropout	23.56%
DCGAN (ours) + L2-SVM	22.48%
Supervised CNN with the same architecture	28.87% (validation)

〈그림 14〉 1000개의 레이블 있는 SVHN분류 결과[21]

4. CycleGAN

CycleGAN은 이미지 대 이미지 변환을 위해 만들어졌다. 즉, 하나의 이미지가 다른 이미지와 매핑되는 것을 뜻한다. 예를 들어, 남자와 여자의 이미지-이미지 변환 과정을 거치게 하면 매핑 함수에 따라 특징을 추가하거나



〈그림 15〉 Cycle GAN를 수행한 모습[22]

$$LGAN(G,DY,X,Y) = E_{y\sim p_{data}(y)}[\log DY(g)] + E_{y\sim p_{data}(y)}[\log(1 - DY(G(x)))]$$

〈수식 3〉 적대적 손실[22]

$$L_{cyc}(G,F) = E_{x\sim P_{data}(x)}[\| F(G(x)) - x \|_1] + E_{y\sim P_{data}(y)}[\| F(G(y)) - y \|_1]$$

〈수식 4〉 주기 일관성 손실[22]

$$L(G,F,DX,DY) = LGAN(G,DY,X,Y) + LGAN(F,DX,Y,X) + \lambda L_{cyc(G,F)}$$

〈수식 5〉 전체 목적함수[22]

제거하여 남자 이미지를 여자 이미지로 또는 그 반대로 변환할 수 있는 매핑 함수를 찾는다. 예측된 출력과 실제 출력은 손실을 최소화하였다.

〈그림 15〉는 얼룩말과 말, 여름과 겨울의 이미지를 CycleGAN을 통해 변환한 결과를 보여준다. 이 접근 방식은 두 가지 매핑(G: X → Y 및 F: Y → X)과 각각의 적대적 판별자를 포함하며, 이 매핑은 대상 도메인의 실제 이미지와 구별할 수 없는 이미지를 생성한다. 번역 주기(X→Y→X)가 일관되도록 하기 위해 주기 일관성 손실이라는 개념을 도입하여 다른 도메인으로 번역된 이미지가 다시 원래 도메인으로 번역될 때 원본 이미지와 매우 유사해야 함을 보장하고자 한다.[22]

CycleGAN은 적대적 손실(Adversarial Loss)을 두 매핑 함수에 적용하여 매핑 함수 G: X → Y와 그에 대응하는 판별자 DY에 대해, 목적함수는 〈수식 3〉 같이 표현된다. 이 식에서, G는 X 도메인의 이미지 x를 Y 도메인의 이미지로 변환하려고 시도하며, DY는 실제 Y 도메인의 이미지 y와 G에 의해 변환된 이미지 G(x)를 구별하려고 한다. G는 이 목적함수를 최소화하려고 하고, 반면에 DY는 이를 최대화하기에 이러한 적대적 과정은 G가 Y 도메인의 이미지와 구별할 수 없는 이미지를 생성하도록 유도한다. 〈수식 4〉는 주기 일관성 손실(Cycle Consistency Loss)은 이미지가 원래 이미지로 되돌아가기 위해 사용되는 것으로 X 도메인의 각 이미지 x에 대해, 이미지 번역 주기는 x를 원래 이미지로 되돌려야 할 경우 즉, x → G(x) → F(G(x)) ≈ x로 표현하며, 그와 반대로 Y 도메인의 각 이미지 y에 대해 G와 F는 y → F(y) → G(F(y)) ≈ y로서 이러한 행동을 독려하기 위해 사용하는 개념으로 다음의 미니맥스를 최적화하는 것이다.[22]

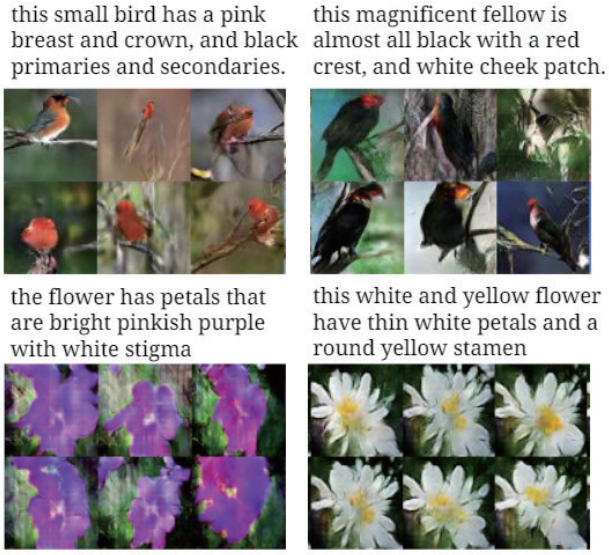
$$G^*, F^* = \arg\min_{G,F} \max_{D_x, D_Y} \mathcal{L}(G, F, D_X, D_Y)$$

〈수식 6〉 미니맥스전략을 이용한 최적화 목적[22]

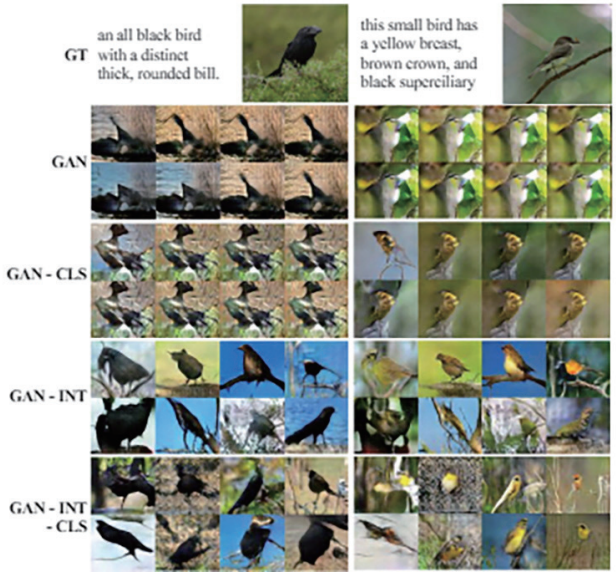
참고로, 이 모델은 UC 버클리의 Jun-Yan Zhu와 한국인 박태성 등이 제안한 모델이다.

5. Generative Adversarial Text to Image Synthesis

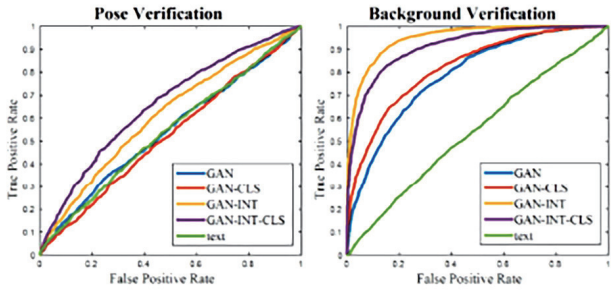
최근 미드저니(Midjourney)와 스테이블 디퓨전(Stable Diffusion) 등이 출현하여 이를 활용한 프롬프터가 고액 연봉을 받으며 가장 각광받는 직업 중 하나로 떠올랐다. 그것의 원류는 바로 미시건 대학의 Scott Read, Zeynep Akata 등이 발표한 'Generative Adversarial Text to Image Synthesis'라고 말할 수 있다. 그들은 GAN을 통해 데이터 셋에서 텍스트 설명에 가장 가까운 이미지를 찾고 유사한 이미지를 생성하는 최초의 연구자가 되었다.



〈그림 16〉 텍스트로 생성된 이미지의 예[4]



〈그림 17〉 제로 샷(zero-shot) 조건의 이미지 생성[4]



〈그림 18〉 ROC 곡선

COCO 데이터 셋으로 훈련된 모델들은 당시로서는 매우 그럴듯한 이미지를 생성했지만, 세부적인 일관성이 부족했다.

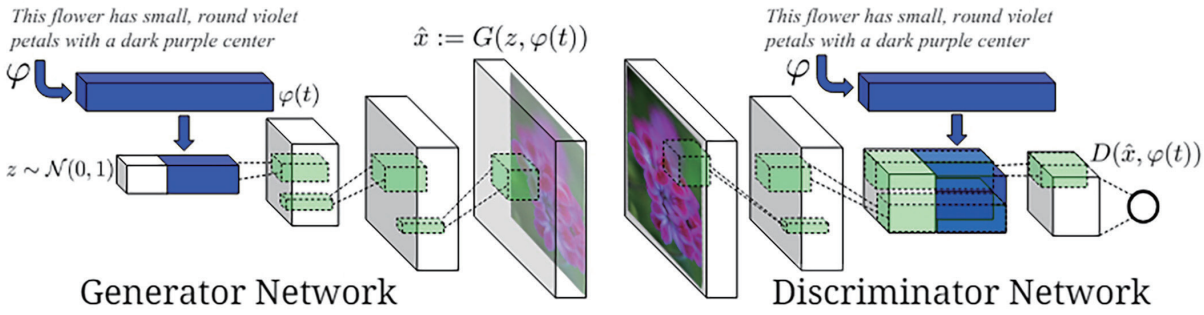
〈그림 16〉은 제로 샷과 훈련 데이터 셋과의 비교 결과를 보여주고 있다. 왼쪽 캡션은 제로샷 카테고리 액션, 오른쪽은 훈련 셋에서 가져온 결과이다. 이들은 주로 꽃과 새의 이미지 생성에 GAN을 사용하였으며, 텍스트의 설명으로부터 가능성 있는 이미지를 자동 생성하는 것에 초점을 맞추었다.[4]

생성기는 노이즈 벡터 z에 연결되고 이어서, 정상적인 디컨볼루셔널 네트워크에서처럼 추론 된다. 생성기 G를 통해 피드포워드하며, 합성 이미지 $\hat{x}$가 $\hat{x} \leftarrow G(z, \phi(t))$를 통해 생성된다. 이미지 생성은 양질의 텍스트와 노이즈 샘플에 조건을 부과한 생성기 G에서의 피드포워드 추론한다. 판별기는 공간 배치 정규화를 따르는 여러 계층의 스트라이드 2번의 컨볼루션을 수행한 후 누설 ReLU를 사용한다.[4]

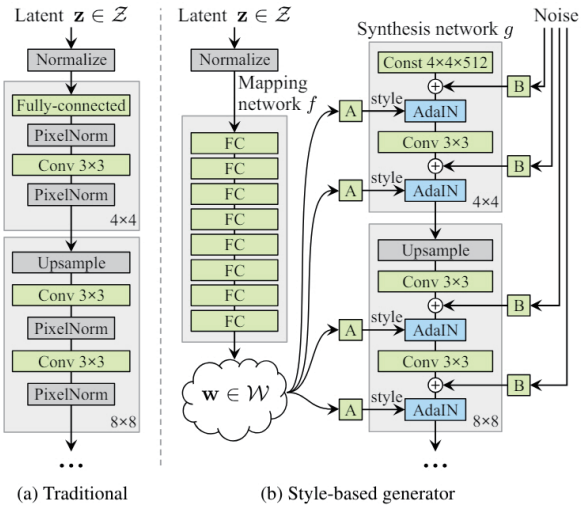
ROC 곡선은 예측된 스타일 벡터 간의 코사인 거리를 이용하여 같은 스타일 혹은 다른 스타일의 이미지 쌍을 평가하는 데 사용된다.

6. 스타일 GAN(StyleGAN)

스타일 GAN은 Tero Karras 등이 2018년에 발표한 논문 'Progressive Growing of GANs for Improved Quality, Stability, and Variation' 일명, PGGAN로 불리는 모델에서 생성기를 개선 시킨 모델이다. Style-based 생성기 부분을 제외하면 나머지 구조는 PGGAN과 동일하다. 따라서, 우선 PGGAN에 대해서 살펴보고자 한다. PGGAN은 4x4 이미지로부터 시작해서 1024x1024까지 점진적으로 레이어를 추가하면서 학습하여 안정적으로 고해상도 이미지를 만들 수 있는 모델이다. 총 9개의 레이어가 있으며, 낮은 해상도에서는 얼굴형이나 머리 스타일 등의 대략적인 특징을 만들고 이후 높은 해상



〈그림 19〉 Text-conditional convolutional GAN 구조[4]



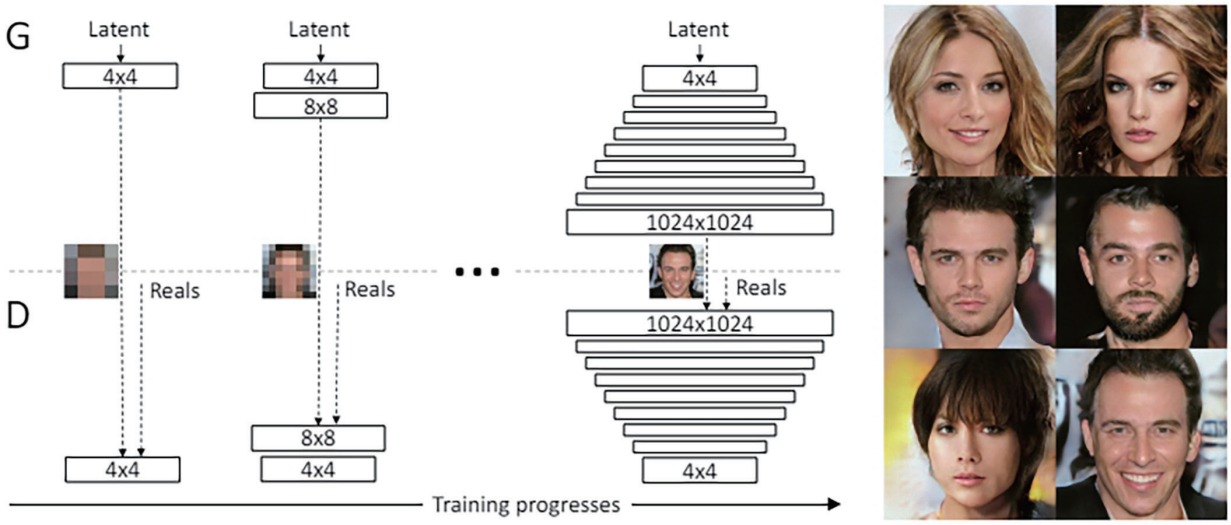
〈그림 20〉 전통적GAN과 StyleGAN 생성기비교

도로 진행됨에 따라 세밀한 특징들을 만들어 간다. 이를 Progressive Growing 구조라고 한다.

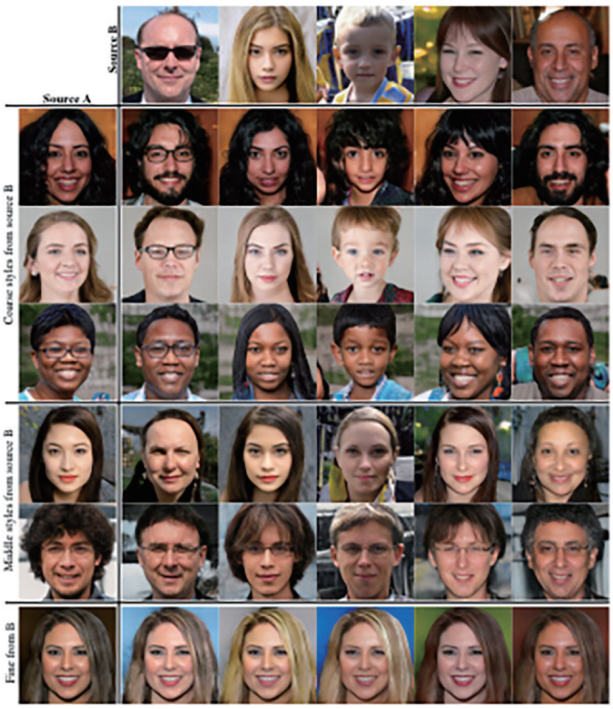
StyleGAN은 PGGAN에 Styled-Based 생성기라는 스타일 제어를 위해 중간 잠재 공간(Intermediate Latent Space)과 적응 인스턴스 정규화(AdaIN, Adaptive Instance Normalization)를 사용하여 이미지를 생성한다. 일반적인 GAN은 잠재 코드(Latent Code)라는 이미지의 특징을 압축된 형태로 표현하는 숫자 벡터를 사용하는데, 이 방식의 제한적인 스타일 제어가 가능했기 때문에 중간 잠재 공간이라는 개념을 도입해 잠재 코드를 변환하여 이미지 생성에 더 적합한 형태로 만드는 역할을

〈그림 22〉 FFHQ(Flicker Faces-HQ) 데이터 세트로 학습된 스타일 기반 생성기를 사용하여 생성된 이미지

수행하는 것으로 이미지의 스타일, 해상도, 특징 등을 제어하는데 사용된다. 매핑 네트워크(mapping network)를 통해서 중간 잠재 공간으로 변환되며, 이는 다시 합성 네트워크(synthesis network)를 통해 실제 이미지로 생



〈그림 21〉 PGGAN의 구조



〈그림 23〉 Style Mixing의 예

성된다. 합성 네트워크는 〈그림 20〉에서 보는 바와 같이 다소 복잡하며, 여러단계의 합성곱 레이어로 구성되어 점진적으로 이미지의 해상도와 디테일을 높여나간다. 여기서, AdaIN을 각 단계마다 거치게 되며, 각 단계에서 잠재 공간 정보를 이미지에 적용하여 스타일을 조정하게 된다. 이러한 복잡한 구조는 학습과 생성 과정에 시간과 에너지 그리고, 가용공간들에서 많은 리소스를 필요로 하게 된다.

$$AdaIN(x_i, y) = y_{s,i}\frac{x_i - \mu(x_i)}{\sigma(x_i)} + y_{b,i}$$

위의 수식은 AdaIN을 나타내고 있다. 학습된 아핀변환 w를 스타일 $y = (y_s, y_b)$로 합성 네트워크 g의 각 컨볼루션 레이어 후에 AdaIN 작업을 제어하는 것으로 각 특징맵 x_i는 별도로 정규화되고, 스타일 y가 변환되고, 편향항이 추가 된다.

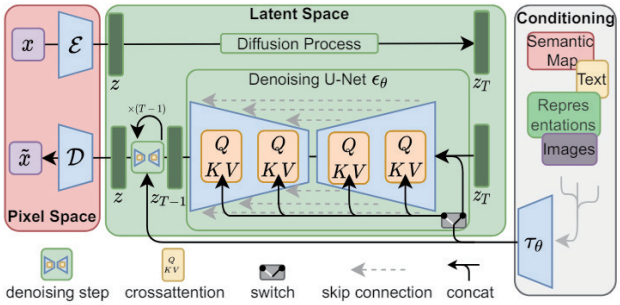
Ⅳ. 최근 연구 동향

앞서 설명한 GAN의 종류 이외에도 미처 설명하지 못한 BigGAN, starGAN, EBGAN, StackGAN, 3DGAN 등 수많은 GANs이 존재한다. NeurIIP 2021에서 Nvidia가 발표한 'Alias-Free Generative Adversarial Networks'를 통해 styleGAN3까지 진화되었으며 이제는 실제 이미지와 식별이 불가능할 정도에 이르게 되었다. 요즘은 한 달이 아닌 한 주가 멀다 하고 새로운 모델들이 쏟아져나오고 있다. 최근 들어 가장 많이 사용하는 미드저니와 스테이블 디퓨전 두 양대 산맥 중 하나인 스테이블 디퓨전의 기본 모델이 되는 확산 모델(Diffusion Model)을 살펴보고자 한다.

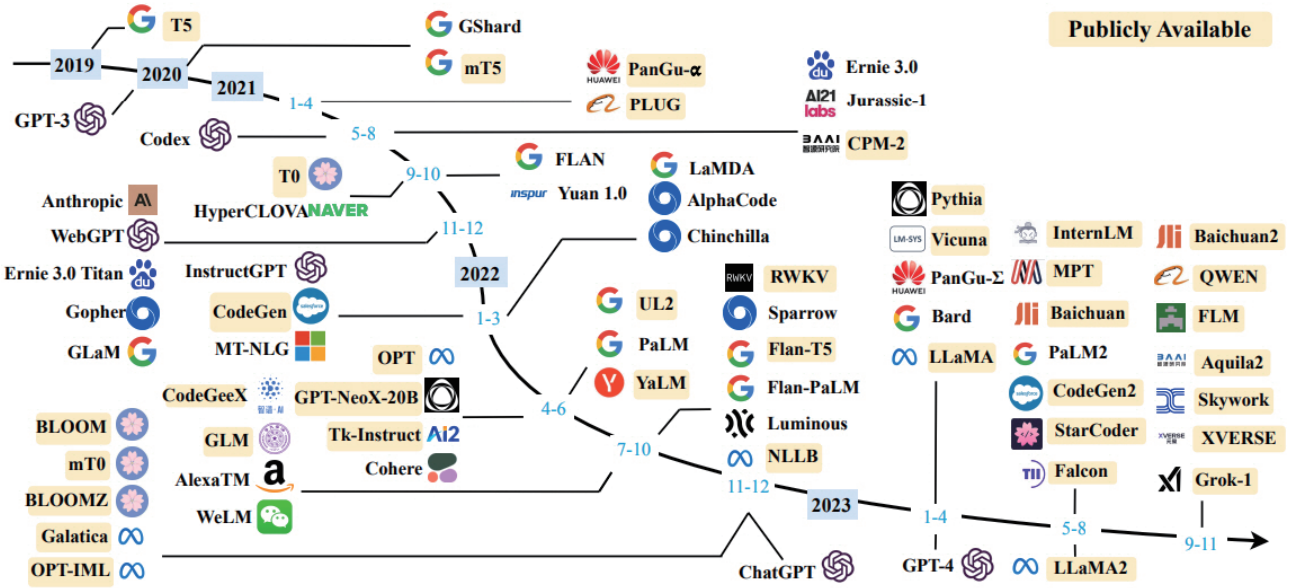
확산 모델(Diffusion Model)은 기본적으로 U-net이라 불리는 오토인코더를 통해서 가우시안 노이즈를 생성하고 제거하는 과정을 통해서 이미지를 생성하는 모델이다. 특히, VAE(Variational AutoEncoder)를 통해서 이미지를 재구성하는데 입력 이미지를 내부의 저차원 잠재 공간으로 인코딩하고, 다시 그 잠재 공간에서 이미지를 재구성하는 방식으로 이 과정에서 이미지의 중요한 특징들을 학습하고, 불필요한 노이즈를 제거해 간다. 이렇듯 VAE의 재구성 과정은 원본 데이터의 주요 특징을 유지하면서 불필요한 변동성을 제거하는 방식으로 진행되며, 이 과정에서의 이미지의 선명도와 디테일을 향상시키게 된다. 또한, 훈련 과정 속에서 이미지 데이터의 확률적 모델을 학습함으로써 이미지의 품질을 조절하면서 현실적이고 자연스러운 이미지를 만들고 있다.[30] 2010년 후반부턴 본격적으로 GAN+VAE를 결합한 모델인 StackGAN, SEGAN[34] 등 다양한 형태로 진화하고 있는 중이다.

$$L_{DM} = E_{x,\epsilon\sim N(0,1),t}[\| \epsilon - \epsilon_\theta(x_t, t) \|_2^2]$$

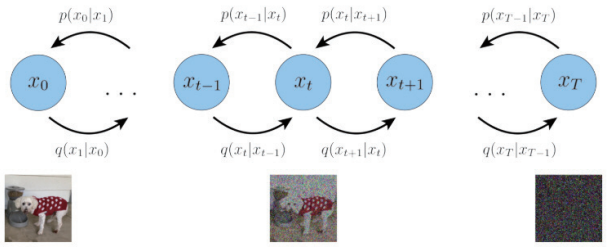
확산모델(Diffusion Model)은 정상 분포된 변수를 점



〈그림 24〉 Latent Diffusion Model



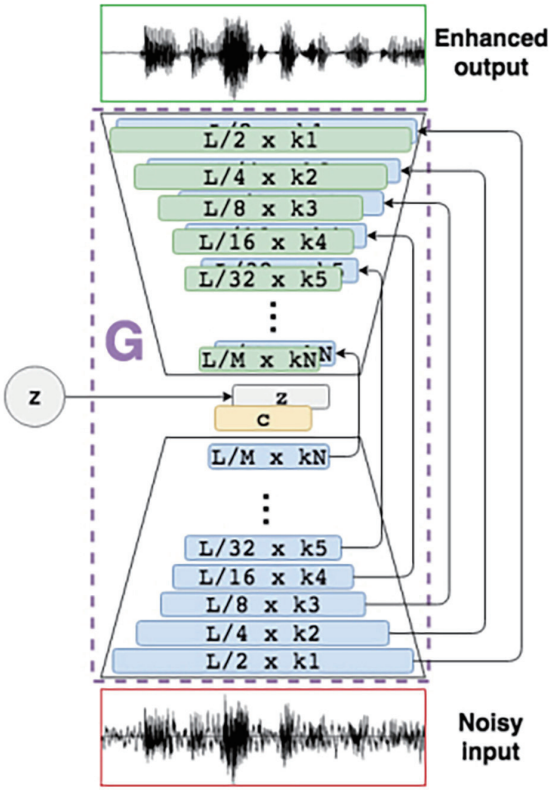
〈그림 25〉 기존 대규모 언어 모델의 시계열 분포[32]



〈그림 26〉 Variational Diffusion Model

진적으로 노이즈 제거하여 데이터 분포 p(x)를 학습하도록 설계된 확률적 모델로서 이는 고정된 마르코프 체인의 역과정을 학습하는 것에 해당된다. 일반적인 확산 모델들은 동등하게 가중된 노이즈 제거 오토인코더의 연속으로 해석할 수 있다.

위 그림에서 x_0, x_T 그리고, x_t가 의미하는 것은 x_0는 자연 상태의 이미지, x_T는 순수한 가우시안 노이즈로 완전히 변화된 상태를 그리고, x_t는 x_0와 x_T 그 중간상태를 의미한다. 따라서, 결국, 위 과정은 $q(x_t \mid x_{t-1})$ 전이 과정을 통해 원본 이미지가 노이즈로의 점진적 변환과 그 역과정에서의 중요한 역할을 수행함을 보여준다. 가우시안의 분산은 일정하거나 확산 과정 중에 변할 수 있으며, 각 단계에서 추가되는 노이즈의 양을 제어한다. 이러한 전이 모델의 변환 과정을 통해 노이즈를 추가하는 과정 속에서 모델링함으로써 자연 이미지를 비롯한 다양한 데



〈그림 27〉 SEGAN의 구조(Encoder-decoder architecture for speech enhancement (G network))

이터 유형의 복잡한 분포를 학습할 수 있게 된다. 이를 통해 원하는 이미지에 가까운 것을 얻을 수 있다.

V. 전망과 결론

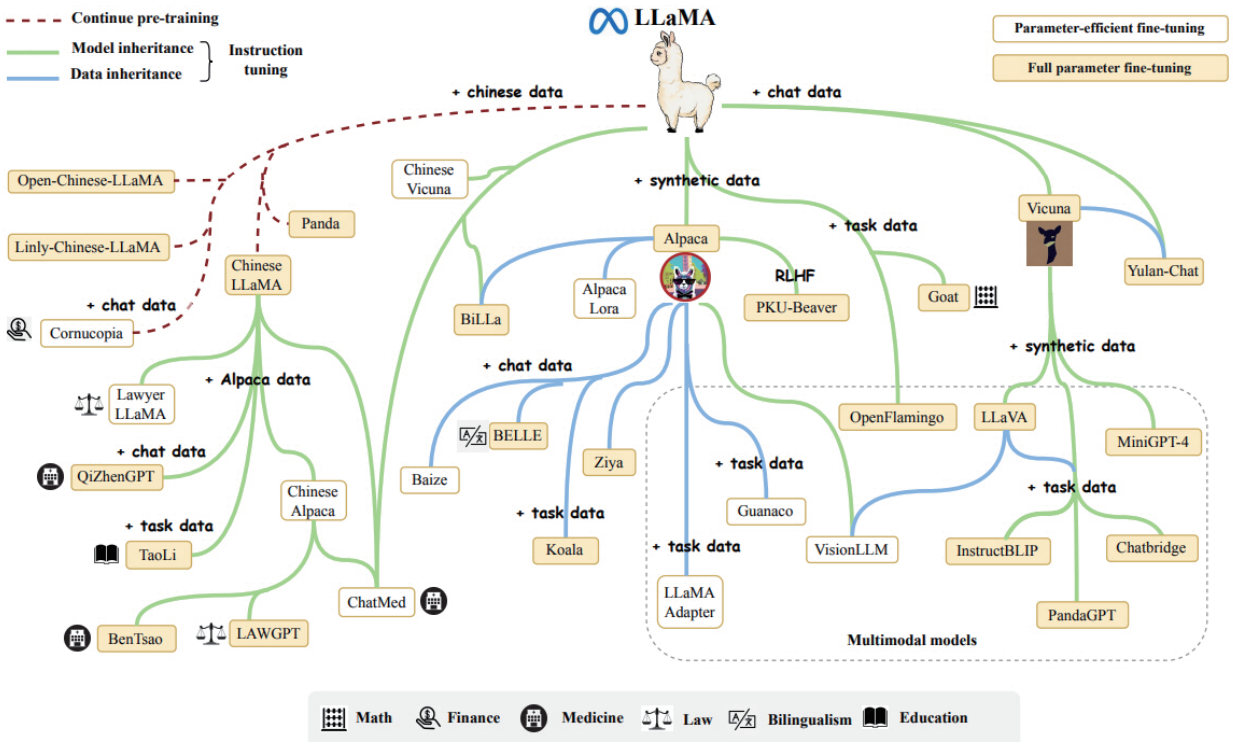
이제는 GANs를 넘어서 거대언어모델(LLM)을 기반으로한 다중모달(Multi-modal)이 트랜드가 되었다. ChatGPT 4.0과 Anthrophic의 Claude3와 최근 발표된 메타(meta)의 LLama3에 이르기까지 다양한 언어모델들이 대부분 다중모달을 지원하고 있거나 지원을 계획 중에 있다. 특히, LLama3의 경우 LLAMA3-8B, n LLAMA3-70B 등의 다양한 모델들을 무료로 제공하고 있으며, 상대적으로 적은 크기의 모델 사이즈를 가지고도 기존의 거대모델과 비슷한 성능을 내고 있다고 알려져 있다.[31]

향후 생성 인공 지능의 전망에 대해서 잠시 고민해 보면 현 시점에서 기대하는 발전 방향은 두 가지로 예상되어 진다. 첫 번째는 더 나은 성능을 위한 모델 개량 내지 새로운 기초 모델의 개발이고, 둘째는 모델의 소형화를 통한 생성 인공 지능의 온디바이스(On-Device)화로 판단된다. 그중에서도 개인적으로 더 끌리는 부분은 바로 소형화를 통한 온디바이스화이다. 그런 면에서 메타의 Llama3의 확장성을 통한 다양화와 타 산업 군에게 미치는 확산 속도 측면에서 인공지능 신산업의 성장 동력으로 작용하는데 더 큰 영향을 기여할 것으로 판단하는 바이다. 그뿐만 아니라, 향후 생성형 인공지능의 성능 향상을 위한 모델 개량과 새로운 기초모델 개발 등에도 많은 학자들이 노력을 기울이고 있기에 조만간 더 우수한 성능의 모델들이 새로운 기초 모델로 자리매김하게 되는 좋은 소식이 들려올 것으로 기대하는 바이다. 생성형 인공지능이 앞으로 우리생활 속에 적극적으로 개입하며 많은 생활상의 변화를 야기할 것으로 예상된다. 부정적 측면과 긍정적 측면이 모두 있겠지만, 법과 제도 그리고, 사회적 합의를 통해 인류가 지속 가능하게 하는 방향으로 기술 발전을 도모해야할 것이다.

참 고 문 헌

[1] Goodfellow Ian, et al. "Generative adversarial nets." Advances in neural information processing systems 27 (2014)

[2] Mehi Mirza, Simon Osindero, "Conditional Generative Adversarial Nets", arXiv.org 6 November 2014



〈그림 28〉 LLaMA에서 수행된 연구 작업의 진화 그래프[32]

https://doi.org/10.48550/arXiv.1411.1784

[3] Scott Reed, Zeynep Akata, Xinchen Yan, Lajanugen Logeswaran, Bernt Schiele, Honglak Lee, "Generative Adversarial Text to Image Synthesis", ICML 2016

[4] Yihan Cao, Siyu Li, Yixin Liu, Zhiling Yan, Yutong Dai, Philip S. Yu, Lichao Sun, "A Comprehensive Survey of AI-Generated Content (AIGC): A History of Generative AI from GAN to ChatGPT", J.ACM, 2018 https://doi.org/10.48550/arXiv.2303.04226

[5] Ashish Vaswan et al. "Attention Is All You Need", NIPS 2017

[6] Minh-Thang Luong, Quoc V. Le, Ilya Sutskever, Oriol Vinyals, and Lukasz Kaiser, "Multi-task sequence to sequence learning", arXiv preprint arXiv:1511.06114, 2015

[7] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio, "Neural machine translation by jointly learning to align and translate", CoRR, abs/1409.0473, 2014

[8] Jianpeng Cheng, Li Dong, and Mirella Lapata, "Long short-term memory-networks for machine reading", arXiv preprint arXiv:1601.06733, 2016

[9] Kyunghyun Cho, Bart van Merrienboer, Caglar Gulcehre, Fethi Bougares, Holger Schwenk, and Yoshua Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation", CoRR, abs/1406.1078, 2014

[10] Junyoung Chung, Çaglar Gülçehre, Kyunghyun Cho, and Yoshua Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling", CoRR, abs/1412.3555, 2014

[11] Cheng-Zhi Anna Huang, Ashish Vaswani, Jakob Uszkoreit, Noam Shazeer, Ian Simon, Curtis Hawthorne, Andrew M. Dai, Matthew D. Hoffman, Monica Dinculescu, Douglas Eck, "Music Transformer: Generating Music with Long-Term Structure"ICLR 2019

[12] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin transformer: Hierarchical vision transformer using shifted windows," in Proceedings of the IEEE/CVF international conference on computer vision, pp. 10012-10022, 2021.

[13] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark, G. Krueger, and I. Sutskever, "Learning transferable visual models from natural language supervision," 2021.

[14] Liunian Harold Li, Mark Yatskar, Da Yin, Cho-Jui Hsieh, Kai-Wei Chang, "VisualBERT: A Simple and Performant Baseline for Vision and Language", arXiv:1908.03557

[15] Z. Wang, J. Yu, A. W. Yu, Z. Dai, Y. Tsvetkov, and Y. Cao, "Simvlm: Simple visual language model pretraining with weak supervision," arXiv preprint arXiv:2108.10904, 2021.

[16] W. Su, X. Zhu, Y. Cao, B. Li, L. Lu, F. Wei, and J. Dai, "VI-bert: Pre-training of generic visual-linguistic representations," arXiv preprint arXiv:1908.08530, 2019.

[17] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," Advances in neural information processing systems, vol. 28, 2015.

[18] L. Zhou, H. Palangi, L. Zhang, H. Hu, J. Corso, and J. Gao, "Unified vision-language pre-training for image captioning and vqa," in Proceedings of the AAAI conference on artificial intelligence, vol. 34, pp. 13041-13049, 2020.

[19] H. Tan and M. Bansal, "Lxmert: Learning cross-modality encoder representations from transformers," arXiv preprint arXiv:1908.07490, 2019.

[20] J. Lu, D. Batra, D. Parikh, and S. Lee, ViLBERT: Pretraining Task-Agnostic Visiolinguistic Representations for Vision-and Language Tasks. Red Hook, NY, USA: Curran Associates Inc., 2019

[21] Alec Radford & Luke Metz, Soumith Chintala, "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks", ICLR 2016

[22] Jun-Yan Zhu, Taesung Park, Phillip Isola, Alexei A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks", arXiv:1703.10593v7

[23] Tero Karras, Samuli Laine, Timo Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks", arXiv:1812.04948v3

[24] Tero Karras, Timo Aila, Samuli Laine, Jaakko Lehtinen, "Progressive Growing of GANs for Improved Quality, Stability, and Variation", arXiv:1710.10196v3

[25] Andrew Brock, Jeff Donahue, Karen Simonyan, "Large Scale GAN Training for High Fidelity Natural Image Synthesis", arXiv:1809.11096v2

[26] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, Alexei A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks", arXiv:1611.07004v3

[27] Calvin Luo, "Understanding Diffusion Models: A Unified Perspective", arXiv:2208.11970v1

[28] Jonathan Ho et al, "Denoising Diffusion Probabilistic Models", NeurIPS 2020

[29] Tero Karraset al, "Alias-Free Generative Adversarial Networks", NeurIPS 2021

[30] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, Björn Ommer "High-Resolution Image Synthesis with Latent Diffusion Models", arXiv:2112.10752v2

[31] Wei Huang et al, "How Good Are Low-bit Quantized LLaMA3 Models? An Empirical Study", arXiv:2404.14047

[32] Wayne Xin Zhao et al, "HA Survey of Large Language Models", arXiv:2303.18223v13

[33] Diederik P. Kingma, Max Welling, "An Introduction to Variational Autoencoders", arXiv:1906.02691v3

[34] Santiago Pascual, Antonio Bonafonte, Joan Serrà, "SEGAN: Speech Enhancement Generative Adversarial Network", arXiv:1703.09452v3

권혁준

- 2008년 8월 ~ 2022년 11월 이에스앤씨 연구소 수석연구원
- 2024년 1월 ~ 국가위기관리학회 사이버대응연구회 위원장 및 편집위원회 이사
- 2022년 12월 ~ 현재 한국폴리텍대학 대전캠퍼스 메타버스콘텐츠과 교수

〈관심 분야〉
Digital Twin, Quantum Computing/ Quantum Error Correction, Metaverse Evolution & Direction

ChatGPT와 거대 언어 모델

Ⅰ. 서 론

"자연은 살아있는 물질을 구조화하는 단 하나의 방법만을 발견했다. 그러나, 아직 자연에서 전혀 발생하지 않은 더 간단하고, 더 유연하며, 더 빠른 또 다른 방법이 있다. 나는 오늘, 생명의 발전을 가져올 수 있는 이 두 번째 방법을 발견했다."[1] 1920년에 쓰여진 카렐 차페크의 희곡 'RUR (로숨의 유니버설 로봇)'에 나오는 대사이다. 한 세기 전에 쓰여진 작품에서 처음으로 '로봇'이라는 단어가 사용되었을 때, 사람들은 인공적인 심장이 혈관을 따라 피를 순환시키는 모습을 상상했을지언정 코일과 배터리가 전선을 따라 전류를 순환시키는 모습을 떠올리진 못했을 것이다. 차페크의 작품 속에서 인공 로봇은 인간 사회 속에서 살아가며, 그는 이로 인해 발생할 수 있는 윤리적, 철학적 질문들을 남긴다. 오늘날, 차페크의 상상이 현실의 일부가 되었으며, 거대 언어 모델(large language model, LLM)은 그 상상을 초월하여 컴퓨터 안에서 동작하는 인공지능으로서 인간과 유사한 수준으로 언어를 이해하고 생성할 수 있는 능력을 보여주고 있다.

〈그림 1〉은 멀티모달 (multimodal) 인공지능인 DALL · E 3[2]를 활용하여 생성한 'RUR'의 상상도이다. 사용자가 제공하는 텍스트 설명인 프롬프트(prompt)에는 '카렐 차페크의 희곡 RUR의 상상도를 그려줘'가 입력되었으며, 해당 작품에 관련한 어떠한 정보도 프롬프트에는 제공되지 않았다. 그럼에도 불구하고 인공지능은 해당 글자 뭉치에서 문맥을 이해한 것처럼 희곡 RUR의 모습과 같이 인공적으로 생성된 인간과 같은 로봇을 대량으로 생산하는 미래 세계를 배경으로 한 그림을 그려낼 수 있었다. 이는 인간이 창조해 낸 언어 논리와 예술 작품을 인공 장치인 반도체의 전자 배치와 그 장치를 제어하는 정보 통신의 기술로 표현할 수 있음을 보여주는 대표적인 사례이다.

김동옥
삼성전자

〈그림 1〉 인공지능이 생성한 카렐 차페크의 희곡 'RUR'의 상상도 이미지

차페크의 작품에서도 볼 수 있다시피 인공지능에 대한 개념은 20세기 초반에서도 확인할 수 있다. 인공지능에 대한 본격적인 학술 연구와 개발은 앨런 튜링과 플랑크 로젠블라트의 논문이 출간된 1950년대부터 본격적으로 시작되었다고 할 수 있다[3-5].

튜링은 '기계가 생각할 수 있는가 (Can machine think)?'라는 질문을 제기하여 기계의 지능을 평가하는 방법으로 '튜링 테스트 (Turing test)'를 제안했다. 튜링 테스트는 인간 평가자가 기계와 인간 참가자와 각각 대화를 하면서 어느 쪽이 기계인지를 판단하는 실험이다. 만약 대화를 통해서 상대방이 기계인지 인간인지 판별할 수 없다면, 해당 기계는 '지능적'으로 대화하는 것으로 간주하고, 인공지능이라고 말할 수 있다.

로젠블라트는 사람의 신경망을 모사하여 인공신경망 (artificial neural network)를 만들었고, 데이터의 입력과 처리, 그리고 출력의 구조를 만들어 '퍼셉트론 (perceptron)'이라 이름지었다. 퍼셉트론은 다양한 학문에서 영향을 받아 기계 학습 (machine learning)의 핵심 모델인 합성곱 인공신경망 (convolutional neural network, CNN)을 탄생시키는 초석이 된다.

인공지능 연구는 2016년 Google DeepMind Challenge Match를 통해 등장한 알파고(Alpha Go)[6]와 2022년 OpenAI를 통해 혜성처럼 등장한 ChatGPT[7] 이후 대중의 삶에 깊고 넓게 녹아들었다. 인공지능 연구자들이 튜링 테스트를 통과하고자 부단히도 노력하던 과거의 발자국들이 이제는 인공지능에 문외한 사람들도 ChatGPT와 같은 거대 언어 모델이 인간과 유사한 수준의 응답을 구현할 수 있다는 것에 쉽게 동의할 수 있도록 만들어 주었다.

이제 거대 언어 모델은 일상 속에서 쉽게 접할 수 있게 되었고, 빅테크 대기업들은 속속들이 자신의 모델을 발표하고 있다. OpenAI의 ChatGPT 외에도 Google의 BERT (Bidirectional Encoder Representations from Transformers)[8], PaLM (Pathways Language Model)[9, 10], Gemini[11, 12], Meta의 LLaMA (Large Language Model Meta AI)[13-15], 그리고 Anthropic의 Claude[16, 17] 등이 대표적인 사례이다. 국내에서도 Naver CLOVA의 HyperCLOVA[18]나 KakaoBrain의 KoGPT[19] 등과 같이, OpenAI의 GPT-3[20] 모델에 한국어 데이터를 학습시킨 모델을 공개했다.

거대 언어 모델의 춘추 전국 시대에 이 모든 모델을 살펴보기는 어렵지만, 거대 언어 모델의 부흥을 이끌었던 OpenAI의 GPT (Generative Pre-trained Transformer)를 중심으로 거대 언어 모델의 역사와 발전 과정을 기술하고자 한다.

Ⅱ. 거대 언어 모델의 발전 과정

언어는 인위적인 개념으로, 인류의 발전 과정에서 만들어진 추상적인 체계이다. 언어는 자연에 존재하지 않지만 기호로써 표현될 수 있다. 이 체계는 같은 의미라도 다른 기호로 표현될 수 있으며, 각 기호에 따라 서로 다른 규칙을 갖기도 하고, 규칙과 표현이 시간이 지남에 따라 달라지기도 한다. 또한 주어진 규칙과 기호 표현에 대해서 무한히 많은 의미를 창조할 수도 있다.

이러한 언어의 특성은 언어를 엄밀한 규칙으로 정의하기 어렵게 만들며, 특히 주어진 절차대로 명령을 수행하는 컴퓨터 프로그램을 이용한 언어 모델링도 굉장히 어렵게 만들었다.

언어 모델의 역사는 20세기 중반으로 거슬러 올라간다. 1960년대 MIT의 컴퓨터 과학자 조셉 와이젠바움이 개발한 ELIZA는 초기 컴퓨터 기반의 언어 모델의 한 형태로 볼 수 있다[21]. ELIZA는 사용자가 제공하는 텍스트 설명에서 키워드를 추출하고, 이를 기반으로 문맥을 유추

하여 적절한 반응을 생성하는 챗봇(chatbot) 프로그램이다. ELIZA의 대화 방식은 심리치료 상담을 모방하여 사용자가 스스로 대화를 이어나갈 수 있도록 되묻는 반응을 하도록 고안되었다. 비록 ELIZA는 복잡한 언어 이해를 실현하지는 못했지만, 컴퓨터가 인간의 언어를 어느 정도 모방함으로써 언어 모델의 구현뿐만 아니라 초기 인공지능 대화 시스템의 발전에 기여한 중요한 연구 사례이다.

비슷한 시기에 발전한 n-gram 모델은 정보 통신과 수리적 확률 모델을 기반으로 유의미한 문장을 생성했다[22]. n-gram은 문장을 구성하는 단어를 파편화하여 n 번째로 올 단어를 앞선 n-1개의 단어의 나열을 통해서 유추한다. 이 방식은 전체 문장에 대한 이해 없이 문장의 지역적인 특성으로만 새로운 문장을 생성하였으며, n-gram을 기반으로 하는 모델들은 스펠링 검사, 자동 완성, 텍스트 요약 등의 간단한 자연어 처리 응용 분야에서 활용될 수 있었다.

이러한 문장, 혹은 데이터 시퀀스의 지역적인 특성은 20세기 인공 신경망의 발전과 더불어 순환 신경망(recurrent neural network, RNN) 형태로 발전했다. 인공 신경망이 주어진 입력과 출력을 통해서 학습한다면, 순환 신경망은 과거의 입력이 현재의 입력과 출력에 영향을 주는 방식으로 동작한다. n-gram 모델에 문맥 학습을 포함한 것이다. 이는 언어의 기호 표현에 존재하는 규칙이 문맥을 형성하는 것과 유사하다. 그러나 순환 신경망처럼 순차적으로 데이터를 처리하는 경우, 데이터 시퀀스가 길어짐에 따라 학습에 많은 시간이 소요되고, 신경망을 따라 전달되는 학습 오차가 너무 작아지거나 커져서 학습 효과가 떨어지는 한계가 있었다.

이처럼 시퀀스 내 간격이 먼 데이터 간 정보가 전달되지 않는 현상을 장기 의존성(long-term dependency)라고 한다. 20세기 말, 호흐라이터와 슈미트후버는 순환 신경망의 장기 의존성 문제를 해결하고자 순환 신경망에 게이트 메커니즘을 추가하여 정보의 흐름을 제어하였다[23]. LSTM(long short-term memory)으로 불리는 이 모델은 RNN의 한 형태로, 입력, 망각, 출력의 3개 게이트를 추가했다. LSTM은 유의미한 정보가 담긴 셀(cell)과 출력 값을 구분하여 새로운 입력 값과 과거의 출력 값을 통해 셀의 정보를 망각하거나, 입력하거나, 혹은 출력 값을 생성할 때 가져올 수 있었다. 이를 통해 불필요한 정보는 차단하고 필요한 정보는 유지하도록 모델을 학습했다.

LSTM은 RNN의 학습 오차 문제를 해결하여 긴 시퀀스를 잘 학습할 수 있었으나, 모델 구조가 복잡해지고 학습 시간이 길어 계산 비용이 높은 문제를 여전히 갖고 있었다. 시퀀스 내 데이터의 상관 관계를 순차적으로 계산하여 병목 현상이 발생했던 것이다. 이는 병목 현상은 상관 관계를 병렬로 계산하는 Attention 메커니즘이 등장하면서 자연스럽게 해결되었다[24-25].

Attention 메커니즘은 새로운 데이터 입력 이전의 데이터로부터 생성된 셀과 출력 값만을 고려하는 것이 아닌, 앞선 모든 셀의 출력 값을 다음 셀의 값을 계산할 때 사용함으로써 현재의 데이터가 모든 이전의 데이터와의 상관 관계를 각각 병렬로 계산할 수 있었다. 과거의 데이터 중에 어떤 데이터에 주의 (attention)를 기울일 필요가 있는지를 계산하여 데이터 간 상관 관계를 통해 주어진 데이터 시퀀스의 문맥을 형성하는 핵심 키워드를 추론할 수 있었다.

2017년에 공개된 Transformer는 Attention 메커니즘의 계산 대상을 자기 자신을 포함하여 입력 시퀀스 내의 모든 데이터에 적용하도록 확장한 Self-Attention 메커니즘 기반으로 한 아키텍처이다[25]. Self-Attention을 여러 번 수행하는 Multi-Head Self-Attention 기능은 입력 시퀀스 내 각 데이터 간의 관계를 동시에 다양한 관점에서 평가할 수 있다. 이를 통해 모델은 문장 내의 장기 의존성을 해결하고, 복잡한 문맥적 관계를 이해할 수 있었다. Multi-Head Self-Attention은 한 번의 Attention 과정에서 놓칠 수 있는 정보를 포착할 수 있었고, 서로 다른 Self-Attention 'head'는 서로 다른 언어적 의미에 대한 정보를 계산할 수 있었다. 예를 들어, 한 head는 문법적 관계를, 다른 head는 의미적 연결을, 또 다른 head는 특정 문맥에 따른 뉘앙스를 포착할 수 있다. 이러한 정보는 모두 결합되어 모델이 더욱 풍부하고 다양한 문맥적 이해를 기반으로 입력 시퀀스를 처리할 수 있도록 했다.

Transformer는 Multi-Head Self-Attention 메커니

즘을 통해 병렬 처리의 이점을 최대화하고, 이러한 혁신적인 접근법은 자연어 처리의 한계를 뛰어넘는 새로운 가능성을 열었다. 기존의 제한적인 언어 모델의 성능을 크게 향상시켜 언어 모델의 연구와 응용에 깊은 영향을 미쳤다.

언어 모델이 발전하고 있는 반 세기 동안, 언어 모델을 개발하고 평가하는데 사용된 데이터와 연산 처리 장치에도 많은 발전이 있었다. 초기의 언어 모델은 제한된 계산 자원으로 인하여 비교적 간단한 모델과 소규모 데이터에 국한되어 있었다. 그러나 컴퓨터 과학의 발전이 효율적인 아키텍처와 정보 처리 기법을 이끌 수 있었고, 그래픽 처리 장치(graphics processing unit, GPU)의 등장과 더불어 연산 처리 장치의 눈부신 성장으로 인하여 언어 모델 분야에서도 점점 더 거대한 모델과 많은 양의 데이터를 다룰 수 있게 되었다.

분산 컴퓨팅 시스템과 클라우드 컴퓨팅 서비스의 발전은 언어 모델 개발에 필요한 계산 자원을 보다 쉽게 접근하고 활용할 수 있게 만들었다. 이러한 시스템을 통해 여러 대의 컴퓨터가 네트워크를 통해 연결되어, 복잡하고 거대한 모델 학습 작업을 여러 개의 처리 단위로 나누어 병렬로 처리할 수 있었다.

이러한 복합적인 기술의 발전은 고성능 하드웨어 위에서 대규모 언어 데이터를 학습하는 Transformer 기반의 거대 언어 모델이 등장할 수 있는 발판이 되어 주었다.

III. ChatGPT의 등장

1. Generative Pre-trained Transformer (GPT)

언어 모델에서 긴 입력 시퀀스를 처리할수록 시퀀스 내 데이터의 다양한 상관 관계를 통해서 문법, 의미, 문맥 등 다양한 추상적인 언어 표현 특성을 이해할 수 있었다. 그러나 거대한 모델을 매번 학습할 수는 없었고, 언어 표현 데이터는 그 특성상 의미에 대한 정답이 있는 데이터의 양은 매우 빈약했다.

Transformer 논문이 공개된 후의 이듬해, 2018년, OpenAI는 이러한 상황을 고려하여 언어 표현을 생성하는 (generative) 모델을 정답이 없는 대규모 데이터셋에서 사전 학습 (pre-trained) 한 Transformer를 만들어 발표했다[26]. 이 언어 모델이 'GPT-1 (Generative Pre-trained Transformer 1)'이다.

GPT-1은 텍스트 데이터의 특성을 살려 BooksCorpus와 1B Word Benchmark 데이터를 활용하여 비지도 사전 학습을 진행하였다. BooksCorpus는 다양한 장르의 7,000권이 넘는 도서로써 긴 글에 대한 문맥을 이해하는데 도움을 주었고, 1B Word Benchmark 데이터는 텍스트를 문장 단위로 자르고 임의로 섞은 데이터로써 다양한 패턴을 학습할 수 있도록 했다. 이 때, 비지도 사전 학습은 시퀀스 데이터를 학습하듯 주어진 n개의 연속된 단어에 대해서 그 다음에 올 n+1번째 단어를 예측하는 방식이다.

GPT-1은 사전 학습 이후에 좀 더 효과적으로 작업을 수행할 수 있도록 정답이 있는 데이터에 대해 모델을 미세 조정 (fine tuning) 하였다. 언어 모델이 수행할 수 있는 작업들로는 문장 논리 추론, 의미에 대한 질의, 유사도 판단, 텍스트 분류 등이 있다. 만족스러울 수준은 아니었지만, Transformer를 기반으로 한 GPT-1은 Transformer 이전의 게이트 메커니즘을 기반으로 한 LSTM 등의 모델보다 다양한 언어 작업에 대해서 괄목할만한 성과를 보여주었다. 그러나 GPT-1은 여전히 계산 비용이 높고 미세 조정이 어려운 단점이 있어 Google이 비슷한 시기에 공개한 BERT에 밀려 각광을 받지는 못했다.

BERT도 GPT와 마찬가지로 Transformer 아키텍처를 기반으로 한 언어 모델이다. BERT는 GPT와는 달리 양방향성 (bidirectionality)을 적용하여 주어진 텍스트의 좌우 문맥을 고려함으로써, 단어의 의미를 더욱 정확하게 파악할 수 있었고, 여러 세부 작업들에서 GPT-1보다 더 뛰어난 성능을 보여주었다.

BERT는 마스크된 언어 모델(masked language model)을 통해 데이터 시퀀스를 순차적으로 예측하는 것이 아닌, 임의의 위치에서 예측하는 방법으로 수행했다. 주어진 텍스트에 대해 15%의 단어를 임의로 선택해서 마스킹(masking)을 하고 해당 위치에 들어갈 단어를 예측하게 하는 방법이다. 특히, 15%의 단어 중 대부분은 마

스킹을 하고, 일부 소수의 데이터는 랜덤으로 변경하거나 혹은 변경하지 않는 방법으로 예측 난이도를 높였다.

학습 데이터도 BooksCorpus 외에 그보다 3배 이상의 단어 개수를 가진 Wikipeda를 이용하였다. BERT는 모델 신경망에 여러 작업에 특화된 층을 추가하는 방식으로 쉽게 다양한 작업에 미세 조정을 할 수 있었다.

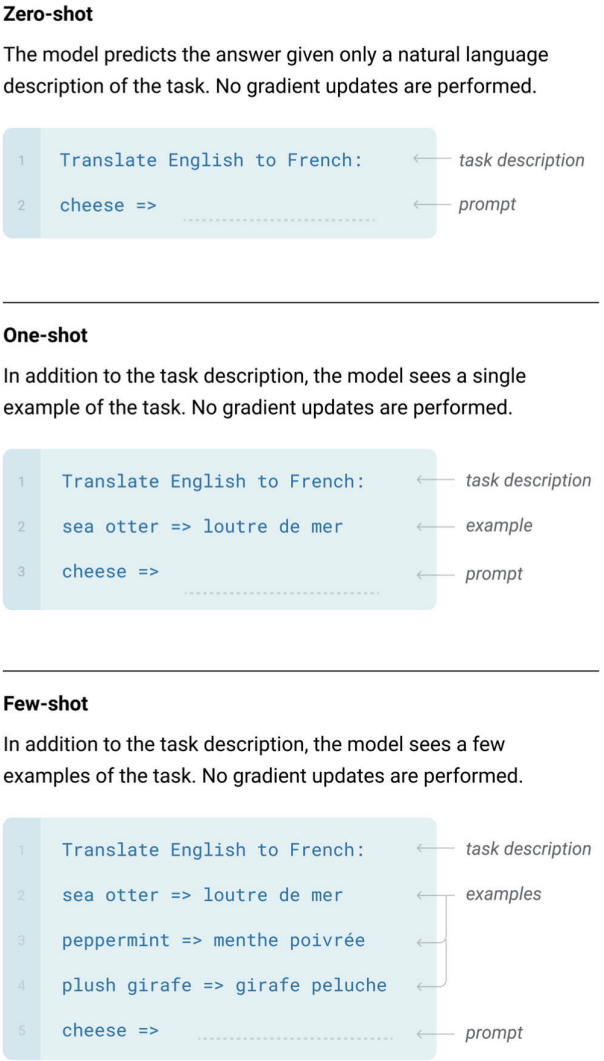
BERT의 등장은 OpenAI에게 위기이자 동시에 새로운 기회가 되었다. 2019년 초, OpenAI는 GPT-1의 후속 모델로 GPT-2를 공개했다[27]. GPT-2는 특정 작업의 지도 학습이 유능한 제너럴리스트 (competent generalist) 보다는 편협한 전문가 (narrow expert) 같은 언어 모델을 만든다고 주장했다. 특정한 작업에서 조금만 변화가 생겨도 성능이 대폭적으로 하락한다는 의미이다.

OpenAI는 이를 보완하고자 소셜 미디어 플랫폼인 Reddit을 스크랩하여 연결된 웹 페이지를 수집해 가공하고, 학습에 활용할 40 GB에 해당하는 텍스트 데이터를 정리했다. GPT-2는 WebText라고 명명된 이 방대한 텍스트 데이터를 비지도 학습 방법으로 학습하고, 지도 학습 형태의 미세 조정 없이 다양한 작업에 바로 적용되었다. 어떤 작업을 수행할지는 모델의 입력값으로 받을 수 있도록 설계되었다. 사용자는 프롬프트를 통해 입력할 수 있었다.

이러한 변화는 GPT-2의 논문 제목 'Language models are unsupervised multitask learners'에서도 알 수 있다. 비지도 학습 방식으로 다양하고 방대한 데이터를 학습한 GPT-2는 다양한 작업에 더욱 복합적으로 대응할 수 있어 Wikipeda, Books, Newspaper 등으로 학습한 당시의 언어 모델들보다 훨씬 효과적인 성능을 보여주었다. 그러나 작문 등에서 좋은 결과를 보여준 반면, 질의 응답이나 독해 등 미세 조정을 한 모델에 비교되는 결과를 보여주기도 하였다.

OpenAI는 비지도 학습에 더해 프롬프트의 문맥을 학습할 수 있는 모델로 GPT-3를 공개했다[20]. 이 때, GPT-3의 논문 제목은 GPT-2와 유사한 'Language models are few-shot learners'로 작성되었으며, 여기서 few-shot은 적은 양의 문맥 학습이 포함된다는 의미이다. 문맥 학습 과정에서 예시가 포함되지 않는 경우는 zero-shot, 하나의 예시가 포함된 경우는 one-shot, 그리고 적은 수의 예시가 포함된 경우를 few-shot이라고 부른다.

GPT-3의 few-shot은 미세 조정을 대체하고, 모델의 변수를 업데이트 하거나 바꾸지 않는다. 〈그림 2〉는 GPT-3 논문에서 보여준 문맥 학습의 예제이다[20]. 어떤 작업을 해야 할지, 그리고 그것에 대한 예시가 무엇인지를 예시의 개수를 달리하여 보여준다. GPT-3의 few-shot을 이용한 문맥 학습은 미세 조정 없이도 언어 모델이 다양한 작업에서 더욱 뛰어난 성능을 보여주는 것을



〈그림 2〉 GPT-3에서 소개한 문맥 학습 예제. 영어를 불어로 번역하면서 0개 (zero-shot), 1개 (one-shot), 3개 (few-shot)의 예시를 포함하였다.

확인시켜 주었으며, 세 자릿수의 덧셈과 뺄셈, 미국 대학 입학 능력 평가 시험인 SAT Analogies, 그리고 나아가 기존에 존재하지 않았던 신조어를 활용하는 작업에서까지 준수한 성능을 보여주었다.

GPT 모델이 발전하면서 모델의 규모 및 학습 데이터의 양이 증가했다. GPT-3의 경우는 1,750억개의 변수로 구성된 모델과 570 GB에 달하는 데이터를 기반으로 학습하였다. 이는 GPT-2에서 활용되었던 잘 정제된 WebText 데이터뿐만 아니라 상대적으로 질이 떨어지는 텍스트 데이터인 CommonCrawl 데이터를 포함하였다. 대신 질이 높은 텍스트 데이터와 섞는 식으로 데이터의 질과 양, 그리고 다양성을 확보하였다.

그러나 모델과 데이터 규모의 증가는 느린 학습 속도와 많은 연산 비용 문제를 여전히 해결하지 못했고, 이러한 비용에 비해 독해 등 특정 작업에서는 여전히 성능이 확보되지 못했다. 게다가 정제되지 않은 수많은 데이터가 학습에 사용되면서 GPT-3는 성별, 인종, 종교 등 여러 가지 사회적인 분야에 대해 편향된 응답을 보여주었다.

Transformer 아키텍처의 등장으로 데이터 시퀀스의 깊은 이해를 기반으로 하는 언어 모델이 만들어 질 수 있었다. 언어 모델 학습을 위한 텍스트 데이터는 그 특성상 의미에 대한 정답이 있는 데이터를 충분히 확보하는 것이 어려웠고, 이로 인하여 언어 모델은 지도 학습 대신 비지도 학습으로 성능이 개선되었다. 비지도 학습은 많은 양의 데이터를 활용할 수 있게 하였지만 언어 모델이 실시간으로 학습하기에는 계산 비용이 커서 사전 학습을 활용하게 되었다. 게다가 학습에 사용된 데이터의 질을 전부 관리하기도 어려웠다. 비록 정제되지 않은 데이터로 인하여 언어 모델이 편향되거나 잘못된 응답을 생성할 수는 있었지만, 언어 모델은 주어진 프롬프트에 대해 '그럴 듯한' 응답을 생성하는 것은 가능했다. 이러한 배경에서 Transformer 기반의 사전 학습된 언어 생성 모델인 GPT가 개발되었다.

〈표 1〉은 앞서 소개한 4개의 거대 언어 모델의 출시 년도, 모델 변수 개수, 그리고 학습 데이터 규모를 비교한다. 후속 모델일수록 기하급수적으로 거대해진 모델 규모와 학습 데이터 규모를 확인할 수 있다. 거대 언어 모델의 성능이 점점 좋아진 것은 단순히 모델의 구조 및 학습 방법의 개선뿐만 아니라 규모 그 자체의 효과도 무시할 수 없다.

〈표 1〉 주요 거대 언어 모델의 출시 년도, 모델 변수 개수, 학습 데이터 규모 비교.

모델 이름	출시 년도	모델 변수 개수 (억개)	학습 데이터 규모 (GB)
GPT-1 [26]	2018	1	5
BERT [8]	2018	3	–
GPT-2 [27]	2019	15	40
GPT-3 [20]	2020	1750	570

하지만 단순히 모델과 데이터의 규모를 키우는 것은 학습 비용과 데이터 정제라는 언어 모델의 고질적인 문제와 직결된다. 정답이 부족하고 질이 떨어지는 데이터만으로도 인간이 구사하는 언어와 유사한 결과를 만들 수 있는 모델의 필요성이 제기되었다.

2. ChatGPT

GPT-3의 데뷔 이후, OpenAI는 프롬프트를 좀 더 효과적으로 활용할 수 있는 모델을 연구하기 시작했다. GPT-3가 학습한 방대한 양의 데이터는 문맥에 따른 언어 작업을 훌륭하게 수행했지만, 사용자의 의도를 읽어내지 못하고 거짓되거나 편향된, 혹은 유해한 응답을 생성하는 경우가 있었다. OpenAI는 이런 불일치를 해결하고자 GPT-3를 세 가지 기술로 미세 조정한 새로운 모델, InstructGPT를 공개했다[28]. InstructGPT는 GPT-3 이후 2년 만인 2022년에 공개되었으며, 사용자의 의도와 언어 모델의 결과를 일치시키고자 사람이 직접 모델 학습 과정에 참여하였다. 이 방법을 RLHF (reinforcement learning from human feedback)라고 부른다.

GPT-3의 미세 조정에 사용된 3가지 기술은 다음과 같다. 첫 번째는 SFT (supervised fine-tuning)으로, 주어진 프롬프트에 대해 사람이 직접 기대하는 응답을 작성하는 것이다. 이렇게 정제되어 만들어진 프롬프트와 응답 데이터는 지도 학습 과정으로 GPT-3를 직접 미세 조정한다.

두 번째는 RM (reward model)로, 주어진 프롬프트

와 SFT로 만들어진 여러 개의 견본 응답에 대해 사용자가 선호도를 평가하는 것이다. 이 때 선호도는 보상(reward)으로 주어진다. 응답의 선호도는 RM을 학습하는 데이터로 활용된다.

세 번째이자 마지막 기술은 PPO (proximal policy optimization)으로, 주어진 프롬프트에 대한 응답을 RM으로 평가하여 높은 선호도의 응답이 출력되도록 모델을 최적화하는 것이다. 이는 보상을 기반으로 하는 강화 학습 (reinforcement learning)과 원리가 같다. RM의 학습과 PPO는 반복적으로 수행되어 점점 더 정교한 언어 모델을 만들어 나갈 수 있다.

정리하자면, InstructGPT는 주어진 프롬프트를 보고 그에 대한 응답을 추론한다. 이 응답을 RM이 평가하여 높은 선호도의 응답이 나오도록 모델이 강화 학습 기반으로 미세 조정된다. 여기서 미세 조정에 사용되는 응답과 그에 대한 선호도가 사람에 의해서 결정된다. 사람의 지시 (instruction)가 보상으로써 모델의 응답에 반영되는 것이다.

이 방법으로 미세 조정한 결과, InstructGPT는 GPT-3에 비해 더 진실된 응답을 생성하는 데는 성공하였다. 그러나 편향되거나 유해한 출력 시퀀스에 대해서는 GPT-3와 비교하여 크게 개선되었다고 말할 수 있는 결과를 확인할 수 없었다.

OpenAI는 GPT-3.5를 기반으로, InstructGPT를 학습하는데 사용한 방법을 적용해 '대화'에 특화된 모델인 ChatGPT를 InstructGPT 논문이 공개된 지 약 10개월 만인 2022년 11월 30일에 서비스 형태로 공개하였다[7]. GPT-3.5는 GPT-3 이후로 만들어진 GPT 모델이고, ChatGPT의 학습에는 SFT, RM, PPO가 적용되었다.

InstructGPT의 학습 방식과의 차이점은 SFT를 적용할 때, 사람이 프롬프트와 응답을 합친 '대화'를 모두 작성하도록 하여 지도 학습 데이터를 준비하였고, 이를 InstructGPT 데이터와 섞어서 데이터의 양과 질을 동시에 확보했다. 또한 RM을 학습할 때도 사람과 챗봇의 대화를 준비하였으며, 모델이 작성한 응답과 사람이 작성한 응답을 임의로 선별하여 PPO로 모델을 미세 조정하였다.

ChatGPT의 모델이나 학습 데이터에 대해서는 공개되지 않았지만, ChatGPT는 등장 이후 대중들에게 많은 관심을 받았다. 이로 인하여 언어 모델의 능력과 한계가 사회 속으로 급속도로 전파되기 시작했다. OpenAI는 ChatGPT를 출시한지 두 달 만에 1억명이 넘는 사용자를 확보했으며 ChatGPT Plus 서비스를 통해 수익화를 시작했다[29].

ChatGPT는 기존의 언어 모델들에 비해 많이 개선됐음에도 불구하고, 여전히 고질적인 문제점들을 갖고 있었다. 그럴듯하지만 만들어진 사실, 즉 거짓된 응답을 하는 환각 (hallucination) 문제, 사용자의 의도가 동일하더라도 프롬프트의 변화에 따라 응답의 질이 크게 달라지는 프롬프트 의존 문제, 필요 이상으로 장황하게 응답하는 문제, 적절하지 않은 응답을 유도했을 때 여전히 편향되거나 유해한 응답을 하는 문제, 그리고 응답에 관한 내용이 2021년 9월까지 공개된 데이터에 한정된다는 문제가 남아있었다.

3. GPT-4V(ision)

ChatGPT는 연구 목적 외에 상업적 목적이 있는 언어 모델 서비스였다. OpenAI는 ChatGPT를 개선하면서도 모델에 대한 자세한 언급은 하지 않았다. ChatGPT를 공개한지 100일쯤 되었을 시점, OpenAI는 GPT-3, GPT-3.5보다 훨씬 강력한 언어 모델인 GPT-4를 공개했다[30]. GPT-4는 InstructGPT나 ChatGPT와 마찬가지로 사람에 의한 지도 미세 조정이 있었으며, OpenAI가 공개한 자료에 따르면 6개월 간 수행되었다고 한다. 즉, GPT-4의 지도 학습은 InstructGPT와 ChatGPT의 공개 시점 사이에 이루어지고 있었다.

GPT-4는 ChatGPT와 마찬가지로 모델의 아키텍처나 규모, 학습 데이터에 대해 구체적인 정보가 공개되지 않았다. 이는 GPT-4 역시 상업적 목적으로 활용되기 때문이기도 하고 지속적으로 서비스가 업데이트 되기도 하지만, 가짜 뉴스나 사기 등 거대 언어 모델이 악용될 수 있는 사회적, 윤리적 문제를 같이 고려한 것으로 생각된다.

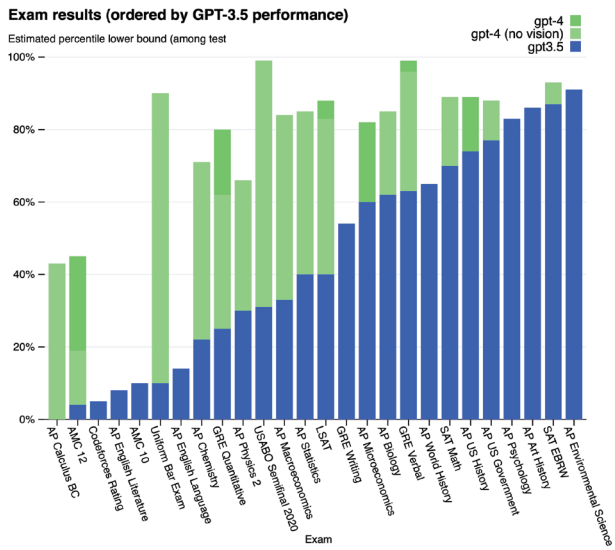
그럼에도 불구하고 GPT-4가 어떻게 기존 모델들이 갖고 있는 문제를 개선했는지는 살펴볼 수 있다. GPT-4는 기존의 모델들이 갖고 있었던 환각 문제와 편향되거

나 유해한 응답 문제를 해결하고자 RBRM (rule-based reward model)을 도입하였다. 이 새로운 RM은 GPT-4의 답변이 적절하거나 그렇지 않은지를 평가하고, 부적절한 프롬프트를 거절하는 것에 보상을 주었다. 이를 통해 언어 모델의 응답 문제를 크게 개선할 수 있었다.

GPT-4의 가장 큰 특징은 멀티모달이다. 단순히 텍스트 형태의 프롬프트뿐만 아니라 그림이나 ABC 기보법으로 작성된 음악 데이터를 입력으로 받아 그 의미나 문맥을 기술하는 응답을 생성할 수 있었다. 이 때, 응답은 이미지를 생성할 수 있는 TikZ, SVG (Scalable Vector Graphics) 코드로도 생성할 수 있었다[31].

GPT-4를 공개한 시점과 같은 해, 2023년 9월, OpenAI는 나아가 시각 (vision) 정보도 처리할 수 있는 GPT-4V를 공개했다[32]. GPT-4V는 한 장 이상의 이미지를 프롬프트로 입력 받을 수 있고, 해당 이미지는 텍스트 데이터와 마찬가지로 그 문맥과 작업이 추론되어 그에 적절한 응답이 생성된다.

GPT-4V는 단순히 이미지를 인식하는 것뿐만이 아니라 이미지 내 글자들이나 이미지 위에 오버레이 되어 있는 도형이나 문구를 같이 인식할 수 있고, 손글씨 수식이나 순서도, 그래프, 표, 유명인 사진, 유명한 장소, 유명한 음식, 의료 영상, 동영상의 프레임들 해석 등 복잡한 작업도 수행할 수 있음을 보여주었다[33].



〈그림 3〉 GPT-4의 학술 및 전문 시험 성능 평가.
대학 수학 능력 시험뿐만 아니라 법학, 의학, 프로그래밍 시험에서도 유의미한 성과를 보여주었다.

〈그림 3〉은 GPT-4의 응답을 채점한 학술 및 전문 시험의 점수를 보여준다[30]. 이를 통해 GPT-4V의 성능을 평가할 수 있었으며, 여러 대학 수학 능력 시험에서 높은 성적을 보여주었다. 또한 법학, 의학, 프로그래밍 시험에서도 유의미한 성적을 보여주었다.

IV. 거대 언어 모델의 의의와 전망

거대 언어 모델의 발전은 인공지능 기술의 가장 중요한 전환점 중 하나로 여겨지고 있다. 이 모델들은 지금도 굉장히 빠르게 진화하고 있으며, 인간과 유사한 수준으로 언어를 이해하고 생성하는 능력을 보여주었다. 특히 특정 작업 혹은 시험에서는 특출난 성능을 보여주기도 했다. 이러한 거대 언어 모델은 의료, 법률, 교육 등 전문적인 분야에도 응용될 수 있는 가능성을 보여주고 있다.

또한, 멀티모달 모델의 발전으로 단순히 텍스트 데이터에 대응하는 것뿐만이 아니라 이미지를 입력으로 받거나 출력으로 생성하는 것까지 가능했다. 〈그림 1〉에서와 같이 인공지능이 주어진 프롬프트 '카렐 차페크의 희곡 RUR의 상상도를 그려줘'에 대해 실제 희곡의 내용을 반영하여 그림을 그릴 수 있는 것도 GPT-1을 개발할 때부터 학습에 사용된 다양한 문학 서적 내용과 InstructGPT를 개발할 때부터 사용된 지도 미세 조정 기술이 주어진 이미지 출력 데이터에 대해 높은 선호도를 평가하여 반영할 수 있었기 때문이다.

이러한 성과는 '이해'와 '앎'의 개념을 다시 정의하게 만든다. 인간의 경우에도 무언가를 이해했거나 알고 있음을 시험으로 평가하고 있다. 같은 기준을 적용한다면 인공지능이 언어를 기반으로 한 추상적인 대상이나 이미지를 기반으로 한 물리적인 대상에 대해 인간과 유사하거나 혹은 특정 대상에 대해서는 더 뛰어난 추론을 할 수 있다고 말할 수 있다.

미국의 철학과 교수 존 설은 기계가 튜링 테스트를 통과하더라도 지능이 있다고 결론 내릴 수 없다고 주장했다[34]. 그는 중국어 패턴에 대한 매뉴얼이 있다면, 이론적으

로 어떤 사람이 중국어를 전혀 할 줄 모르더라도 다른 사람들에겐 중국어를 이해하고 있는 것처럼 보일 수 있다고 주장했다.

거대 언어 모델 역시 어떠한 개념을 이해한 '지능'인지, 혹은 주어진 패턴을 학습하고 모델을 최적화한 '모방'인지 구분하기 어렵다. 이를 구분하기 어렵게 만드는 이유 중 하나는 같은 질문을 인간에게도 할 수 있고, 지능을 가진 인간과 모방을 하는 인간도 구분하기 어렵기 때문이다.

인공지능 모델 평가에 대한 심리적, 철학적 논쟁은 굉장히 중요하다. 인간과 인간의 지능에 대한 깊은 이해를 바탕으로 InstructGPT 혹은 ChatGPT를 지도 학습한 것과 같이 보상 모델을 만들어 인공지능을 미세 조정 할 수 있기 때문이다. 이는 인공지능을 더욱 인간 다운 반응 혹은 응답을 하는 것을 가능하게 한다.

그러나 거대 언어 모델의 발전에 따른 기술적 성과 외에 윤리적, 사회적 문제가 동시에 대두되고 있다. 모델의 발전에 따라 편향되거나 유해한 응답을 하지 않도록 모델이 개발되고 있지만 이 문제는 근본적인 수준에서는 아직 해결되지 않았으며, 이로 인한 가짜 정보의 생성과 확산 등의 문제가 있다. 이러한 문제를 해결하려면 모델 개발 과정에서의 투명성을 높이고, 다양한 배경을 가진 사람들의 참여를 확대하여 미세 조정 단계에서 편향되지 않은 데이터를 확보하고 지도 학습을 할 필요가 있다.

추가로, 거대 언어 모델과 이를 기반으로 한 서비스의 발전에 따른 사회적 영향을 예측하고, 적절한 규제와 정책을 마련하는 것도 필요하다. 예를 들어, 인공지능 기술이 발전하고 사회 전반에 영향을 끼침에 따라 가치의 생산과 가치 판단 과정을 인공지능에게 넘김으로써 수동적인 삶의 태도를 갖게 되는 것을 경계할 수 있다.

비교적 최근인 지난 3월, Anthropic의 거대 언어 모델인 Claude 3가 지능 지수 시험 (IQ Test)에서 101을 기록했음이 알려졌다[35]. 지능 지수 (IQ)는 평균이 100이 되도록 평가하는 시험으로, 통계적으로 인간의 절반은 IQ가 100 이하라는 의미이다. 거대 언어 모델의 발전에 따라 여러 모델들이 다양한 형태의 지능 지수 시험에서 다수의 인간보다 높은 성취를 보여줄 가능성이 있다. 그러한 시기에 인간의 사회적 기여에 대한 문제도 미리 고민할 필요가 있다.

V. 결론

인공지능의 개념은 이미 100년 전부터 구체화된 개념으로 묘사되고 있었으며 컴퓨터 과학과 연산 처리 장치의 발전으로 괄목할만한 성과를 이뤄냈다. 그 중에서도 Transformer 아키텍처를 기반으로 한 거대 언어 모델은 인간의 지식과 창의력을 확장하는 도구로써 엄청난 잠재력을 보여주었다. 방대하면서도 잘 정돈된 언어 데이터의 수집과 효과적인 학습 방법에 대한 연구, 그리고 그것을 실현시킬 수 있는 연산 처리 장치의 조합으로 다양한 거대 언어 모델이 등장하고 오늘날에도 서비스되며 다방면에서 활용되고 있다.

이 모델들이 제공할 수 있는 맞춤형 교육, 신속한 정보 접근, 빠른 문제 해결 등은 인간 사회에 긍정적인 변화를 가져올 수 있다. 그러나 이러한 긍정적인 가능성을 실현하기 위해서는 기술적 진보뿐만 아니라 철학적, 윤리적, 사회적 고려가 병행되어야 한다. 거대 언어 모델 기술의 연구, 개발이 가속화되고 있는 지금, 거대 언어 모델의 과거를 통해 거대 언어 모델의 현재를 이해하고, 거대 언어 모델의 미래를 대비할 필요가 있다.

참고문헌

[1] Karel Capek, R. U. R.: Rossum's Universal Robots, Aventinum, 1920, Prague.

[2] James Betker, et al., "Improving image generation with better captions", Internal Report, 2023.

[3] Alan M. Turing, "I.–Computing machinery and intelligence", Mind, 1950.

[4] Frank Rosenblatt, "The perceptron: a perceiving and recognizing automaton (Project PARA), Cornell Aeronautical Laboratory, 1957.

[5] Frank Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain", Psychological Review, 1958.

[6] David Silver, et al., "Mastering the game of Go with deep

neural networks and tree search", Nature, 2016.

[7] OpenAI, https://openai.com/blog/chatgpt

[8] Jacob Devlin, et al., "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv:1810.04805, 2018.

[9] Chowdhery, et al., "PaLM: Scaling language modeling with pathways", arXiv:2204.02311, 2022.

[10] Google, "PaLM 2 technical report", arXiv:2305.10403, 2023.

[11] Google Gemini Team, "Gemini: A family of highly capable multimodal models", arXiv:2312.11805, 2023.

[12] Google Gemini Team, "Gemini 1.5: Unlocking multimodal understanding across millions of tokens of context", arXiv:2403.05530, 2024.

[13] Hugo Touvron, et al., "LLaMA: Open and efficient foundation language models", arXiv:2302.13971, 2023.

[14] Hugo Touvron, et al., "Llama 2: Open foundation and fine-tuned chat models", arXiv:2307.09288, 2023.

[15] Weizhe Yuan, et al., "Self-rewarding language models", arXiv:2401.10020, 2024.

[16] Anthropic, "Model card and evaluations for Claude models", Internal Report, 2023.

[17] Anthropic, "The Claude 3 model family: Opus, Sonnet, Haiku", Internal Report, 2024.

[18] Boseop Kim, et al., "What changes can large-scale language models bring? intensive study on HyperCLOVA: billions-scale Korean Generative Pretrained Transformers", arXiv:2109.04650, 2021.

[19] Ildoo Kim, Gunsoo Han, Jiyeon ham, and Woonhyuk Baek, "KoGPT: KakaoBrain Korean(hanful) Generative Pre-trained Transformer", https://github.com/kakaobrain/kogpt, 2021.

[20] Tom B. Brown, et al., "Language models are few-shot learners", arXiv: 2005.14165, 2020.

[21] Joseph Weizenbaum, "ELIZA-a computer program for the study of natural language communication between man and machine", Communications of the ACM, 1966.

[22] Claude E. Shannon, "A mathematical theory of communication", Bell System Technical Journal, 1948.

[23] Sepp Hochreiter, Jurgen Schmidhuber, "Long Short-Term Memory", Neural Computation, 1997.

[24] Dzmitry Bahadanau, Kyunghyun Cho, and Yoshua Bengio, "Neural machine translation by jointly learning to align and translate", arXiv:1409.0473, 2014.

[25] Ashish Vaswani, et al., "Attention is all you need", arXiv:1706.03762, 2017.

[26] Alec Radford, et al., "Improving Language Understanding by Generative Pre-Training", Internal Report, 2018.

[27] Alec Radford, et al., "Language models are unsupervised multitask learners", Internal Report, 2019.

[28] Long Ouyang, et al., "Training language models to follow instructions with human feedback", arXiv:2203.02155, 2022.

[29] OpenAI, https://openai.com/blog/chatgpt-plus

[30] OpenAI, "GPT-4 technical report", arXiv:2303.08774, 2023.

[31] Sebastien Bubeck, et al., "Sparks of artificial general intelligence: early experiments with GPT-4", arXiv:2303.12712, 2023.

[32] OpenAI, "GPT-4V(ision) system card", Internal Report, 2023.

[33] Zhengyuan Yang, et al., "The dawn of LMMs: preliminary explorations with GPT-4V(ision)", arXiv:2309.17421, 2023.

[34] John R. Searle, "Minds, brains, and programs", Behavioral and Brain Sciences, 1980.

[35] Maxim Lott, "AIs ranked by IQ; AI passes 100 IQ for first time, with release of Claude-3", Maximum Truth, https://www.maximumtruth.org/p/ais-ranked-by-iq-ai-passes-100-iq, (2024, April 3rd).

김동옥

- 2014년 2월 KAIST 물리학과, 수리과학과 학사
- 2022년 2월 KAIST 물리학과 박사
- 2022년 5월 ~ 현재 삼성전자 DS부문 설비기술연구소 책임연구원

〈관심 분야〉
Artificial Intelligence, Machine Learning, Deep Learning, Computer Vision, Image Processing

실물 환경에서의 데이터 기반 강화학습 적용 방안

Ⅰ. 서 론

8년 전, 구글 딥마인드의 알파고(AlphaGo)와 이세돌 9단 간에 있었던 세기의 바둑 대결은 인공지능 학계뿐만 아니라 일반 사회에도 인공지능 기술의 발전과 그 가능성에 대한 큰 의미를 제공했다. 이 대결은 인공지능이 인간 전문가를 능가할 수 있는 수준에 도달했음을 보여주는 중요한 사례이자, 특히 복잡한 전략이 요구되는 바둑과 같은 분야에서도 인공지능의 가능성을 입증하는 계기가 되었다.

알파고를 학습시키는 과정은 크게 두 가지로 나눠서 볼 수 있다. 먼저 인간 전문가의 기보 데이터를 학습하여, 특정 상황에서 어떤 수를 둘지 예측하는 지도학습(Supervised Learning)이 동작하며, 이때 예측된 결과를 바탕으로, 다양한 바둑판 상황을 경험하고 각 상황에서의 최적의 수를 찾도록 정책을 지속적으로 개선시키는 강화학습(Reinforcement Learning)을 통해 학습이 된다. 특히 앞의 대국 사례를 통해 강화학습이 복잡한 문제를 해결하고, 인간 전문가의 수준을 능가할 수 있는 잠재력에 주목하고, 이를 실제 생활 응용 분야에 접목시키고자 하는 노력이 현재까지 이어져 오고 있다. 특히 자율 주행 차량, 로봇 공학, 에너지 관리, 추천 시스템 등에서는 강화학습이 적용되면서 성능을 개선시키는 사례들이 학계에 발표되고 있다.

하지만 강화학습을 실제 생활에 적용하는데 있어서 여전히 중요한 도전 과제들이 존재한다. 불과 몇 년전만 하더라도, 좋은 성능을 보여주는 대부분의 강화학습 알고리즘은 '온라인' 학습을 전제로 하며, 이는 에이전트가 환경과 직접 상호작용을 하면서 데이터를 수집하고 학습하는 것을 의미한다. 이 학습 과정 안에는 실시간으로 대량의 데이터를 처리하고, 때로는 탐색을 통한 시행착오 과정이 포함되어 있다. 예를 들어, 자율 주행 차량이 실제 도로에서 강화학습을 통해 학습한다고 가

강찬석
LG전자

정하면, 분명 사고를 통해서 뭔가를 배우긴 하겠지만, 이때는 막대한 비용과 안전상의 위험을 동반하게 된다.

이러한 배경에서 데이터 기반 강화학습, 혹은 오프라인 강화학습 (Offline RL)이 주목받기 시작했다. 오프라인 강화학습은 사전에 수집된 데이터셋을 활용하여 학습하는 방식으로, 에이전트가 환경과의 직접적인 상호작용 없이도 의사 결정 정책을 개선할 수 있도록 해준다. 이 접근 방법을 통해서 데이터 수집 과정에서 발생할 수 있는 비용과 위험을 크게 줄여주며, 사전에 수집한 대규모 데이터를 효율적으로 활용할 수 있는 가능성을 열어줬다. 특히 의료, 금융, 제조업 등의 분야에서 오프라인 강화학습은 실제 문제 해결에 적용될 잠재력을 가지고 있다.

본 기고문에서는 강화학습의 기본 원리와 온라인 강화학습의 한계를 넘어서 오프라인 강화학습이 어떻게 실제 문제 해결에 기여할 수 있는지에 대해 소개하고자 한다. 또한 오프라인 강화학습의 현재 연구 동향과 주요 도전 과제, 그리고 자사 적용 사례를 통해서 실물 환경에서 오프라인 강화학습을 적용해 문제 해결에 폭넓게 적용될 수 있는 길을 모색해보고자 한다.

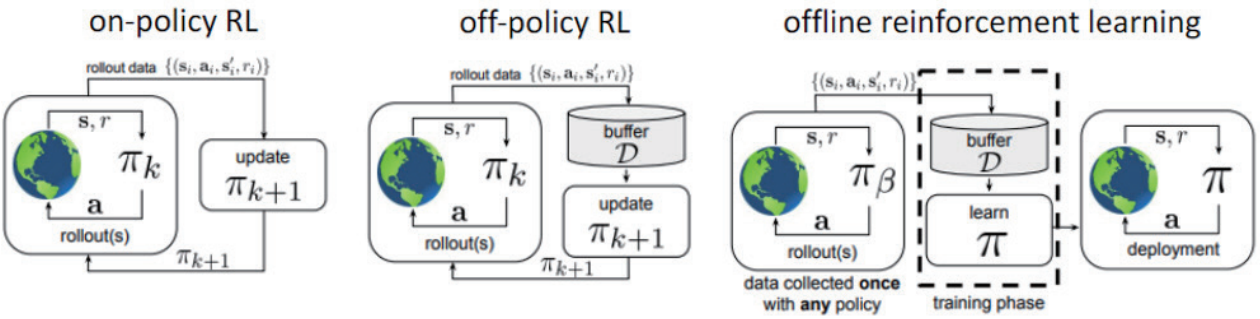
Ⅱ. 오프라인 강화학습

레이블의 존재 여부에 따라 학습이 이뤄지는 지도학습이나 비지도학습과는 다르게, 강화학습은 에이전트가 주어진 환경에서 상호작용을 하면서, 시행착오를 통해 최적의 행동 전략, 즉 정책을 학습하는 학습 방법론 중 하나이다. 강화학습의 핵심 목표는 학습이 이뤄지는 환경상에서 에이전트가 최대의 보상을 얻을 수 있는 정책(policy)을 찾는 것이며, 이 과정에서 에이전트(agent)는 환경(environment)의 상태(state)를 관찰하고, 행동(action)을 취한 후, 그 결과로부터 보상(reward)를 받게 된다. 이런 상호작용을 통해 에이전트는 시간이 지남에 따라 최적의 정책을 학습하게 된다.

강화학습을 학습 방법론으로 사용했을 때의 이점으로는 우선 명시적인 레이블 데이터가 필요하지 않아 데이터 수집이 용이하다는 점이다. 또한 환경에서 얻은 상태 데이터를 바탕으로 학습이 이뤄지기 때문에 실제 환경에 적용할 수 있는 정책 자체를 학습할 수 있다. 이런 특성을 활용하여 최적화 문제나 로봇 제어, 게임 같은 다양한 분야에 효과적으로 적용할 수 있다.

하지만 다양한 상태에 대한 이상적인 정책을 학습하기 위해서는 지도학습이나 비지도 학습에 필요한 데이터보다 훨씬 더 많은 데이터를 요구한다는 단점이 존재한다. 그리고 실제 환경과 같이, 게임이나 어떤 상황을 가상으로 모사할 수 있는 시뮬레이터가 존재하지 않는 환경이라면 해당 환경을 모델링하는 것이 매우 복잡할 수 있으며, 이로 인해서 학습 과정이 어려워지는 요인이 될 수 있다. 마지막으로 앞에서 언급했다시피 강화학습의 목표는 에이전트가 최대의 보상을 받을 수 있는 정책을 찾는 것인데, 어떤 상태가 좋고 나쁜지 판단할 수 있는 적절한 보상 함수가 설계하는 것은 현재까지도 강화학습을 실제 환경에 적용하는데 큰 어려움으로 남아있다.

강화학습 알고리즘은 크게 on-policy와 off-policy로 구분할 수 있다. On-policy 방법은 에이전트가 현재 수행하고 있는 정책(behavior policy)에 따라 행동을 선택하고, 이를 통해서 학습을 진행하는 방식이다. 대표적



〈그림 1〉 온라인 강화학습(on-policy와 off-policy)과 오프라인 강화학습의 일반적인 형태 [1]

인 알고리즘으로는 SARSA (State-Action-Reward-State-Action) 가 있는데, 보통 이런 방식에서는 에이전트가 경험하는 모든 상황이 현재 평가하는 정책과 직접적으로 관련되어 있어야 한다.

반면, off-policy 방법은 에이전트가 행동을 취하는 정책(behavior policy)를 활용하여, 실제 학습하고자 하는 정책(target policy)를 평가하고 최적화하는 것을 말하며, Q-learning이 대표적인 알고리즘으로 알려져 있다. Off-policy 방법에서는 현재의 정책이 아닌 다양한 정책으로 학습하는 기법이기 때문에, on-policy에 비해서 최적의 정책을 학습할 수 있는 유연성이 생긴다.

Q-learning에 신경망이 붙으면서 발전된 형태인 Deep Q-Network (DQN)[2] 이 나오면서 off-policy의 특성을 살린 experience replay의 역할이 중요해졌다. 단어 뜻에 내포되어 있는 것처럼, experience replay는 학습 과정에서 에이전트가 겪은 경험(State, Action, Reward, Next state 등)을 메모리에 저장하고, 이를 재사용하면서 학습을 진행하는 방법이다. 이를 사용하면, 이전의 경험을 여러 번 재사용함으로써, 더 적은 수의 에피소드로부터 더 많이 학습할 수 있다. 특히 데이터를 얻기 어렵거나 비용이 많이 드는 환경에서 유용하다. 또한 과거의 경험을 무작위로 샘플링하면서, 데이터간의 상관 관계를 줄이고, 학습 과정을 안정시킬 수 있는 장점을 가진다.

지금까지 설명한 강화학습 알고리즘들의 전제는 모두 환경과의 상호작용이 온라인으로 이뤄진다는 것이다. 그런데 만약 앞에서 언급한 experience replay을 온라인이 아닌 오프라인 설정에서 사용할 수 있다는 가정을 해보면 어떨까? 다시 말해, 실제 환경과 상호작용이 없는 상태에서도 이전에 쌓은 경험을 바탕으로 정책을 학습시킬 수 있지 않을까 하는 아이디어가 생기고, 여기에서 출발한 개념이 바로 오프라인 강화학습이다.

오프라인 강화학습은 배치 강화학습(Batch RL)[3]라고 불리기도 하며, 수집된 대규모의 데이터셋을 활용하여 에이전트를 학습하는 방법이다. 어떻게 보면, 앞에서 언급한 off-policy를 확장시킨 형태이므로 fully off-policy라고 표현하기도 한다. 오프라인 강화학습은 특히 데이터 수집에 비용이 많이 들거나 위험한 환경, 예를 들어 의료나 자율 주행 등의 분야에서 유용하다.

무엇보다 오프라인 강화학습이 최근들어 주목받게 된 이유는 이렇게 데이터 기반으로 학습하는 방법론을 통해 강화학습의 적용 범위를 크게 확장할 수 있는 잠재력을 부여한다는 점이다. 이를 통해 기존의 온라인 강화학습을 실제 환경에 적용시킬 때 겪는 한계에 대한 새로운 접근 방법을 제공하며, 강화학습 자체의 실용적인 적용 가능성을 더욱 넓힐 수 있게 되었다.

하지만 오프라인 강화학습이 가지는 한계도 역시 존재한다. 강화학습에서 정책을 개선시킬 수 있는 근간 중 하나는 에이전트가 실제 환경에 대한 "탐색(exploration)"을 수행하는 것인데, 온라인 강화학습에서는 탐색에 대한 피드백을 "온라인"으로 받을 수 있기 때문에, 환경에 대한 다양한 경험을 토대로 정책을 실시간으로 개선시킬 수 있다. 하지만 오프라인 강화학습의 경우, 이런 탐색 과정을 데이터셋 내에서 수행해야 하며, 데이터셋이 환경을 표현할 수 있을 만큼 충분하지 않거나, 다양한 경험이 포함되어 있지 않다면 성능을 개선시킬 수 있는 여지가 온라인 강화학습에 비해 부족하다. 또한 본질적으로 학습 데이터의 분포와 실제 경험하는 데이터의 분포가 다르면서 발생하는 Data Distribution Shift 문제가 있다.

실제 환경에서는 이 Data Distribution Shift의 문제가 더 부각되어 나타난다. 〈그림 2〉는 실제로 오프라인 강화학습을 적용하기 위해서 수집된 데이터(파란색)와 해당 데이터로 학습된 모델을 실제로 돌렸을 때 수집된 온라인 데이터(빨간색)를 t-SNE (t-distributed stochastic neighbor embedding)[3]라는 기법을 통해서 차원을 감소시켜 시각화를 해본 그림이다. 그림에서 확인할 수 있다시피, 왼쪽의 경우는 오프라인 학습시 사용한 데이터의 분포가 온라인에서 수집된 데이터의 분포를 대부분 충족시킨다. 다시 말해, 오프라인에서 데이터로 학습된다 할지라도, 실제 수행했을 때 비슷한 상태를 관찰하기 때문에 적어도 성능이 저하될 가능성이 적다. 반대로 오른쪽의 경우는 온라인에서 수집된 데이터 분포가 오프라인에서 사용된 데이터의 분포 이외의 영역을 차지하고 있는 것을 확인할 수 있다. 다시 말해, 오프라인 강화학습을 통해서 학습이 잘 되었다 할지라도 온라인에서 실제 경험하

지 못한 상태(counterfactual query)를 관찰할 경우, 이상적인 행동을 할 수 없게 된다. 특히 시간이 지남에 따라 환경이 변화할 때, Distribution shift로 인한 모델의 성능 저하는 더욱 두드러지게 나타나게 된다. 이 때문에 에이전트가 학습한 환경과 실제 환경 사이의 불일치를 극복할 수 있는 전략이 필요하다.

III. 실물 환경 강화학습의 어려운 점

강화학습을 실제 환경에 적용한 사례는 점점 늘고 있지만, 영상이나 NLP와 같은 분야에 비하면 그 적용 사례가 현저히 작은 편이다. 이렇게 실물 환경 강화학습 (Real-World Reifnorcement Learning)이 실사례화되기 어려운 점에 대해서 소개한 논문[4]이 있어 이 중 몇 가지를 정리해보고자 한다.

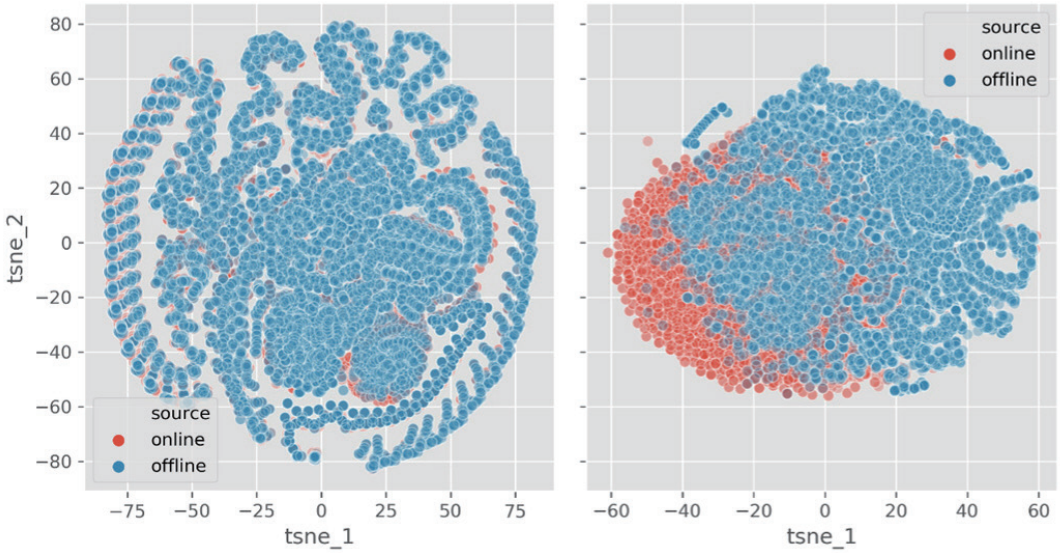
1. 제한된 샘플에서의 학습

강화학습 실제 적용 사례로 언급되는 게임이나 로봇의 경우는 ALE[5]나 MuJoCo[6]와 같은 환경을 모사할 수 있는 시뮬레이터가 존재하기 때문에, 가상 환경상에서 임의의 정책을 사용하여 탐색하고, 해당 데이터를 바탕으로 정책의 성능을 개선시킬 수 있었다. 하지만 대부분의 실물 환경은 일반적으로 데이터를 생성하는데 있어 비용적인 문제와 위험성 문제를 가지고 있으며, 실제 환경에서 정책의 평가와 개선이 이뤄지게 된다. 이 때문에 시뮬레이터를 사용하는 방식에 대비해 탐색이 제한적이며, 얻어지는 데이터도 low-variance, 즉 전체 데이터에 대비해 다양한 특성을 반영하지 못하게 된다. 이 때문에 오프라인 혹은 온라인 상에서 Off-policy 방법과 같이 다른 모델이 수집한 정책을 기반으로 학습하는 형태가 실물 강화학습에서 적용되는 방법론 중 하나이다.

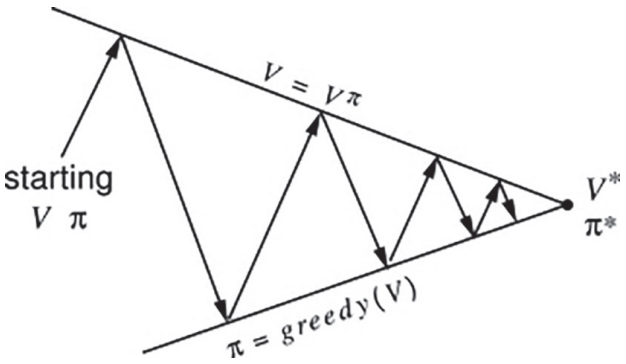
또한 실제 환경 상에서의 학습 주기는 시뮬레이터를 사용하는데 비해 현저하게 길다. 특히 온라인 광고에 적용되는 추천 시스템이나 의료/의학 계열과 같은 분야의 경우는 학습이 적용되면서 실제 효과로 나타나는 데까지 수개월까지 걸리는 경우도 존재한다. 그렇기 때문에, 실제 환경에서 학습하기 위해서는 적어도 샘플 효율성도 좋으면서, 시스템에 적용되었을 때의 성능도 역시 좋은 알고리즘이 적용되어야 한다.

2. Offline/Off-policy 환경에서의 성능 평가

강화학습이 다양한 환경에서의 일반성과 높은 성능을 보이는 근본적인 원리 중 하나는 정책의 평가(policy evaluation)와 정책의 개선(policy improvement)이 반복되면서 최종적으로 이상적인 정책을 찾아가는 방법(Policy Iteration)이다.



〈그림 2〉 오프라인 데이터셋과 온라인에서 수집된 데이터간의 분포 차이



〈그림 3〉 Policy Iteration [7]

이중, 정책의 평가 단계는 개선될 정책과 비교할 수 있는 현재 정책의 성능 수준을 나타내기 때문에 필요한 단계이다.

문제는 실물 환경에서 이 평가 단계가 실제 환경에서 이뤄져야 한다는 것이다. 앞에서 언급했던 것처럼 실제 환경에서 탐색을 하던가 어떤 정책을 수행한다는 것은 비용적인 문제나 위험성의 문제가 동반한다. 그렇기 때문에 실제 환경에서 정책을 수행하지 않고도 정책을 평가할 수 있는 Off-Policy Evaluation (OPE)[8] 방법이 필요하다. 쉽게 표현하면, 하나의 정책(behavior policy)으로부터 수집된 데이터를 사용하여 다른 정책(target policy)의 성능을 평가하는 방법인데, 실제 환경에서의 시행착오 없이 정책의 성능을 평가할 수 있게 해주므로, 비용이 많이 들거나 위험한 환경에서의 학습에서 유용하게 사용할 수 있다. Importance Sampling이나 Direct Method, Doubly Robust와 같은 기법들이 OPE의 대표적인 알고리즘이며, 각각의 기법별로 장점과 한계가 존재한다. OPE가 가지는 일반적인 도전과제에 대해 언급하자면, 우선 OPE의 정확도는 OPE에 사용된 데이터의 품질과 다양성이 좋다는 가정을 전제로 두고 있기 때문에, 만약 수집된 데이터가 target policy의 다양한 상황을 충분히 커버하지 못한다면, 성능 추정이 부정확할 수 있다는 점이다. 또한 소개한 Importance Sampling 기법은 high-variance를 가지는 반면, Direct Method는 high-bias를 가지면서 OPE의 오류가 발생하게 되는데, 이 variance와 bias간의 균형을 맞춘 OPE 방법에 대한 연구도 필요하다. 그리고 마지막으로, 실제 환경과 같은 복잡성을 띤 환경에서는 고려하고 있는 변수뿐만 아니라 다양한 요소들이 상호 작용하기 때문에 OPE를 통한 정책 평가가 조금 더 어려울 수도 있다.

이를 해결하기 위해서 앞에서 소개한 기법 이외에 오프라인 강화학습에서 간단하게 사용할 수 있는 Fitted Q-Evaluation (FQE)[9]나 Regularizatrion 기법이 적용된 DIstribution Correction Estimation (DICE)[10] 같은 기법들도 존재하지만, 근본적으로 오프라인 데이터만을 가지고 OPE를 수행하는데 있어 한계가 존재하기 때문에, OPE를 명시적으로 수행하지 않고도 성능을 개선할 수 있는 방법[11], [12] 들도 최근에 소개되었다.

3. 불특정적이고, 다중 목표를 띄는 보상 함수

강화학습의 궁극적인 목표는 최종 기대 보상의 합이 최대가 되는 정책을 찾는 것이지만, 실제 환경에서는 종종 보상 함수가 명확하게 지정되지 않거나(unspecified), 여러 목표를 동시에 만족시켜야 하는(multi-objective) 경우가 많다. 강화학습의 사례로 많이 소개되는 게임의 경우는 점수라고 하는 명확한 보상함수가 존재하지만, 자율주행 자동차의 경우는 단순히 목적지에 도달하는 것 뿐만 아니라, 안전, 연료 효율성 등 다양한 요소를 고려해야 한다. 이처럼 실제 환경을 강화학습 환경으로 변환하는 문제는 복잡하며, 고려해야 할 모든 중요 요소를 단일 보상 함수로 표현하기 어렵다. 더불어 인간과 관련되어 있는 테스크일 경우 인간의 주관적 선호도가 보상에 영향을 주게 되는데, 이런 선호도를 정량화하여 표현하는 것 역시 강화학습을 실제 환경에 적용하기 어려운 장애물 중 하나이다. 또한 여러 목표를 동시에 만족해야 하는 multi-objective의 특성을 가지기 때문에 각 목표에 대한 보상 사이에서 적절한 균형을 찾는 것도 어려운 일이다.

이를 해결하기 위한 방안으로 다양한 기법들이 연구되고 있지만, 최근 주목받고 있는 내용 중 하나는 인간의 피드백을 직접적으로 학습 과정에 통합하여 보상 함수를 동적으로 조정하여 생성하는 Reinforcement Learning from Human Feedback (RLHF)[13] 이다.

RLHF에서는 인간의 선호도를 직접적으로 학습 데이터로 사용하는데, 예를 들어 인간 평가자가 제공한 행동 중

선호하는 행동에 대한 정보를 사용하여, 에이전트가 선호도를 반영한 보상 함수를 학습할 수 있다. 이 방법을 통해서 목표가 복잡하거나 모호한 상황에서도 인간의 직관을 모델에 통합할 수 있는 방법을 제공한다. 또한 multi-objective 문제에서 시간에 따라 변화하는 목표의 중요도나 환경의 변화를 반영하여, 보상 신호를 동적으로 조정할 수 있다는 특징도 가진다. 물론 인간의 피드백을 기반으로 하기 때문에, 피드백의 질이나 양, 혹은 제공하는 과정에서의 일관성도 있어야 하며, 인간 평가자의 편향이나 주관적인 판단이 결과에 영향을 미칠 수 있다는 한계도 존재하지만, 적절히 설계된 RLHF 시스템은 실물환경의 문제 해결에 있어 강화학습의 적용 범위를 확장시킬 수 있다.

IV. 최근 동향 및 자사 적용 사례

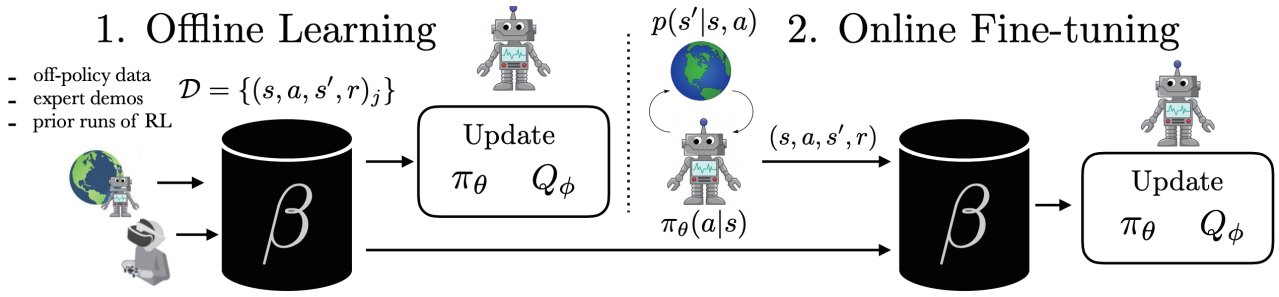
앞에서 소개한 바와 같이 오프라인 강화학습은 오프라인 데이터만 활용하여 학습하는 형태로써, 본질적으로 Data Distribution shift 문제를 가지고 있으며, 이를 해결하기 위해 대안으로 제시된 Conservative Q-Learning (CQL)[14] 와 같은 Conservative 접근 방식 역시 sub-optimality를 가지면서 궁극적인 해결 방안을 제공하지 못했다.

이를 해결하기 위해 최근 활발하게 연구되고 있는 주제 중 하나는 오프라인 강화학습에서 발생하는 sub-optimality를 해결하기 위해서 온라인 fine-tuning 기법을 활용하여, 조금 더 학습에 도움이 될만한 고급 데이터를 사용하는 Offline-to-Online 강화학습이다. 기존의 온라인 강화학습과의 차이라면, 온라인 학습은 초기에 환경에 대한 정보 없이 탐색을 통한 학습이 이뤄지는 반면, Offline-to-Online 방법은 오프라인 강화학습을 통해 사전 학습된 에이전트를 온라인 상에서 세부 학습을 통해 성능을 향상한다는 점이다. 결과적으로 오프라인 강화학습이 가지는 샘플 효율성과 온라인 강화학습이 가지는 성능 개선 측면이 적절히 조합된 방법론이라 할 수 있다. 관련하여 오프라인 학습 모델을 활용하는 방법[16]이나 오프라인 데이터만을 가지고 온라인 학습을 수행하는 기법[17] 등 다양한 기법들이 최근 소개되고 있으므로, 강화학습 관련 연구자들이 참고해서 보면 좋을 것 같다.

자사에서도 오프라인 강화학습을 실제 환경에 적용하는 연구를 진행해 왔으며, 최근 가정에서 많이 사용되는 생활 가전 중 하나인 세탁기를 대상으로 오프라인 강화학습을 적용하여 출시하였다. 앞에서 소개한 실물환경 강화학습의 어려움 역시 세탁기를 대상으로 한 강화학습 환경에서도 동일하게 나타났다.

세탁기에서도 세탁과 탈수에 대한 본질적인 성능 요구와 더불어 옷감 손상 줄임이나 에너지 절감과 같은 multi-objective의 요구 조건이 들어있다. 이런 요구조건 해결을 위한 노력이 지속되어 왔지만, 현재의 개발 경향은 관련 전문가가 실제 세탁기를 놓고 수많은 시행 착오를 통해서 이상적인 제어 변수를 찾는, 흔히 말하는 'user-in-the-loop' 접근 방식을 취하고 있다. 이 또한 세탁기의 내부 구조가 점점 복잡해지고, 세탁기의 내부 환경을 표현할 수 있는 시뮬레이터의 부재로 인해 한계를 가지게 되었다.

이 때문에 처음으로 강화학습을 통해 세탁기의 이상적인 제어 방법을 찾는 시도를 하게 되었으며, 특히 탈수 구간에서 발생하는 진동이 기준 이상을 벗어나지 않으면서



〈그림 4〉 Offline-to-Online 강화학습 [15]

목표 RPM까지 모터를 가속하는 문제를 강화학습을 활용하여 해결하고자 하였다.

추가로 고려했던 부분 중 하나는 실제로 세탁기가 시중에 판매되었을 때, 성능 개선을 어떻게 할 수 있느냐는 점이다. 일반적으로 세탁기에는 저성능의 MCU가 탑재되어 있고, 해당 환경에서는 온라인 학습을 통해서 에이전트의 성능을 개선시키기 힘들다. 이로 인해서 실제 세탁기가 양산되는 시점에는 실험실 환경에서 수집된 데이터 기반으로 오프라인 강화학습을 통해 초기 탑재 모델을 학습한 후, 추후에 실생활에서 발생한 데이터를 수집하여 재학습하여 배포하는 일종의 workflow (Delayed Online Update, DOU)를 제안하였다.[18] 이 방법을 통해 기대할 수 있는 효과는 저성능의 MCU가 탑재된 대부분의 생활가전에서도 ThinQ나 SmartThings와 같이 제어 데이터를 주고 받을 수 있는 구조가 갖춰져 있다면, 지속적인 데이터 수집을 통해 실환경에서의 가전의 본질 성능 개선이 가능하다는 점이다. 물론 앞에서 소개한 Offline-to-Online과 같은 극적인 성능 개선을 기대하기는 어렵지만, 온라인 학습시 필요한 컴퓨팅 자원 등을 고려했을 때 조금 더 실물 환경 강화학습에 실용적인 방법이라고 할 수 있겠다.

개발을 위해서 강화학습 개발자와 세탁기 전문가가 협업하여 강화학습 환경, 즉 Markov Decision Process (MDP)에 필요한 State와 Action, 보상함수, 그리고 기타 환경에 필요한 구성 요소를 정의하였고, 실험실 환경에서 해당 프로세스를 활용하여 개발을 진행하였다. 두 번의 프로세스를 수행한 결과, 내부 개발 지표는 초기 모델 대비 16% 개선됨을 확인할 수 있었고, 특히 세탁기 전문가가 만든 제어 방식과 비교한 성능 지표상에서는 15% 정도의 개선 효과를 확인할 수 있었다.

V. 전망과 결론

강화학습이 게임이나 로봇 제어, 자율 주행, 의료, 금융과 같은 분야에서 인간 대비 우월한 성능을 보여주면서 인공지능 분야에서 중요한 연구 영역으로 자리잡았다. 무엇보다도 지도학습이나 비지도학습과는 다르게 복잡한 의사 결정 문제를 해결하고, 유동적인 환경 변화에 대해서 최적의 행동 전략을 학습하는 능력은 강화학습을 다양한 응용 분야에서 고려해볼만한 방법 중 하나로 만들었다.

실물환경 강화학습과 같은 분야에서의 연구도 최근 연구 사례를 통해 지속적으로 대중에 공개되고 있으며, 현재 시점에도 강화학습을 현실 세계의 다양하고 복잡한 문제를 해결하는데 있어 효과적으로 적용할 수 있는 방법에 대해 연구되고 있다.

또한 강화학습 시스템의 안전성과 윤리적 사용도 중요한 연구 주제 중 하나이다. 인간과 상호작용하는 시스템에서의 안전한 학습 방법, 편향의 최소화, 그리고 윤리적 결정을 내릴 수 있는 알고리즘의 개발이 미래에는 중요한 요소로 작용할 것이다.

또한, 다중 목표 강화학습 (Multi-Objective RL)과 다중 에이전트 강화학습 (Multi-Agent RL)은 팀 기반의 협업, 경쟁, 그리고 복잡한 목표 사이의 균형을 찾는 알고리즘 개발을 통해서 더욱 발전할 것이다. 이와 함께 인간의 피드백과 선호도를 학습 과정에 통합하는 RLHF와 같은 접근 방식은 인간 중심의 의사 결정 시스템을 개발하는데 기여할 것으로 생각한다.

강화학습을 공부하고, 실제 환경에 적용하려 노력하는 연구자로써, 미래의 강화학습 발전 방향은 개인적으로 기대가 되며, 관련 분야의 연구와 개발은 인간의 삶을 향상시키는데 중요한 기여를 할 것이라고 생각한다. 특히 앞에서 소개한 실생활 문제로의 적용, 안전성과 윤리적 고려, 그리고 인간과의 상호작용을 중심으로 한 학습 방법의 발전을 통해서 강화학습이 더욱 강력하고, 실용적이며, 사회적으로 책임있는 기술로 만들어 나갔으면 하는 기대감을 가져본다.

참 고 문 헌

[1] Levine, Sergey, et al. "Offline reinforcement learning: Tutorial, review, and perspectives on open problems." arXiv preprint arXiv:2005.01643 (2020).

[2] Mnih, Volodymyr, et al. "Playing atari with deep reinforcement learning." arXiv preprint arXiv:1312.5602 (2013).

[3] van der Maaten, L.J.P.; Hinton, G.E. “Visualizing High-Dimensional Data Using t-SNE.” Journal of Machine Learning Research 9:2579-2605, 2008.

[4] Dulac-Arnold, Gabriel, Daniel Mankowitz, and Todd Hester. "Challenges of real-world reinforcement learning." arXiv preprint arXiv:1904.12901 (2019).

[5] Bellemare, Marc G., et al. "The arcade learning environment: An evaluation platform for general agents." Journal of Artificial Intelligence Research 47 (2013): 253-279.

[6] Todorov, Emanuel, Tom Erez, and Yuval Tassa. "Mujoco: A physics engine for model-based control." 2012 IEEE/RSJ international conference on intelligent robots and systems. IEEE, 2012.

[7] Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). The MIT Press.

[8] Precup, Doina. "Eligibility traces for off-policy policy evaluation." Computer Science Department Faculty Publication Series (2000): 80.

[9] Le, H. M., Jiang, N., Agarwal, A., Dudík, M., Yue, Y., & Daumé III, H. (2019). Batch Policy Learning under Constraints. In Proceedings of the 36th International Conference on Machine Learning (ICML 2019)

[10] Yang, Mengjiao, et al. "Off-policy evaluation via the regularized lagrangian." Advances in Neural Information Processing Systems 33 (2020): 6551-6561.

[11] Kostrikov, Ilya, Ashvin Nair, and Sergey Levine. "Offline reinforcement learning with implicit q-learning." arXiv preprint arXiv:2110.06169 (2021).

[12] Brandfonbrener, David, et al. "Offline rl without off-policy evaluation." Advances in neural information processing systems 34 (2021): 4933-4946.

[13] Christiano, Paul F., et al. "Deep reinforcement learning from human preferences." Advances in neural information processing systems 30 (2017).

[14] Kumar, Aviral, et al. "Conservative q-learning for offline reinforcement learning." Advances in Neural Information Processing Systems 33 (2020): 1179-1191.

[15] Nair, Ashvin, et al. "Awac: Accelerating online reinforcement learning with offline datasets." arXiv preprint arXiv:2006.09359 (2020).

[16] Zhang, Haichao, We Xu, and Haonan Yu. "Policy expansion for bridging offline-to-online reinforcement learning." arXiv preprint arXiv:2302.00935 (2023).

[17] Ball, Philip J., et al. "Efficient online reinforcement learning with offline data." International Conference on Machine Learning. PMLR, 2023.

[18] Chanseok, Kang., et al. "Data-Driven Reinforcement Learning for Optimal Motor Control in Washing Machines" IEEE Conference on AI (CAI), 2024. (to be appeared)

강찬석

- 2013년 2월 중앙대학교 전자전기공학부 학사
- 2015년 2월 서울대학교 컴퓨터공학부 석사
- 2012년 6월 ~ 2012년 12월 한국과학기술연구원 (KIST) 연구원
- 2019년 4월 ~ 2019년 8월 카네기멜론대학 방문연구원
- 2014년 12월 ~ 현재 LG전자 인공지능연구소 선임연구원

〈관심 분야〉
실물환경 강화학습, 오프라인 강화학습, 임베디드 시스템

반도체 공정 장비에서 인공지능 기술의 역할과 활용

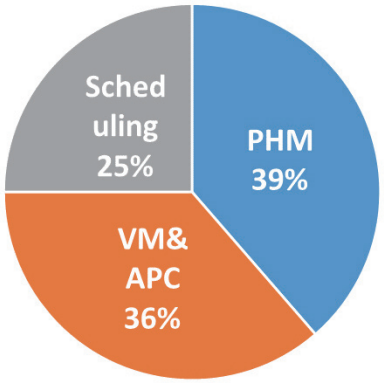
1. 서론

최근 수 년 동안, 인공지능(Artificial Intelligence, AI)은 제4차 산업혁명의 큰 축으로서 다양한 종류의 연구 분야와 산업 분야에서, 이를 각 분야에서 활용하고자 하는 노력을 기울이고 있다[1,2]. 특히, 최근 1~2년 사이에는 OpenAI社에서 개발한 ChatGPT 서비스가 출시되어, 사용자의 질문 분야에 무관하게 전문가 수준으로 응답하는 특징으로 많은 사용자의 주목을 받았다[3,4]. 이에 따라, 세계적으로 여러 정보통신 회사에서는 ChatGPT의 근간이 된 초거대 언어 모델(Large Language Model, LLM)을 기반으로 한 여러 인공지능 서비스들을 앞다투어 시장에 출시하고 있다. 이러한 서비스들은 AIaaS(AI as a Service)로 불리우며, 사용자는 어렵고 복잡한 인공지능 기술을 이해할 수 없어도 인공지능 기술 기반의 서비스를 실 생활에서 사용할 수 있다.

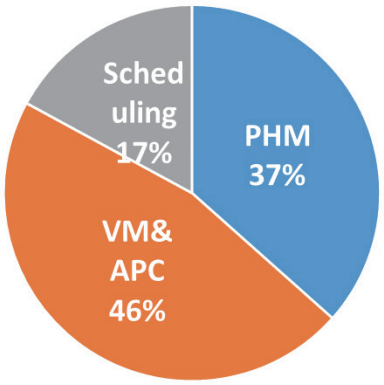
현 시점에서 인공지능은 사람들의 일상을 넘어서, 그 동안 전문가의 영역이라 간주되었던 복합적이고 어려운 산업 분야에서도 활용되고 있다. 특히, 최첨단 분야의 집약체라 할 수 있는 반도체 분야에서도 인공지능 기술을 기반으로, 기존의 비효율적인 작업과 인력 및 시간 낭비를 줄여 생산성과 효율성을 높이려 하고 있다[5,6]. 최근 5년간, 웨이퍼 제조사와 장비 제조사에서는 장비 건강도 예지 및 부품 생명 주기 추정, 웨이퍼(wafer) 품질 관리, 설비 운영 관리(maintenance), 반도체 제조 스케줄링 등에 인공지능 기술을 적용하려는 연구가 진행되고 있다.

〈그림 1과 2〉는 2016년부터 2021년까지, 웨이퍼 제조사〈그림 1〉과 장비 제조사〈그림 2〉에서 인공지능 기술을 기반으로 진행한 연구들을 주제 별로 도식화한 자료이다. 이 때, 조사 방식은 구글 학술 검색을 사용하였으며, 논문 저자들의 이메일 도메인을 기반으로 제조사들을 구분하였다. 〈그림 1〉과 같이, 장비 건강도를 추정하기 위한 인공지능 기

김홍연
㈜원익IPS



〈그림 1〉 Research statistics for AI (Chip maker)



〈그림 2〉 Research statistics for AI (Equipment maker)

반의 PHM(Prognostic and Health Management) 기술을 주로 연구하고 있으며, 다음으로 웨이퍼의 품질, 즉, 두께(thickness), 응력(stress), 잔여물(particle), 불량(defect) 등의 수준을 추정하는 VM(Virtual Metrology) 기술 연구를 진행 중이다. 이 때, VM과 같이 사용되는 기술인 APC(Advanced Process Control)는 반도체 공정 중 웨이퍼의 품질이 하락했을 때, 이를 사전에 정의한 사양에 만족하도록 공정을 자동으로 제어한다. 이외에도, 공정 스케줄링(Scheduling) 관점에서의 연구들도 꾸준히 진행되고 있으며, 웨이퍼 제조사에서는 반도체 장비가 설치된 환경인 Fab에서 장비 간 웨이퍼 이송(transfer)을 수행하는 OHT(Overhead Hoist Transport)의 작업 스케줄링을 연구하고 있다. 이와 다르게, 장비 제조사에서는 장비의 챔버(chamber) 간 웨이퍼 이송을 담당하는 VTM(Vacuum Transfer Module)에 대한 작업 스케줄링 연구를 진행하고 있다.

앞서 기술한 연구 내용을 정리하면, 반도체 분야에서의 인공지능 연구 주제는 크게 4가지로, 1) 장비 건강도 관리(Prognostic and Health Management, PHM), 2) 웨이퍼 품질 추정(Virtual Metrology, VM), 3) 자동 공정 제어(Advanced Process Control, APC), 그리고 4) 제조 스케줄링 최적화(Scheduling Optimization)와 같이 구분될 수 있다. 지금부터는 각 연구 주제가 해결하고자 하는 반도체 공정에서의 문제점과 해당 연구 주제의 필요성을 기술한다.

우선, 장비 건강도 관리 기술인 PHM[5,6]은 반도체 장비 운영에 필수적이다. 반도체 장비는 자연 현상에서는 거의 발생할 수 없는 수준의 초고온, 초고압, 초전압 상태를 인위적으로 생성하여 고품질의 웨이퍼를 다량으로 생산하기 때문에, 누적 가동 시간이 증가할수록 장비 내 상태(status)는 기하급수적으로 악화되고, 또한 부품 별 생명 주기도 단축된다. 이는 웨이퍼의 품질에 직간접적으로 영향을 주기 때문에, 고품질의 웨이퍼를 생산해내면서 장비에서 불량이 발생하지 않는 상태, 즉, 적정 수준의 건강도(health index)를 명확하게 파악하는 것은 매우 중요하다. 가령, 장비의 공정 별로 건강도를 추정할 수 있다면, 건강도가 일정 수준보다 하락했을 때만 정비를 수행하는 조건부 정비(Condition-Based Maintenance)가 가능하며, CBM과 예측 모델을 병합하여 예지 정비(Predictive Maintenance, PdM)로의 확장을 도모할 수 있다. CBM과 PdM 모두 과도한 인적 비용과 시간적 비용을 소요하는 정비 횟수를 감소시켜서, 실제 장비의 생산량을 향상시킬 수 있다.

그러나 실제 반도체 공정 운영에서는 주기적으로 정비를 수행하는 TBM(Time-Based Maintenance), 또는 PM(Periodic, or Preventive Maintenance) 방식으로, 공정 운영자의 전문 지식과 노하우, 정비 이력을 기반으로 정비 시점을 결정한다. 이에 따라, 적정 상태의 장비 운영을 위해서는 과도한 인적 및 시간적 비용이 투입되어야 한다. 뿐만 아니라, 이전 공정이 다음 공정에 영향을 주기 때문에, 공정 운영자는 이전 공정의 특성을 정확하게 파악해야만 적절한 시점에 정비를 수행할 수 있다.

다음으로, 웨이퍼 품질 추정을 위한 VM 기술은 여러 장비들을 통해 생산된 반도체 중 결함이 없는 합격품의 비율인 수율(yield)을 향상시키기 위한 기술이다[7,8]. 반도체 공정 과정은 기본적으로 여러 장비들로 여러 종류의 박막을 증착하고(deposition), 이를 식각(etch) 공정들의 연속으로 표현할 수 있다. 이 과정에서 주기적으로 모니터링(monitoring) 작업을 수행하여, 웨이퍼의 품질이 사전 정의된 사양(specification)에 만족하는지를 확인한다. 이 작업은 일반적으로 수십~수백 웨이퍼마다 진행되며, 전체 웨이퍼 중 일부만을 선별하여 여러 종류의 계측기에서 두께, 응력, 잔여물, 불량 등의 품질을 계측한다. 그러나, 이 작업은 계측기의 측정 정밀도, 작업자의 숙련도, 작업 주기에 따라 반도체 수율에 영향을 끼칠 수 있다. 예를 들어, 총 100개의 웨이퍼를 생산하고자 1, 50, 100번째 웨이퍼마다 모니터링 작업을 수행한다고 가정했을 때, 가령 모니터링 사이의 웨이퍼인 2번째부터 49번째, 그리고 51번째부터 99번째의 웨이퍼의 품질은 보장하지 못하는 문제가 있다. 만약, 위 상황에서 20번째 웨이퍼가 불량이 나더라도 작업자는 이를 인지하지 못하기 때문에, 해당 불량 웨이퍼는 이후 공정의 장비들을 그대로 통과해 반도체로 생산되어 수율 감소뿐만 아니라 불필요한 공정 시간과 재원을 소비하게 된다. 또한, 웨이퍼의 두께, 응력 등이 사전 정의된 사양을 벗어나는 경우에도 동일한 문제가 존재한다. 이를 방지하기 위해, 모니터링 주기를 감소시켜서 더욱 많은 표본의 웨이퍼에 대해 평가하는 방법도 있지만, 모니터링 작업에 소요되는 비용도 증가하기 때문에 이는 근본적인 해결책이 되지 않는다.

이와 같이, 웨이퍼 품질을 정확하게 추정하는 것은 곧, 수율 향상과 비용 감소에 직접적으로 영향을 주기 때문에, 정확도가 높은 VM 기술은 웨이퍼 제조사 입장에서 필수적인 기술이다. 또한 장비 제조사 입장에서도 VM 기술은 최적화된 공정 레시피(recipe)를 설계하기 위한 실험계획법(DOE, Design of Experiment) 시간을 감소시킬 수 있기 때문에, 인적 및 시간적 비용 절감 측면에서 매우 중요하다. 이 때, 레시피는 공정 수행 방법을 명시한 일련의 공정 작업 순서를 의미하고, 공정에 필요한 재원의 양, 온도, 압력 등의 수치가 기입되어 있다.

VM 기술은 웨이퍼의 품질을 추정한다면, APC는 웨이퍼의 품질을 사양에 만족하도록 공정 수행 중 공정 파라미터를 자동으로 제어하는 기술을 의미한다. 반도체 공정 중에는 모니터링을 수행하여 웨이퍼의 품질을 관리하는데, 가령, 웨이퍼 품질의 하락이 확인되면, 공정 전문가가 각종 파라미터와 계측된 웨이퍼 품질을 분석하거나 과거 경험 및 이력을 기반으로 제어 가능한 파라미터를 선별하여 미세 조정을 수행한다. 이러한 미세 조정 과정은 공정과 장비의 이해도가 높은 전문가가 수행하며, 해당 전문가의 수준에 따라 직접적으로 품질과 수율에 영향을 준다.

이와 다르게, 웨이퍼의 품질이 아닌 적정 품질의 웨이퍼를 양산할 수 있는 장비 상태를 유지하기 위한 기술인 Run-to-Run(R2R) 기술도 APC에 포함된다. 또한, 일반적으로 장비 내 결함을 탐지 및 분류하는 FDC(Fault Detection & Classification) 기술을 내포하기 때문에, 장비에서 FDC로 분류할 수 있는 결함이 탐지되었을 때, 그에 적합한 미세 조정을 수행한다. 예를 들어, 장비 챔버 내 온도가 사전 정의된 기준치 이하로 감소하였을 때, FDC가 이를 탐지 및 분류하고, 이를 기반으로 기준치에 맞는 수치로 R2R이 히터에 인가되는 전력량을 자동 조정한다. 이러한 조정 방법은 사전 각 결함 별 설정된 파라미터와 그 값으로 조정하거나 해당 장비 특성을 반영한 모델을 사용하여 조정할 수도 있다.

마지막으로, 반도체 제조 스케줄링 최적화의 목적은 장비 또는 장비 내 챔버의 가동율을 최대로 하는 작업 순서를 결정하여 Fab 전체 또는 각 장비 별 생산율(Unit Per Equipment Hour, UPEH)을 향상시키는 것이다. 앞서 기술한 바와 같이, 반도체 제조 스케줄링 최적화는 크게 Fab의 OHT 스케줄링과 장비의 VTM 스케줄링으로 구분된다. 먼저, Fab 내 OHT 스케줄링은 장비 간 웨이퍼 이송 시간인 Q-time의 감소, 특정 장비로만 웨이퍼 이송되어 발생되는 병목 현상(bottleneck phenomenon)의 방지, OHT가 어떠한 작업도 수행하지 못하고 서로 장비에 웨이퍼를 이송하거나 무한정 대기하는 상태인 교착 상태(deadlock)의 방지, 장비가 어떤 공정도 수행하지 않는 유휴(idle) 시간의 감소, 그리고 OHT 간의 충돌 방지 등

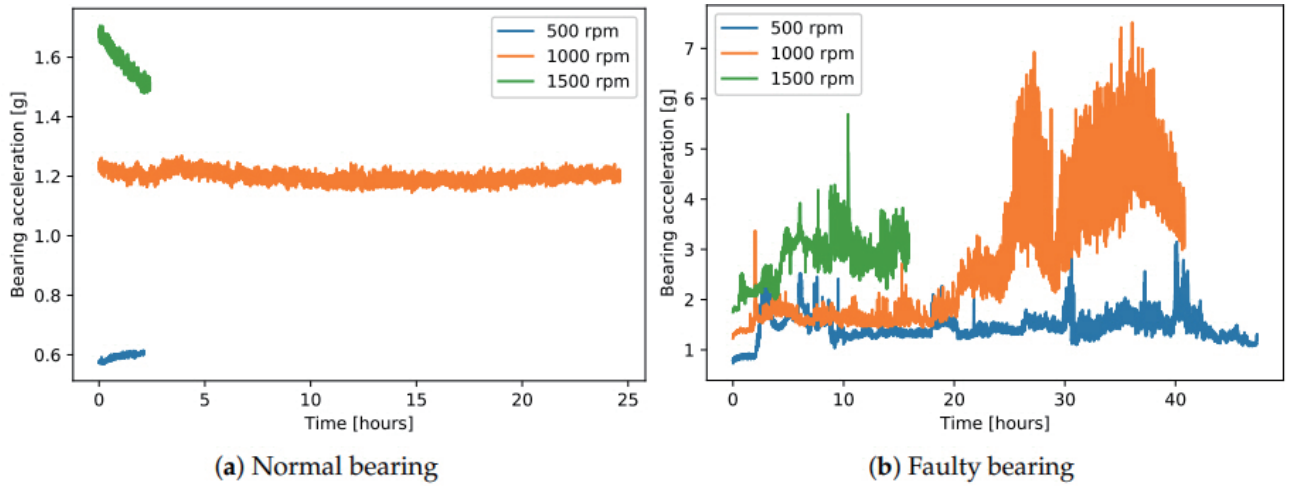
을 주된 목표로 설정된다. 이와 유사하게, 장비 내 VTM 작업 스케줄링은, 앞서 설명한 OHT 작업 스케줄링에서 OHT를 VTM으로, 장비를 챔버로 각각 간주하여 스케줄링 최적화를 수행한다.

이와 같은 스케줄링 방법은 스케줄러 개발자의 장비 이해도 및 숙련도에 따라 최적화 수준이 상이하여 수많은 시행 착오를 거쳐서 스케줄러가 개발된다. 다시 말해, 높은 수준의 스케줄링 개발에는 많은 인적 및 시간적 비용이 투입되어야 하고, 투입되더라도 항상 최적을 보장하기 어려운 문제가 있다. 이와 다르게, 양산 측면에서의 문제점은 OHT 스케줄링과 VTM 스케줄링의 호환이 불가능한 상태에서 OHT 스케줄링만을 최적화하더라도, 결국, VTM 스케줄링의 성능에 종속되는 문제가 존재한다. 가령, 다중 단계(multi-steps)의 공정을 여러 장비들로 수행하고자 한다면, 이는 웨이퍼 Q-Time의 증가로 이어지고, 다시 생산율의 하락으로 귀결될 수 있기 때문에, 일반적으로 단일 장비에서 다중 단계의 공정을 수행하기 위한 공정 레시피를 개발한다. 이러한 경우, Q-Time은 감소되었지만, 장비 내에서의 작업 복잡성(complexity)이 증가하여 VTM 스케줄링 방식의 고도화가 요구된다. 이를 반대로 설명하면, 고도화된 VTM 스케줄러가 존재한다면, 다중 단계의 공정을 단일 설비에서 최적화된 작업 순서로 처리할 수 있기 때문에, Q-Time 감소라는 이점을 그대로 유지할 수 있다.

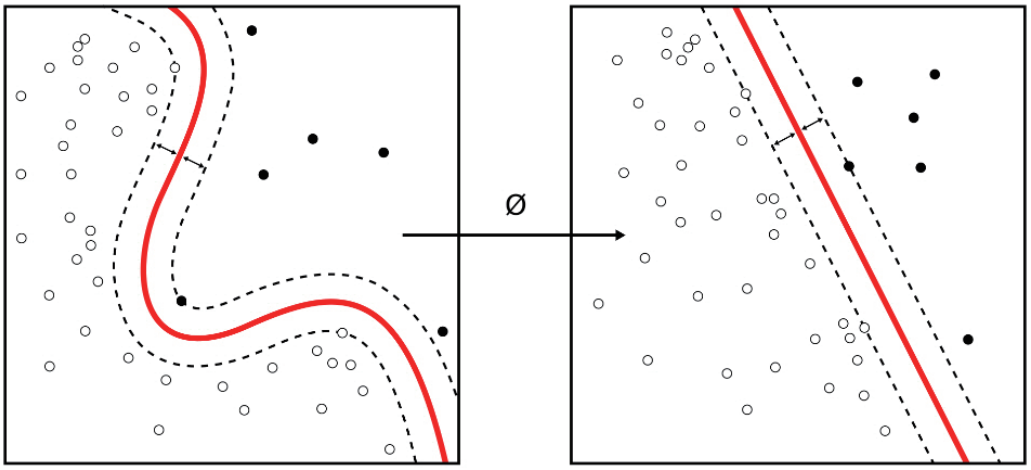
이렇듯 반도체 분야에서는 반도체 공정의 효율, 수율, 그리고 생산율을 향상시키기 위해, 전통적인 공정 및 설계 측면에서의 접근 방법이 아닌 최근 각광받고 있는 인공지능 기술 기반으로 기존 한계점을 접근하고 있다. 지금부터는 반도체 분야에서 해결하고자 하는 문제점과 인공지능 기술을 기반으로 해당 문제를 완화시키거나 해결한 사례를 기술하고자 한다.

2. AI 활용 사례 – 장비 결함 탐지

본 장에서는 반도체 장비의 건강도를 일정 수준으로 유지할 수 있도록 하고 인공지능 기반의 CBM을 수행하기 위한 FDC 기술에 대한 사례를 살펴본다. FDC는 인공지능 분야에서는 이상치 탐지(Outlier detection)의 일종으로 간주하며, 이상치는 정상(normal) 데이터와 일정 크기 이상의 차이를 보이는 데이터를 의미한다. 이 때, 공정 로그 데이터는 시간 흐름에 따른 데이터인 시계열(timeseries) 데이터이므로, 특정 시점 t에서의 단일 이상치(point outlier), 또는 여러 시점에 걸쳐서 연속 이상치(sequence outlier)가 존재할 수 있다. 또한, 반도체 공정은 매우 많은 부품들과 각 부품 별 센서(sensor)가 존재하여 계측되는 데이터도 매우 고차원이다. 일반적으로 수백 종류의 센서가 장비 내 탑재되어 있으며, 온도, 압력, 전류 등의 값을 수 초에서 수 백분의 일초 사이의 주기



〈그림 3〉 Data comparison for Normal and Faulty bearing

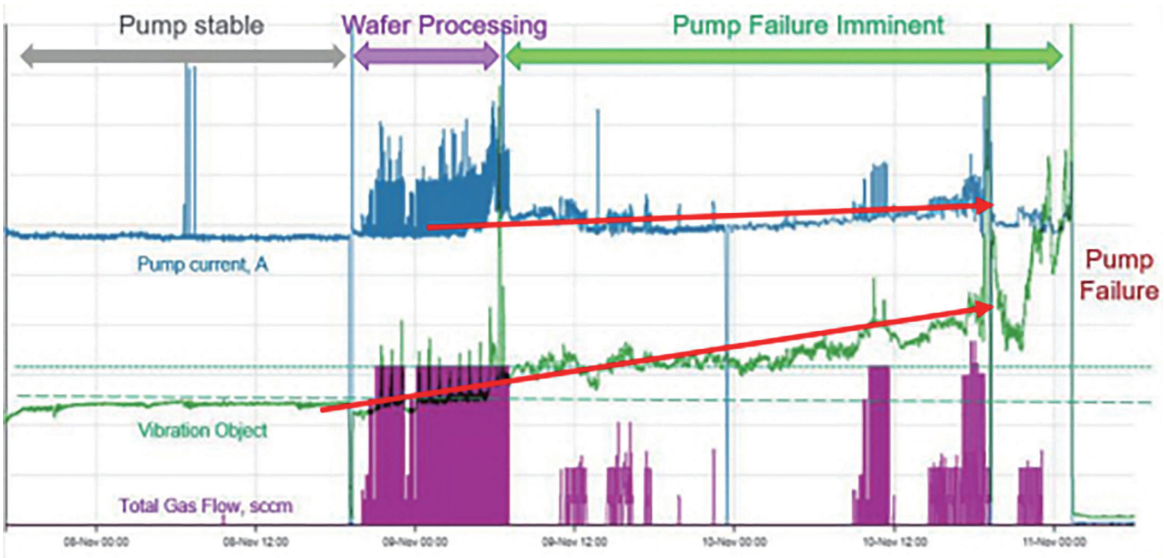


〈그림 4〉 Example for SVM *source: Wikipedia

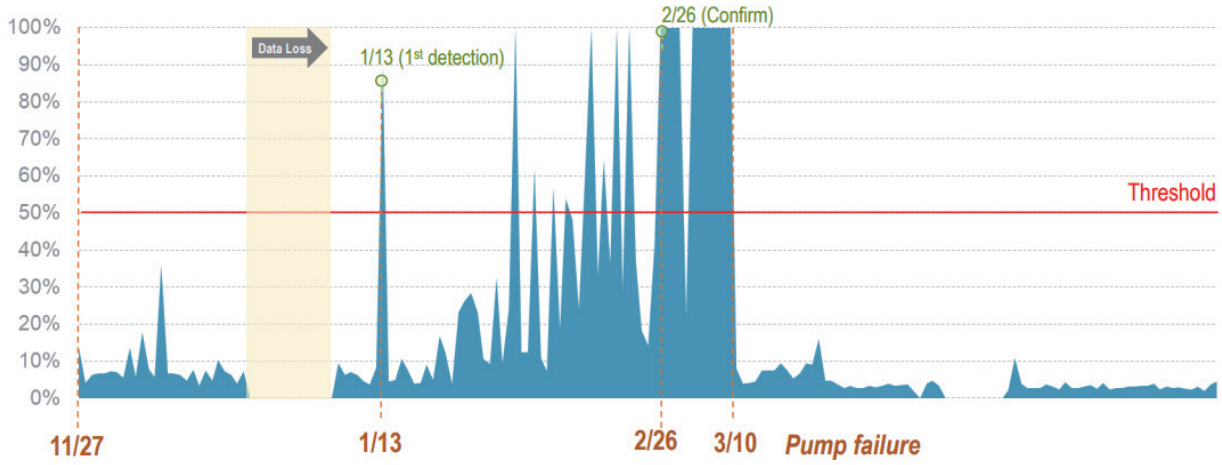
로 계측한다. 정리하자면, 반도체 설비에서 발생할 수 있는 이상치 탐지는 단순히 단일 측정 값에 대한 탐지가 아니라, 고차원의 연속된 데이터 집합에서의 연속된 이상치 탐지를 의미한다.

첫 번째 사례는 장비의 웨이퍼 이동 장치 중 하나인 DEU(Drive and Elevating Unit)의 결함을 탐지하는 모델을 실 환경에 적용하여 평가한 사례이다[9]. 여기서, DEU는 베어링(bearing)과 모터(motor)를 통해 웨이퍼의 회전과 상/하 이동을 가능케 해준다. 이 때, 베어링의 떨림(vibration) 수준, 즉, 결함 수준에 따라 해당 베어링에 부착된 가속도 센서 값의 변동폭이 커지거나 작아진다. 〈그림 3〉은 모터의 회전 속도(Revolutions Per Minute, RPM) 별 베어링의 정상 상태와 비정상 상태에서의 가속도 센서 값을 나타내며, 결함 발생 유/무에 따라 그 데이터 분포가 상이하다.

이러한 분류 과정을 자동화하기 위해, SVM(Support Vector Machine) 기반의 SVR(Support Vector Regression) 모델을 사용하였다. 여기서, SVR 모델은 결함이 발생한 데이터만을 학습하여 학습한 데이터와 유사하거나 상이한 정도를 결함이 발생한 확률로 추정한다. 다시 말해, 베어링의 가속도 값이 결함 데이터 분포와 유사한 분포로 변해가면, 점차 결함 발생 확률이 증가한다.



〈그림 5〉 Example for vacuum pump failure



〈그림 6〉 Health index for vacuum pump

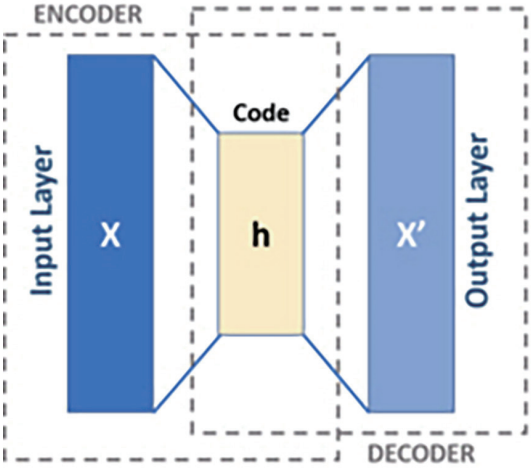
이 때, SVM은 분류 모델 중 하나로, 주어진 다차원 데이터 집합에서 분포가 상이한 여러 개의 부분 집합들 간 간극(margin)을 최대화하는 초평면(hyperplane)을 계산하여, 해당 데이터 집합을 분포가 상이한 여러 개의 부분 집합들로 분류한다. 〈그림 4〉는 한 데이터 집합을 비선형(nonlinear)과 선형(linear)으로 분류한 예시로, 적색 선은 초평면, 점선은 초평면과 부분 집합 간 간극을 의미한다. 이와 유사하게, SVR은 SVM을 시계열 데이터로 확장한 모델이다.

다음 사례는 첫 번째 사례와 유사하지만, 장비 정비 시점을 추정하기 위해 진공 펌프(Vacuum pump)의 결함을 탐지하는 모델을 실제 공정 환경에서 검증한 사례이다[10]. 문제 상황은 연속된 공정이 수행된 이후, 시간이 흐를수록 진공 펌프에 인가되는 전류 값, 진공 펌프의 진동 센서 값, 그리고 진공 펌프로 흐르는 가스(gas)의 유속(Standard Cubic Centimeter per Minute, SCCM)의 진폭 변화가 커지다가, 특정 시점 이후에 진공 펌프가 고장나는 상황이다〈그림 5〉. 여기서, 진공 펌프의 건강도를 0과 1사이의 실수로 나타내어, 사용자가 정의한 임계값(threshold) 이상이 되는 시점을 탐지하는 AI 기반의 모델을 개발하여 실제 공정 중 진공 펌프의 결함을 탐지한다. 이 때, 건강도는 결함이 발생할 확률 값으로, 0에 가까울수록 결함이 없는 상태이고, 1에 가까울수록 결함이 발생할 수 있는 상태이다.

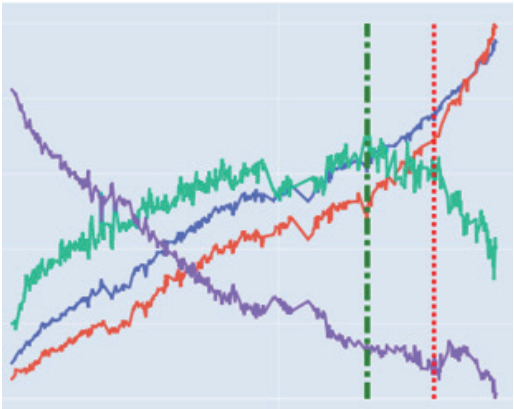
〈그림 6〉은 누적 공정에 따른 진공 펌프의 건강도를 나타낸 그림이고, 이 때, 임계값(적색 선)은 50%(=0.5)로 설정하였다. 결함이 발생한 시점(3월 10일)에 가까워질수록 건강도의 진폭이 커지면서 진폭 주기도 짧아지고 있다. 특히, 이전 시점(2월 26일)에서 결함 시점과 유사한 건강도 진폭 패턴으로 나타나기 때문에, 해당 건강도를 기준으로 진공 펌프에 대한 정비가 가능할 것이라 추정된다.

마지막 사례는 여러 개의 펌프들을 장착한 증착 장비에 대해, 펌프들의 압력 균형 문제가 발생할 수 있는 시점을 추정하는 모델을 실제 공정 환경에 적용한 사례이다. 증착 장비는 재원으로서 가스(gas)를 사용하여 초고온과 초고압으로 웨이퍼에 원하는 물질을 증착하는 공정을 수행한다. 이 때, 챔버 내부에는 증착되지 않고 남는 잔여 가스가 생길 수 있으며, 이는 다음 공정에 영향을 줄 수 있기 때문에, 펌프를 이용하여 장비 밖으로 배출시킨다. 이 과정에서 펌프 배관에는 일부 잔여물이 생길 수 있으며, 이는 펌프들의 압력 균형 문제를 야기시킬 수 있다. 이 균형 문제는 사전 파악이 중요하며 미리 펌프 배관에 대한 정비를 수행해야만 장비 고장을 미연에 방지할 수 있다.

이러한 펌프측 압력 균형 문제를 탐지하기 위해, 과거 해당 문제가 발생했던 데이터의 패턴을 학습하여 탐지 모델을 생성하고, 탐지 모델은 문제 데이터 패턴과 유사한 시점을 압력 균형 문제 시점으로 탐지한다. 이 때, 모델은 문제 데이터 패턴을 정확하게 학습하는 것이 매우 중요하다. 인공지능 분야에서 오토인코더(AutoEncoder,



〈그림 7〉 Structure of AutoEncoder *source: Wikipedia



〈그림 8〉 Example for issue of pumping pressure

AE) 모델은 주어진 입력 데이터를 모방하도록 학습하는 방법으로, 주어진 데이터의 패턴을 최대한 정확하게 학습한다. AE 모델은 인코더(Encoder)와 디코더(Decoder)를 한 쌍으로 갖으며, 인코더는 주어진 입력 데이터를 요약 데이터(code)로 부호화 하는 모델이고, 디코더는 주어진 요약 데이터를 원본 데이터로 복원하는 모델이다〈그림 7〉. 즉, AE 방법은 입력 데이터를 부호화하고 이를 다시 복원하는 과정을 반복적으로 거치면서, 입력 데이터와 원본 데이터의 차이인 재구축 오차(Reconstruction error)를 줄이는 방향으로 학습한다. 이 때, 각 인코더와 디코더를 구성하는 모델은 LSTM(Long Short-Term Memory) 모델로, 데이터 간의 순서 특성을 학습 및 추론할 수 있도록 개발되어 다차원 시계열 데이터 처리에 사용되는 모델들 중 범용적으로 사용되는 모델이다.

AE 모델은 학습한 데이터 패턴과 유사한 패턴을 갖는 데이터에 대해서 복원할 때는 낮은 재구축 오차를 갖지만, 상이한 패턴에 대한 복원에서는 상대적으로 높은 오차를 갖는다. 이러한 특성으로, 본 사례와 같이 문제 상황에 대한 데이터를 알고 있을 때, 유용하게 활용될 수 있다. 〈그림 8〉은 AE 모델을 활용하여 실제 펌프 압력 균형 문제 시점을 탐지한 결과를 데이터 패턴만 유지한 체 다시 그린 그림으로 실제 데이터와 상이할 수 있다. 이 때, 가로축은 시간을, 세로 축은 펌프 관련 파라미터와 재구축 오차를 0과 1사이로 정규화(normalization)한 값이다. 〈그림 8〉에서 각각 청색/적색/녹색 실선은 펌프 관련 파라미터, 보라색 실선은 재구축 오차, 진녹색 점선은 펌프 압력 균형 문제가 조짐이 보이는 시점, 그리고 적색 점선은 탐지 모델이 추론한 문제 시점이다. 이 때, 탐지 시점은 재구축 오차가 골짜기 모양처럼 "V"자 형태로 나타나는데, 이는 최근 데이터 패턴이 문제 데이터 패턴과 가장 유사해졌다가 이후, 그 패턴이 변화하기 때문이다. 다시 말해, 해당 탐지 시점 이후에는 학습한 과거 이력의 문제 데이터 패턴보다 더욱 악화된다고 판단할 수 있고, 이는 펌프 압력 균형이 과거보다 더 심하게 무너지는 것으로 추정할 수 있다.

3. AI 활용 사례 – 웨이퍼 품질 추정 및 자동 프로세스 제어

본 장에서는 인공지능 기반의 웨이퍼 품질을 추정하는 VM 기술과 웨이퍼 품질을 적정 수준으로 유지하는 APC 기술에 대한 사례를 살펴본다. VM과 APC를 인공지능 기술 기반으로 진행하기 위해서는 두 기술 모두 입력에 대한 출력이 존재하는 데이터가 필요하다. 예를 들어, 어떤 공정 로그 데이터에 대한 대응되는 두께 정보가 존재해야 한다. 그러나, 앞서 기술한 바와 같이, 모니터링 작업의 낮은 수행 횟수와 이에 따른 적은 수의 학습 데이터로 인하여, 완전한 데이터 중심(data-driven)의 VM/APC 기술은 사실상 불가능할 수 있다. 이에 따라, 공정의 특성과 챔버의 구조 정보 등을 활용한 물리 기반(physics-based) VM/APC 기술과 데이터 중심 VM/APC 기술을

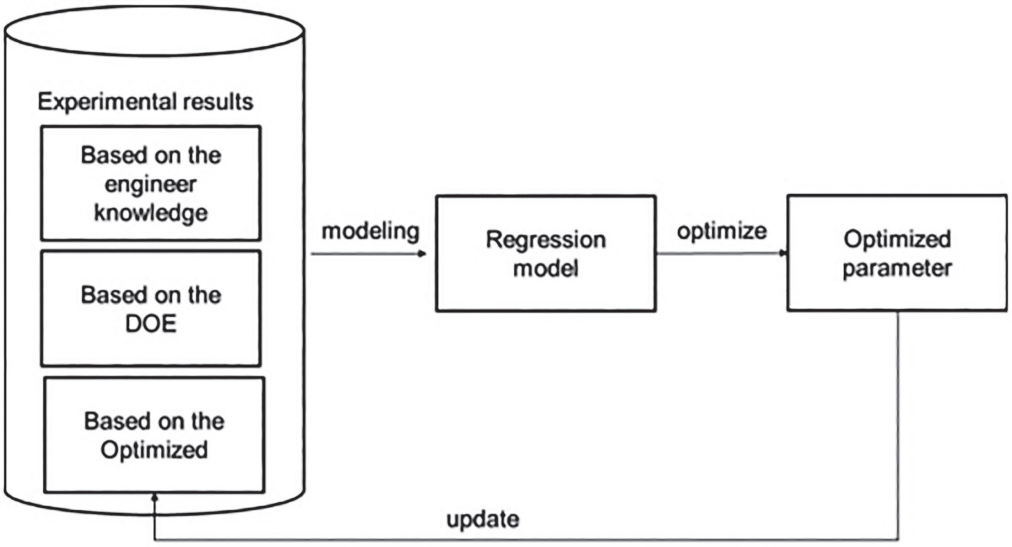
병합하는 학술적 연구가 진행되고 있다. 또한, 산업 분야에서는 물리 기반과 데이터 중심뿐만 아니라 공정 전문가 지식(knowledge)을 정량화하여 학습 데이터로서 활용하여 VM/APC 기술을 발전시키고 있다.

첫 번째 사례는 VLSI(Very Large Scale Integration) 플라즈마(plasma) 식각 공정을 대상으로, 적은 수의 공정 로그 데이터와 공정 전문가 지식을 인공지능 기반의 VM 모델로 개발하여 실제 공정 데이터에서 평가한 사례이다[11]. 식각 결과는 3D 모델 형태로 측정되기 때문에, 이를 학습이 가능한 형태로 변환이 필요하다. 이 사례에서는 각 식각 결과인 3D 모델로부터 총 4가지의 정량적 데이터–회로 너비, 마스크(mask) 너비, 시각 깊이, 잔여 마스크의 양–으로 변환된다. 또한, 공정 로그 데이터에서는 실제 시각을 수행하는 단계(step)들의 데이터만 추출하여, 파라미터와 식각 결과 간 상관 관계 분석을 통해, 학습할 파라미터가 선별된다. 식각 공정에 대한 VM 모델링은 딥러닝(deep–learning) 기반의 회귀 모델을 설계하여, 공정 로그 데이터를 입력, 식각 결과를 출력으로 하도록 학습이 수행된다. 이와 같이, 공정 로그 데이터나 식각 결과를 그대로 학습시키는 것보다 사전 전처리 방법을 어떻게 하느냐에 따라 결과가 상이하다. 예를 들어, 파라미터 간 데이터 범위가 극단적으로 차이가 존재할 때, 추정 결과가 특정 파라미터에 의해 종속된 결과를 얻을 수 있으므로, 0과 1사이의 값으로 정규화가 필요하다.

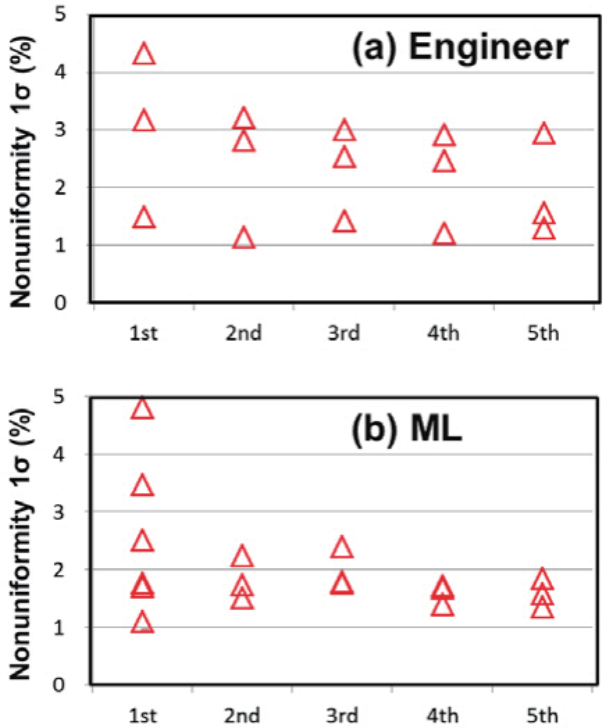
반도체 웨이퍼는 여러 증착 공정과 식각 공정을 거치면서 여러 레이어(layer)들로 구성되고, 이 때, 각 레이어는 집적 회로(Integrated Circuit, IC)를 의미한다. 여기서, 각 레이어가 다른 레이어들에 비해 얼마나 오정렬 되어 있는지에 대한 값을 오버레이(Overlay)로 표현한다. 두 번째 사례는 임플란트 레이어(implant layer)들에 대한 가상 오버레이 계측을 물리 기반과 데이터 중심의 모델을 수행한 사례이다[12]. 이 때, 임플란트 레이어는 총 9개이고 각 레이어 별로 동일한 패턴을 갖는 STI(Shallow–Trench–Isolation) 레이어를 쌓는다〈그림 9〉. 각 임플란트 레이어는 각기 다른 스캐너(scanner)로 오버레이 값이 측정되고, STI 레이어의 오버레이 값은 단일 스캐너로 측정된다. 이 두 번째 사례에서는 각 임플란트 레이어에 대한 스캐너의 측정 값을 추정하기 위해, STI 레이어의 오버레이 측정 값, 다음 번째 임플란트 레이어에 대한 다른 스캐너로 측정한 오버레이 값을 재귀적인 수식으로 설계한다. 이 때, 각 임플란트 레이어와 STI 레이어 간 왜곡(distortion) 등으로 인한 오차가 발생할 수 있기 때문에, 이를 딥러닝 모델을 이용하여 보정할 수 있도록 수식을 수정하였다.

세 번째 사례는 웨이퍼 품질 달성을 위한 공정 최적화 수행에 있어서 인공지능 기술을 이용한 사례이다[13]. 엄밀히 말해, 최적화된 공정 레시피, 즉, 사용자가 요구한 품질의 웨이퍼를 생성하기 위한 제어 파라미터 값을 찾기

〈그림 9〉 Implant layers and STI layer



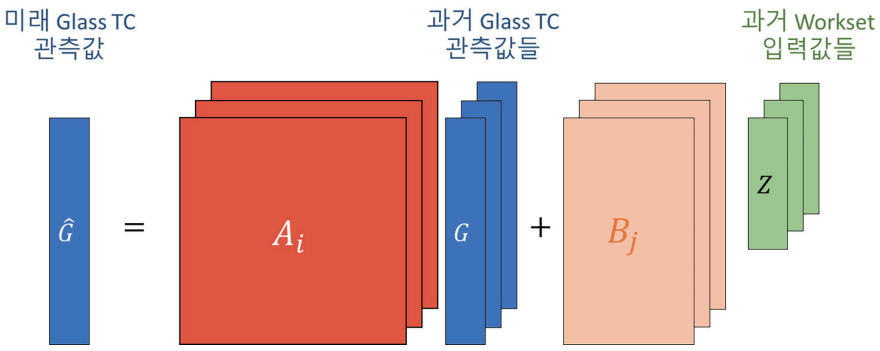
〈그림 10〉 Pipeline to explore optimal parameters



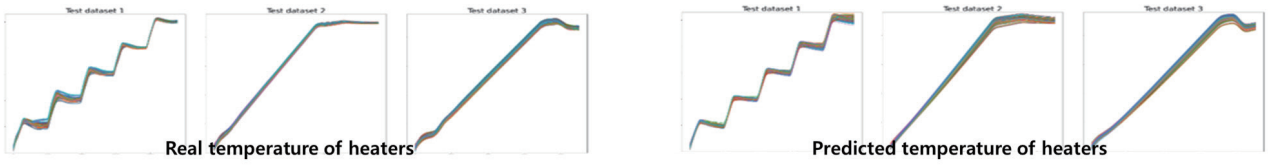
〈그림 11〉 Comparison for finding optimal parameters

위해, 공정 전문가는 수많은 DOE 과정을 수행하게 되며, 이는 과도한 시간이 소요될 수 있다. 본 사례에서는 웨이퍼의 박막에 대한 불균일성(non-uniformity)과 응력 수준을 감소시키기 위한 DOE 과정 시간을 줄이기 위해, 회귀 모델(regression model)을 사용하였다. 〈그림 10〉은 회귀 모델 기반의 특정 목표 수준을 만족하기 위한 최적의 파라미터 탐색하기 위한 과정을 도식화한 그림이다. 여기서, 공정 전문가의 지식, 과거 DOE 결과, 그리고 현재 최적화된 파라미터 값을 입력으로 하여 회귀 모델이 갱신된다. 정확하게는, 각 회귀 모델은 현 웨이퍼의 공정 로그와 이전 웨이퍼에서 최적화한 파라미터 값을 기반으로 각 계측 포인트에 대한 두께를 추정하는 모델이며, 각 계측 포인트 별 모델이 필요하다. 〈그림 11〉은 DOE 횟수가 1부터 5까지 증가함에 따라 전문가와 모델이 얼마나 빠르게 불균일성 목표를 달성할 수 있는 가를 보여주는 결과이다. 여기서, 첫 번째 DOE에서는 전문가와 모델의 차이가 적으나, DOE 횟수를 거듭할수록 모델은 빠르게 최적의 파라미터 값을 찾아서 불균일성을 낮추는 것으로 나타난다.

다음 사례는 장비 챔버 내 적정 온도 유지를 위한 히터 제어를 목표로 APC 달성을 위한 딥러닝 기반의 MPC(Model Predictive Control) 적용 사례이다. 이 때, 기존 반도체 장비보다 상대적으로 제어해야 할 히터 파라미터가 많은 디스플레이 장비를 대상으로 진행하였다. 히터 제어에 있어서 필수적인 기술인 PID(Proportional-Integral-Derivative) 제어는 매 단위 시간마다 히터의 온도를 측정하여 목표 온도와의 차이 값을 계산하고, 해당 차이 값을 줄이는 방향으로 히터를 제어하는 방법



〈그림 12〉 Abstract for MPC model



〈그림 13〉 Comparison between real and predicted temperature

을 의미한다. 본 사례에서는 제어해야 할 히터 온도의 과거 값과 히터 제어 값(workset)을 학습하여 미래 시점의 히터 온도를 예측하는 MPC 모델을 실 디스플레이 장비에 적용하였다. 〈그림 12〉는 적용한 MPC 모델의 개략도이며, 히터 온도와 히터 제어 값이 각각 Glass TC와 Workset이다. 여기서 A_i와 B_j는 실제 학습시킬 모델 파라미터로, MPC 모델의 학습 과정에서 히터 온도 및 제어 값의 패턴이 반영된다. 또한, 단순히, 다음 시점 t+1의 히터 온도를 예측하는 것이 아닌 t+k 시점까지의 히터 온도를 예측하여 더 빠른 히터 제어가 가능하다. 〈그림 13〉은 실제 히터 온도 값과 예측 온도 값을 시각화한 결과이며, 예측 온도는 미래 값 900 step까지 예측한 결과이다. Step이 증가하더라도 사전 정의된 사양내로 예측되는 것으로 나타났다.

4. AI 활용 사례 – 공정 스케줄링 최적화

본 장에서는 반도체 장비의 하드웨어 구조와 공정 레시피 시간에 따라 장비 내 공정 스케줄링의 최적화를 수행하는 인공지능 기술 기반의 적용 사례를 살펴본다. 공정 스케줄링은 여러 장비 조건, 즉, 하드웨어 구조의 복잡도, 공정 조건, 공정 레시피 종류, 그리고 공정 요구 사항에 따라 장비 운영 복잡도가 매우 높아질 수 있으며, 이에 따라 기존 공정 스케줄러의 개선이 필요하다. 그러나, 이 공정 스케줄러의 개선은 장비 운영 복잡도에 따라 난이도가 증가하며, 특히, 규칙 기반(rule-based) 스케줄러의 경우, 매 공정 교칙과 예외 사항을 추가할 때마다 매번 동작 평가와 취약점을 파악해야 되기 때문에, 그 개발 시간이 기하급수적으로 증가한다. 다시 말해, 스케줄러 개발자는 장비 공정 중 발생할 수 있는 모든 경우의 수를 고려하여 스케줄러를 개발해야 하며, 다중 VTM들 간의 교착상태(deadlock)와 기아 상태(starvation) 등의 예외도 동시에 고려해야 한다. 또한, 높은 수준의 UPEH와 챔버 가용율(utilization)도 만족해야 한다.

현재, 반도체 장비의 공정은 여러 장비를 거치지 않고 단일 장비에서 여러 단계의 공정을 거치는 추세로 변화되고 있다. 다양한 이유가 있겠지만, 공정 스케줄링 측면에서는 서론에서 기술한 바와 같이, 반도체 장비 간 웨이퍼 이송으로 인한 병목 현상에 따른 이송 지연이 그 이유가 될 수 있다. 가령, 양산 Fab 내에서 장비 간 웨이퍼 이송에 대한 OHT 스케줄링을 최적화하더라도, 병목 현상이 발생하는 장비 내 공정 스케줄링이 최적화되지 않으면, 최적화된 OHT 스케줄링의 효율을 얻을 수 없다. 이와 반대로, 장비 자체에서 웨이퍼 생산이 늦어지면 OHT의 대

기가 발생할 수 있고, 이는 여러 OHT들이 단일 장비에 대해 경쟁 상태가 발생하여 웨이퍼 이송에 지연이 발생할 수 있다. 따라서, 공정 스케줄링 측면에서 최대한 단일 장비 내에서 더 많은 단계의 공정을 수행하여 OHT의 병목과 경쟁 상태를 유발하지 않는다면, 전체 Fab 생산율을 증가시킬 수 있을 것이다.

최근에는 공정 스케줄링의 효율을 높이기 위한 연구로서, 강화학습(Reinforcement Learning) 기반의 연구가 다수 진행되고 있다. 기존에는 TEG(Timed Event Graph)를 기반으로, VTM이 고정되고 다수의 챔버들로 구성된 클러스터 툴(Cluster tool) 형태의 연구들이 주를 이루었다[14]. 여기서, TEG는 TPN(Timed Petri Net)의 일종으로, 동시 접근과 병렬 시스템에 대한 행동을 분석할 수 있는 수학적 모델링 방법인 페트리 넷(Petri Net)에 시간 특성을 접목한 방법이다.

강화학습 기반 공정 스케줄링 방법의 차별성은 장비 운영 복잡도의 증가로 고려해야 되는 경우들을 모델이 대신 탐색하면서 최적의 스케줄링 규칙을 도출하는 것으로, 스케줄러 개발 시간을 단축시킬 수 있다. 강화학습은 에이전트(agent)라 하는 어떤 행동(action)을 수행하는 수행자가 있고, 이 에이전트는 주어진 환경(environment)의 상태(state)에서 보상(reward)을 최대로 하는 행동을 선택한다〈그림 14〉. 이 때, 결정된 행동은 다음 상태에 영향을 주며, 그에 따라 보상의 수준도 변화될 수 있다. 가령, 장비의 VTM을 에이전트라 가정했을 때, 장비는 환경이고 챔버 내 남은 공정 시간 등은 상태가 될 수 있다. 그리고 VTM이 특정 챔버에서 공정 완료된 웨이퍼를 꺼내는 것을 행동, 이 행동으로 인한 UPEH를 보상으로 간주될 수 있다.

〈그림 14〉 Abstract for Reinforcement Learning (RL)

지금부터 강화학습 기반의 공정 스케줄링을 위한 사례 중 연구 장비에 실제 탑재하여 UPEH를 개선한 사례를 살펴본다. 본 사례는 기존 규칙 기반의 스케줄러 개발 시간이 감소시키고, 공정 운영 조건이 증가하더라도 적정 UPEH를 유지하는 스케줄러를 개발하기 위해, 강화학습 기반의 공정 스케줄러를 개발하여 연구 장비에 탑재 및 검증한 사례이다. 공정 스케줄링을 위한 강화학습 모델의 학습 방법은 에이전트(e.g., VTM)가 에피소드(episode)라 하는 일련의 연속된 행동을 임의로 수행하고, 행동에 대한 가치(valuable)를 평가한다. 이 평가 결과에 따라 다음 에피소드에서의 행동 규칙이 결정되며, 이러한 과정을 UPEH가 수렴할 때까지 반복한다. 정리하자면, 본 사례에서의 강화학습 모델 학습 방법은 행동과 평가를 반복적으로 수행하면서, 현 상태에서 가장 좋은 행동 규칙을 도출한다. 물론, 이것이 최적임을 보장하는 것은 아니지만, 에이전트에게 최대한 모든 경우의 수를 경험하도록 학습되어야만 최적에 근사된 결과를 얻을 수 있을 것이다. 결과적으로, 본 사례에서는 기존 운영되고 있는 스케줄러 대비 약 2.31% 향상된 UPEH를 얻을 수 있었다.

5. 결론

지금까지 장비 결함 탐지, 웨이퍼 품질 추정 및 자동 프로세스 제어, 그리고 공정 스케줄링 최적화 등 반도체 장비에서의 인공지능 기술을 활용한 여러 사례에 대해서 살펴보았다. 인공지능 기술을 반도체 장비에 적용 시도가 증가하는 이유는 반도체 장비의 효율, 수율, 그리고 생산율의 향상 뿐만 아니라, 기존 방법으로는 해결하지 못한 문제점들을 신기술을 통해 시도하고자 하는 부분도 있을 것이다. 그러나, 실제 반도체 장비 환경의 제약 등으로 인하여 단순히 인공지능 기술로만 장비의 문제점들을 해결하기에는 어려울 수 있으며, 사전에 구축되거나 협의되어야 할 선행 작업들이 필요하다.

인공지능 기술을 반도체 장비에 적용함에 있어서, 사전에 고려되어야 할 선행 작업으로는 데이터 수집 및 정제 환경 구축, 데이터 양식(format) 통일화 및 체계적 관리, 그리고 데이터 중심의 회사 규약이다. 반도체 장비의

특성상 각 반도체 장비에서는 매초마다 수십에서 수백개의 고차원 시계열 데이터가 생성되며, 웨이퍼의 품질 데이터는 계측 대상에 따라 이미지 등으로 적재된다. 또한, 반도체 장비의 각 부품에 대한 교체 이력 및 사양과, 반도체 장비에서 발생한 이슈에 대한 후속 조치 방법은 글자(text) 형태로 적재된다. 이와 같이, 반도체 장비를 중심으로 생성되는 데이터는 시계열, 이미지, 글자 등 다양한 형태를 갖으며, 그 수도 매초마다 기하급수적으로 증가된다. 따라서, 이를 수집 및 정제하기 위한 업무 규약 정립과 데이터 관리 시스템 구축이 필요하다. 최근에는 MLOps나 DataOps와 같이, 소프트웨어(software) 개발의 정보 처리 과정을 인공지능이나 데이터 관리에서 활용하는 방법과 오픈소스 프레임워크(opensource framework)들이 배포되고 있기 때문에, 데이터 관리 시스템으로의 활용을 고려해볼 수 있다.

장비 데이터 중 계측 데이터의 경우에는 자동화하여 데이터 관리 시스템에서 수집 및 정제가 가능하지만, 작업자가 직접 입력하는 데이터(e.g., 교체 이력, 후속 조치 방법 등)는 작업자의 업무 방식에 따라 상이할 수 있으며, 경우에 따라, 완전히 상이한 형태로 데이터가 입력될 수 있다. 이는 데이터 관리의 복잡도가 증가하는 것을 의미하며, 이를 해결하기 위해 각 작업자 별로 별도의 데이터 해석(parsing) 기능이 필요하기 때문에, 매우 비효율적인 데이터 관리로 이어질 수 있다. 따라서, 이러한 데이터의 수집은 데이터의 형식이나 식별자 등에 대한 규칙을 업무 규약(protocol)으로 정립하는 것이 올바른 방향이라 할 수 있다.

참고문헌

[1] Jing Zhang, and Dacheng Tao, "Empowering things with intelligence: a survey of the progress, challenges, and opportunities in artificial intelligence of things", Journal of IEEE Internet of Things, Vol. 8, No. 10, pp. 7789–7817, 2020

[2] Imran Ahmed, Gwanggil Jeon, and Francesco Piccialli, "From artificial intelligence to explainable artificial intelligence in industry 4.0: a survey on what, how, and where", Journal of IEEE Transactions on Industrial Informatics, Vol. 18, No. 8, pp. 5031–5042, 2022

[3] Tianyu Wu et. al., "A brief overview of ChatGPT: The history, status quo and potential future development", IEEE Journal of Automatica Sinica, Vol. 10, No. 5, pp. 1122–1136, 2023

[4] Yiheng Liu et. al., "Summary of chatgpt-related research and perspective towards the future of large language models", Journal of Meta-Radiology, Vol. 1, Issue. 2, pp. 100017, 2023

[5] Yassine Qamsane et. al., "A unified digital twin framework for real-time monitoring and evaluation of smart manufacturing systems", IEEE International conference on automation science and engineering (CASE), pp. 1394–1401, 2019

[6] James Moyne et. al., "A requirements driven digital twin framework: Specification and opportunities", Journal of IEEE Access, Vol. 8, pp. 107781–107801, 2020

[7] Jianshe Feng et. al., "An online virtual metrology model with sample selection for the tracking of dynamic manufacturing processes with slow drift", Journal of IEEE Transactions on Semiconductor Manufacturing, Vol. 32, No. 4, pp. 574–582, 2019

[8] Hong-Goo Lee et. al., "Virtual overlay metrology for fault detection supported with integrated metrology and machine learning", "Metrology, Inspection, and Process Control for Microlithography XXIX", Vol. 9424, pp. 610–615, 2015

[9] Federico Pittino et. al., "Automatic anomaly detection on in-production manufacturing machines using statistical learning methods", Journal of Sensors, Vol. 20, No. 8, pp. 2344, 2020

[10] Andreas Neuber, and Jim L'Heureux, "Sustainability improvements in semiconductor manufacturing using smart manufacturing technologies", Applied Materials, 2022

[11] Jack Kenney et. al., "Deep Learning Regression of VLSI Plasma Etch Metrology", arXiv submitted, 2019

[12] Leon van Dijk et. al., "Towards excursion detection for implant layers based on virtual overlay metrology", Semi Advanced Semiconductor Manufacturing Conference (ASMC), pp. 1–6, 2020

[13] Yusuke Suzuki et. al., "Machine learning approaches for process optimization", International symposium on

Semiconductor Manufacturing (ISSM), pp. 1–4, 2018

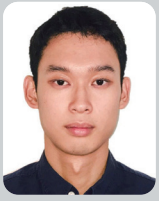
[14] Johannes Trunk et. al., "Model decomposition of timed event graphs under partial synchronization in dioids", International Federation of Automatic Control (IFAC), Vol. 51, No. 7, pp. 198–205, 2018

김홍연

- 2011년 2월 한국기술교육대학교 컴퓨터공학 학사
- 2013년 2월 한국기술교육대학교 컴퓨터공학 석사
- 2019년 2월 한국기술교육대학교 컴퓨터공학 박사
- 2019년 3월 ~ 2021년 1월 롯데정보통신 AI 기술팀
- 2021년 2월 ~ 현재 원익IPS SW팀 과장

〈관심 분야〉
Deep learning, Outlier detection, Prognostic and Health Management, Graph data analysis

인공지능을 활용한 기계의 건전성 예측 및 관리 방법



김관필
중앙대학교

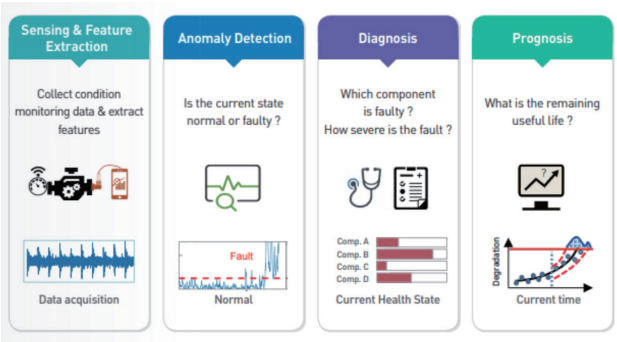
김상범
한국폴리텍대학

Ⅰ. 서론

1차 산업혁명 시대부터 현대의 4차 산업혁명까지 산업계에서 기계는 필수적인 요소로 자리 잡고 있다. 특히 현대 산업에서 기계는 복잡성과 정밀도가 증가하고 있고, 대량생산 체계를 통해 수많은 물건을 소비자에게 제공하고 있다. 복잡해진 기계를 통한 공정에 따라 장비의 안정성과 신뢰성이 중요하며 예상치 못한 기계 결함으로 인한 상품성 저하 또는 생산 지연에 따른 비용과 피해를 최소화하기 위한 많은 노력이 이루어지고 있다.

산업계에서 기계의 유지 보수는 생산 활동 못지않게 중요하며 이에 따라 기계 건전성 예측 및 관리와 같은 관리 전략을 도입하고 있다. 특히, 기업 입장에서 최소한의 유지비용으로 기계의 신뢰성, 가용성, 수익성 및 안정성을 유지[1]하는 것을 목표로 하고 있으며, 이러한 전략은 고장 발생 전에 예방적 조치를 취함으로써 장비의 수명을 연장하고 전체적인 운영 효율성과 상품성을 향상시키는 데 중점을 두고 있다.

기계의 유지 보수 전략은 다음 세 가지 주요 방법으로 발전해왔다. 첫째, 반응형 유지보수는 기계가 고장 나면 수리를 진행하는 전통적인 방식이다. 이 방법은 단기적으로 비용을 절약할 수 있지만, 예기치 않은 고장으로 인한 생산 중단과 같은 리스크가 크다. 둘째, 예방적 유지보수는 정기적인 검사와 부품 교체를 통해 잠재적 고장을 미리 방지하는 접근법이다. 이 방식은 예측 가능한 유지 관리 스케줄을 통해 기계의 수명을 연장할 수 있다. 셋째, 예측적 유지보수는 센서 기술과 데이터 분석을 활용하여 장비의 상태를 지속적으로 모니터링하고, 장비의 고장을 예측한다. 이 방법은 고장이 발생하기 전에 적절한 조치를 취함으로써 비효율적인 유지보수 작업을 최소화하고, 유지보수 비용을 효과적으로 관리할 수 있다. 최근에는 인공지능(AI) 기술이 유지보수 전



〈그림 1〉 PHM의 주요 4 단계[2]

략에 통합되면서 더욱 정밀한 고장 예측과 효율적인 유지보수 계획 수립이 가능해지고 있다.

PHM (prognostics and health management) 시스템은 기계의 건강 상태를 예측하고 관리하는 중요한 유지보수 전략 중 하나로 자리 잡았다. PHM 시스템의 주요 단계는 데이터 수집 및 전처리, 이상 탐지, 고장 진단, 유효수명 예측으로 이루어져 있다. 초기 접근 방법으로 통계적, 물리적, 수학적 방법을 이용해 기계의 상태를 분석하고 고장을 예측했으나, 이러한 방법들은 데이터가 충분하지 않거나 복잡한 비선형 시스템을 모델링하는데 한계가 있었다. 최근 인공지능과 IoT (internet of things) 기술의 발전으로 방대한 양의 데이터가 수집되고 있으며, 이를 바탕으로 한 인공지능 기반의 분석 방법이 도입되고 있다. 이러한 기술 발전에 따라 기계의 유지보수를 위한 다양한 연구가 이루어지고 있으며, 이러한 최신 기술들을 PHM 시스템의 각 단계에 적용함으로써 보다 정확하고 효율적인 고장 예측 및 유지 보수가 가능해지고 있다.

본 기고문에서는 〈그림 1〉과 같이 PHM 시스템의 각 단계별 활발히 사용되는 AI 기반 모델의 원리와 이를 응용한 관련 연구 동향과 AI를 활용한 접근 방법의 한계점을 해결하기 위한 연구 동향에 대해 소개하고자 한다.

Ⅱ. 데이터 수집 및 전처리

기계의 현재 상태를 모니터링 및 분석하기 위해서 데이터 수집 및 전처리 과정이 필수적이다. 하지만, 데이터의 불균형, 서로 다른 도메인 차이로 인한 일반화된 성능이 나타나지 않는 문제가 발생하기도 한다. 본 절에서는 위 문제를 해결하기 위한 기존의 접근 방법과 인공지능을 활용한 방법에 대한 연구를 살펴본다.

다음 하위 절에서 소개할 연구들은 고장 진단과 예측 등에 관련한 연구이나, 데이터 수집 및 전처리에 대한 내용에 집중하여 소개할 것이다.

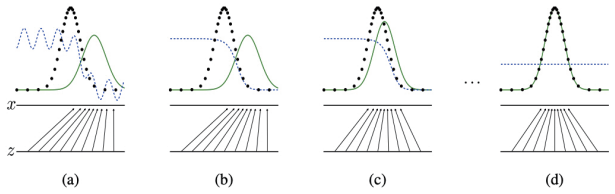
1. 데이터 수집

PHM 시스템에서 상태 모니터링을 위해 데이터 수집은 필수적이다. 최근 임베디드 기술과 클라우드 기술의 급속한 발전으로 인해 다양한 센서로부터 실시간 데이터를 수집하고 클라우드 기반 데이터베이스 시스템에 저장할 수 있게 되었다. 이렇게 수집된 방대한 양의 데이터에서도 데이터의 불균형이 존재한다. 이는 일반적으로 수집되는 데이터는 정상 상태를 나타내는 것이 대부분이기 때문이다. 고장진단, 유효 수명 예측 등 학습 모델에 활용하기 위해서 다양한 종류의 결함 데이터를 필요로 하는데, 이러한 데이터 불균형 문제를 해결하기 위해 실험실에서 인위적인 결함을 구현하여 데이터를 수집하기도 한다. 하지만 결함을 구현하기 위해 많은 비용이 투자될 뿐만 아니라 수많은 조합의 결함을 모두 구현할 수도 없다는 문제가 있다. 이를 보완하기 위해 시뮬레이션을 통한 데이터 수집도 이루어지고 있다. 특히 디지털 트윈[3]을 이용한 방식은 실제 물리적 개체를 디지털 환경에서 시뮬레이션하는 기술로, 이를 통해 실제 환경에서는 구현하기 어려운 다양한 조건을 시험하고, 데이터를 수집할 수 있어 많은 연구가 이루어지고 있다. 하지만 디지털 트윈을 통한 데이터 생성은 데이터 수집에 드는 비용을 절감할 수 있으나, 실제 물리 객체를 정확하게 구현하는데 난이도가 존재해 한계를 가진다.

반면에 인공지능을 활용해 데이터 불균형을 해결하기 위한 방안으로 데이터 증강 또는 생성하는 방법이 존재한다. 기존의 방법과 반대로 전문적인 물리적 구현 또는 이해가 필요가 없는 장점이 존재한다. 일반적으로 데이터 증강 또는 생성은 적대적 생성 신경망 (generative adversarial networks, GAN)[4] 기반 모델을 이용한다. GAN은 생성 모델 (generator, G)과 판별 모델 (discriminator, D)로 구성된다. 생성 모델 G는 실제 데

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{\text{data}}(x)}[\log D(x)] + \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z)))]$$

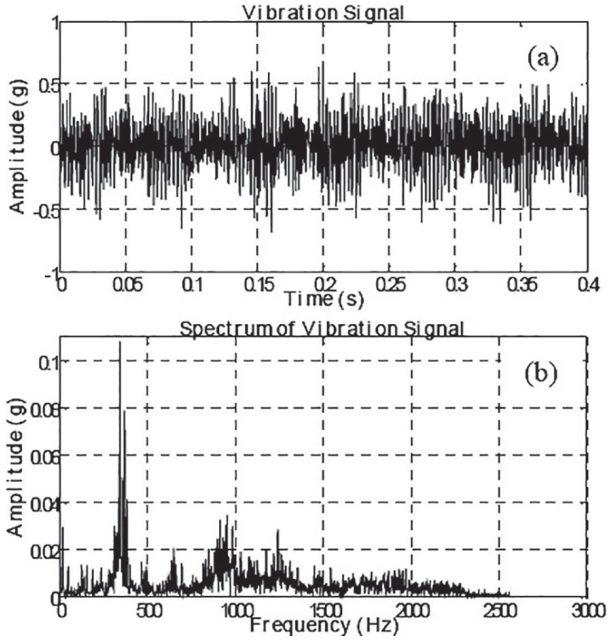
〈수식 1〉 GAN 손실함수



〈그림 2〉 GAN 학습 과정 도식화[4]

이터와 구분할 수 없는 가짜 데이터를 생성하고, 판별 모델 D는 제공된 데이터가 실제 데이터인지 생성된 가짜 데이터인지 분류한다. 두 모델은 적대적 학습을 통해 〈수식 1〉의 손실 함수를 기반으로 학습이 진행된다. G는 생성한 가짜 데이터 G(z)를 D에 입력하고, D(G(z))의 값을 1에 가까워지게 학습하고, G에 의해 생성된 가짜 데이터 G(z)에 대해서는 D(G(z))의 값을 0에 가까워지게 학습이 이루어진다. 학습은 〈그림 2〉와 같이 학습 초기 가짜와 진짜의 분포가 달랐으나, 학습이 이루어지면서 가짜가 진짜의 분포에 맞게 학습이 이루어지고, D는 구분을 할 수 없어 확률이 1/2이 되는 모습을 볼 수 있다.

GAN을 기반 모델로 하는 데이터 생성 관련 연구로, shao et al.[5]은 보조 분류자 생성적 적대 신경망(Auxiliary Classifier Generative Adversarial Network, ACGAN) 기반의 프레임워크를 사용하여 기계적 센서 신호로부터 실제적으로 보이는 일차원 원시 데이터를 생성하는 방법론을 제안했다. Liu et al.[6]은 개선된 Wasserstein 생성적 적대 신경망(Wasserstein Generative Adversarial Networks with Gradient Penalty, WGAN-GP)을 사용하여 실패 데이터를 증강하는 방법을 제안했다. 이 방법은 다차원 센서 데이터를 2차원 회색 이미지로 변환한 후, 불균형 데이터 샘플의 회색 이미지를 WGAN-GP의 입력으로 사용하여 적대적 훈련을 수행해 데이터를 생성했다.



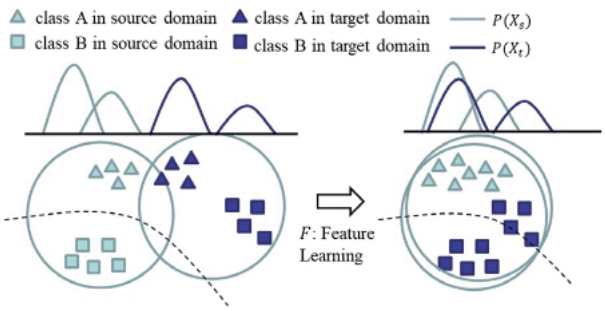
〈그림 3〉 시간 도메인 진동 신호, 주파수 도메인 진동 신호[7]

2. 데이터 전처리

일반적으로 결측치 제거, 데이터의 정규화 및 표준화 등 기본 데이터 전처리 과정 이후, AI 모델 학습, 예측 및 진단에 필요한 고급 전처리 또는 변환을 필요로 한다[8]. 센서로부터 수집되는 원시 데이터는 기계의 현재 상태를 분석하기엔 신호의 잡음 또는 기계의 특징을 나타내는데 한계를 가진다. 원시 데이터에서 보다 많은 정보를 얻기 위해 대표적으로 FFT (fast fourier transform) 변환이 사용된다. FFT는 시간 도메인의 신호를 주파수 도메인으로 변환하는 기법으로, 〈그림 3〉과 같이 주파수 도메인으로 변환 후 각 주파수 대역별 강조되는 진폭 특징을 쉽게 파악할 수 있다. 이 변환은 기계의 다양한 조건에 따라 주파수 대역이 분포되는 방식을 쉽게 파악할 수 있게 도와준다[9]. FFT는 연속적인 신호에 대한 CFT (continuous fourier transform)으로, CFT는 〈수식 2〉와 같이 계산되며, f(t)는 시간 도메인의 연속 함수를 나타내고, F(w)는 해당 함수의 주파수 도메인에서의 복소수 표현을 나타낸다. 적분은 모든 시간에 대해 수행되고, w는 각주파수를 나타낸다. 이 변환을 통해 시간 도메인에서 주파수 도메인으로 전환이 가능하게 된다.

$$F(w) = \int_{-\infty}^{\infty} f(t)\, e^{-iwt}\, dt$$

〈수식 2〉 Fast fourier transform



〈그림 4〉 도메인 적응 시각화[10]

또한 도메인 차이로 인해 발생하는 문제를 해결하기 위한 전처리(변환)도 존재하는데, 도메인 차이란 〈그림 4〉와 같이 환경적, 운용 조건 차이 등으로 인해 AI 모델 학습에 사용된 데이터와 실제 시나리오에서 수집된 데이터 간 차이가 발생하는 것으로, 이 경우 고장 진단, RUL 예측 등 AI 모델이 원하는 성능이 나오지 않을 수 있는 문제가 존재한다. 도메인 차이를 극복하기 위한 대표적인 방법으로 도메인 적응 방법을 사용하는데, 도메인 적응이란 모델이 소스 (학습) 도메인과 타겟 (실제 적용될) 도메인 데이터의 분포 차이를 줄이는 것으로 서술한 적대적 학습 메커니즘을 기반으로 사용된다.

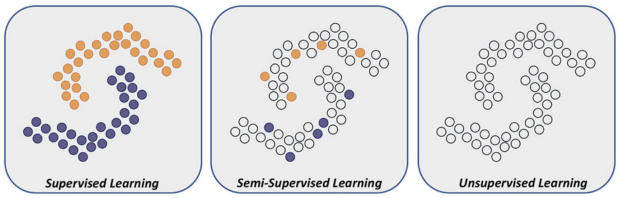
적대적 학습을 기반으로 한 도메인 적응 관련 연구로, Ozdagli et al.[11]은 SHM (structural health monitoring)을 위한 도메인 적응 방법을 제안했다. 이 연구는 레이블이 있는 소스 도메인과 레이블이 없는 타겟 도메인 간의 통계적 차이를 해결하기 위해 DANN (domain adversarial neural network)을 활용한 것으로, 두 도메인 데이터를 잠재적 특성 공간에 일치시켜 예측 정확도를 향상 시킨 연구이다. Jang et al.[10]은 다양한 작업 조건 및 고장 상태에서 RUL 예측을 위한 DDAN (deep domain adaptative network)을 제안했다. 이 방법론은 SCRNN (selective convolutional recurrent neural network)을 사용해 센서 신호에서 시공간 특성을 추출하고, 소스와 타겟 도메인 간 특성 분포를 조정하는 협력적 도메인 정렬 메커니즘을 사용해 도메인 차이를 최소화하는 방법을 제안했다.

Ⅲ. 이상탐지

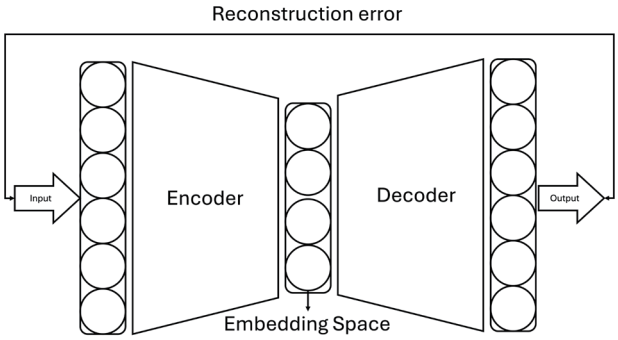
PHM 시스템에서 이상탐지란 기계나 시스템이 예상된 동작과 일치하지 않는 데이터의 패턴을 찾는 것을 말한다[12]. 이상탐지의 주요 목적은 잠재적인 문제를 조기에 발견하는 것으로 이를 통해 예방 조치를 취해 고장을 예방할 수 있게 하는 것이다. 인공지능을 활용한 이상탐지는 대부분 반지도 (semi-supervised) 학습 또는 비지도 (unsupervised) 학습 방식의 이상 탐지 방법을 사용한다.

반지도 방식의 이상탐지는 〈그림 5〉와 같이 레이블이 지정된 소량의 데이터와 레이블이 없는 데이터를 이용해 이상탐지 모델을 학습시키고, 이를 기반으로 이상탐지를 하는 방법을 말한다. Denkena et al.[14]은 LOF (local outlier factor)와 KNN (K-nearest neighbors) 또는 ABOD(angle-based outlier detection)를 앙상블한 모델을 이용해 이상탐지 하는 방법을 제안한 것으로, 특정 통계적 모델과 달리 데이터의 기본 분포에 대한 가정 없이 이상탐지를 하는 방법을 제안했다. Longlong et al.[15]은 hybrid C-LSTM를 이용해 시공간 데이터로 부터 이상을 탐지하고 WFDA (weighted fuzzy-rough density algorithm)을 이용해 실시간으로 이상 탐지 결과를 비교하고 신뢰되지 않는 데이터를 재학습하여 이상탐지 정확도를 개선하는 방법을 제안했다. 반지도 방식은 소량의 레이블이 있는 데이터가 있을 경우 데이터 활용성을 극대화하고 학습 비용을 절감하는 장점이 있으나, 레이블 없는 데이터의 특성을 오해할 위험이 있어 주의가 필요하다.

비지도 방식의 이상탐지는 〈그림 5〉와 같이 레이블이 따로 존재하지 않는 데이터로부터 이상탐지 학습이 이루



〈그림5〉 지도,반지도,비지도 학습에 대한 훈련 데이터 예시[13]



〈그림 6〉 Auto Encoder 구조

어지는 방식이다. 일반적으로 AE(Autoencoder)를 기반으로 하는 모델이 비지도 이상탐지 모델로 제안되고 있는데, AE는 〈그림 6〉과 같이 입력 데이터를 낮은 차원으로 임베딩하고 다시 디코더로 임베딩된 값을 원래 데이터의 차원으로 재구성하여 입력 데이터와 재구성된 데이터 사이의 오차를 기반으로 학습이 이루어져 레이블이 따로 필요하지 않아 비지도 학습이 가능하다. AE는 이상 탐지에도 자주 사용되는 모델인데, AE 학습시 정상 데이터만을 이용해 학습을 진행한 뒤, 이상 데이터를 입력했을 때 임베딩 공간 또는 재구성 오차를 기반으로 이상을 탐지할 수 있다. 재구성 오차는 일반적으로 〈수식 3〉과 같이 평균 제곱 오차 (mean squared error, MSE)를 사용해 입력 값과 출력 값과의 오차를 제곱한 값의 평균 값을 통해 값의 차이를 줄이는 방향으로 학습이 진행된다.

AE 기반 비지도 학습을 통한 이상탐지 관련연구로, yan et al.[16]은 주변소음 하에서도 효과적인 비지도 학습이 가능한 HRCAE (hybrid robust convolutional autoencoder)를 제안했다. PCDF(Parallel Convolutional Distribution Fitting) 모듈을 구성해 다중 센서 정보를 효과적으로 통합하고, 이를 통해 정상 데이터의 분포를 더 잘 맞출 수 있도록 네트워크의 견고성을 높인 연구이다. Li et al.[17]은 AE와 TCN (temporal convolutional networks)을 기반으로 하는 VCEAD (variable cumulative error anomaly detection)를 제안했다. 이는 AE의 재구성 오류의 확률 분포를 사용하여

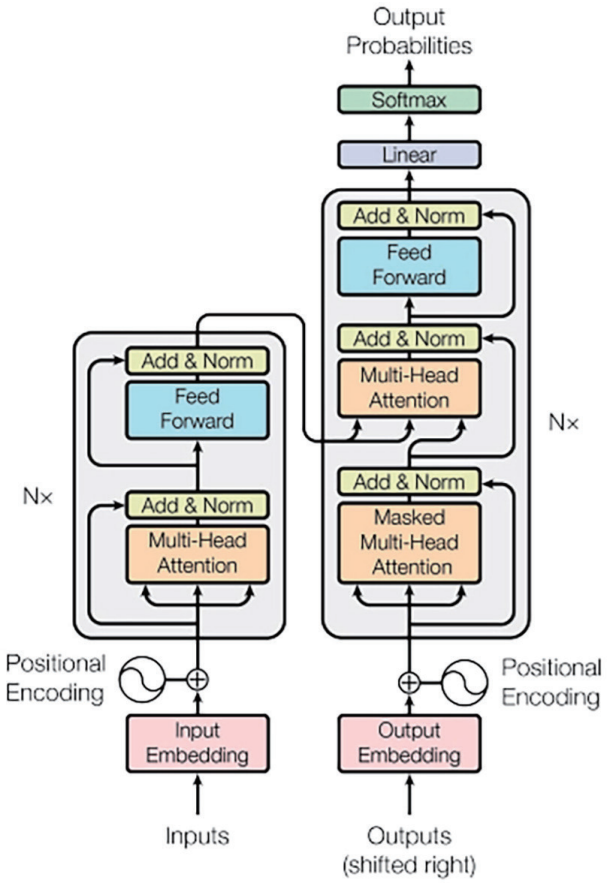
$$E = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y_i})^2$$

〈수식 3〉 Mean squared error

변수 누적 오류를 계산하는 알고리즘을 제안 했으며 이를 기반으로 임계값을 자동으로 설정하고 시계열 예측을 위해 TCN을 활용해 시계열 데이터의 동적인 오류 계산 및 임계값 비교를 통해 복잡하고 변동성이 큰 운영 환경의 영향을 고려한 이상 탐지 방법으로 조기 이상 탐지를 가능하게 했다.

Ⅳ. 고장진단

고장 진단은 결함이 발생한 구체적인 부품이나 과정을 확인하고, 문제가 발생한 원인을 판단하게 된다[18, 19]. 기존의 고장진단은 모델 기반으로 정확한 수학적 모델에 의존하기도 했다. 하지만 이는 정확한 수학적 모델에 의존하기 때문에, 시스템이나 구성요소의 실제 운영 조건과 다양한 고장 메커니즘을 완벽하게 모델링하는 것이 어렵다는 한계가 존재한다[20]. 이러한 한계를 보완하기 위



〈그림 7〉 Transformer 모델 구조[23]

해 데이터 기반 AI 모델을 도입하기 시작했다. AI를 활용한 고장 진단은 SVM (support vector machine), K-NN (K-nearest neighbor), DT (decision tree) 등 분류에 강점을 두는 메소드를 기반으로 한 다양한 방법론[21]이 제안되고 있으며, CNN (convolution neural networks), SAE (stacked autoencoder), DBN (deep belief network), DBM (deep boltzmann machine), RNN (recurrent neural network)과 같은 딥러닝 아키텍처를 기반으로 한 고장 진단 방법론[22]이 제안되고 있다.

특히 딥러닝 아키텍처를 기반으로 하는 모델은 attention[23] 매커니즘을 추가한 방법론이 많이 제안되고 있다. attention은 자연어 처리 분야에서 처음 제안된 방법으로 문장에서 중요한 부분을 동적으로 선택하고 강조하는 기능을 수행한다. 이는 고장 진단 분야에서 다중 센서 데이터와 같은 복잡한 시계열 데이터 속에서 고장과 관련된 패턴 또는 이상 상태를 attention score를 기반으로 효과적으로 식별할 수 있게 해주어 기존의 딥러닝 모델에 비해 더 정확한 진단을 가능하게 한다. 〈수식 4〉는 transformer의 scaled dot-product attention 수식으로 Q (query)는 특정 시점의 센서 데이터가 다른 시점 데이터와 어떤 연관성을 가지는지 찾을 때 사용되고, K (keys)는 연관성을 계산할 대상인 다른 모든 시점의 데이터를 의미한다. V (values)는 각 데이터 포인트에 할당된 값으로, Q와 모든 K의 내적을 수행해 유사도를 계산한다. K 벡터의 차원 수인 d_k로 내적 결과를 스케일링 한다. 이는 소프트맥스 함수의 기울기 소실 문제 방지를 위한 것이다. 이후 특정 시점의 센서 데이터가 다른 시점 데이터와 얼마나 관련성 있는지 확률 값으로 나타내기 위해 소프트맥스 함수를 이용해 나타낸 뒤, V와 곱한 뒤 attention 값을 구하게 된다.

Attention 매커니즘을 활용한 관련 연구로 Wang et al.[24]은 필터를 통해 복잡한 진동 신호에서 특징을 추출하는 CNN, attention 매커니즘을 이용해 중요한 채널 정보를 강조하는 CAM (channel attention mechanism)과

$$\mathrm{Attention}(Q, K, V) = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

〈수식 4〉 Scaled dot-product attention 수식

시계열 데이터에 특화된 GRU (gated recurrent unit)을 통합한 CAT-GRU 모델을 제안했다. CNN은 원시 1차원 진동 데이터에서 추상적 특징을 다중 합성곱 연산을 통해 인코딩하고, CAM은 CNN의 출력에서 채널별 결함의 중요성을 파악해 각 채널에 가중치를 부여해 결함과 관련된 특징을 강조했다. 마지막으로 GRU를 이용해 CNN과 CAM의 출력 특징에서 시간 관련 특징을 포착하여 고장 유형 분류를 하는 방법론을 제안했다.

최근 정확도 높은 모델들이 많이 제안되고 있지만 이러한 모델의 한계도 존재한다. 레이블이 지정된 데이터가 충분하고 도메인이 동일할 경우 고장 진단의 정확도가 높지만, 기존에 수집되지 못한 새로운 유형의 결함이 발생할 가능성이 있어, 결과적으로 고장진단을 어렵게 만들고 진단 모델의 성능을 저하시킬 수 있기 때문이다.

새로운 유형의 고장 식별을 위한 방법으로 새로운 결함 유형에 대한 탐지를 수행한 뒤, 고장을 분류하는 접근 방법이 존재한다. Ge et al.[25]은 다중 스케일 추출기를 통해 새로운 고장 유형의 데이터로부터 특징을 추출하고, 적응형 고장 판별기를 통해 새로운 고장 유형을 구분한 뒤, 새로운 고장 유형에 대해 클러스터링 하는 방법론을 제안했다. 이 연구에서 적응형 결함 판별기는 새로운 결함을 판별하는 핵심 요소로, 판별기는 두 개의 fully-connected 레이어로 구성되어 있으며, 첫 번째 레이어는 입력된 데이터가 알려진 결함 유형인지 판단을 하게 된다. 만약 알려진 결함 유형인 경우 다음 레이어로 넘어가게 되며, 기존에 학습된 모델을 통해 softmax 활성화 함수를 통해 어떤 결함 유형에 속해있는지 판단하게 된다. 새로운 결함 유형일 경우, 이 데이터는 다음 레이어로 넘어가지 않고 수집된 새로운 결함 데이터들에 대해 k-means clustering을 통해 클러스터링이 이루어진다. Zhao et al.[26]은 AE를 이용해 새로운 결함에 대한 분류를 수행한다. 이는 기존에 알려진 정상 또는 결함 데이터를 이용해 AE와 고장 분류기 모델의 학습이 진행된다. 이후 학습된 AE에 새로운 결함이 입력으로 들어갈 경우, AE를 통해 재구성된 결과가 입력 값과의 차이가 발생하게 되는데, 이때 설정된 임계값 보다 클 경우 새로운 유형 데이터로 판단하게 된다.

Brito et al.[27]은 비지도 학습을 통한 결함 분류와 XAI (explainable artificial intelligence)를 이용한 고장 진단 방법을 제안했다. 이 접근 방법은 지도 학습을 통해 새로운 유형의 결함을 분리하는 접근법과는 달리 비지도 학습을 통해 학습을 수행해, 새로운 결함 유형도 탐지하고 분류할 수 있는 방법이다. 이 연구에서는 CBLOF (Clustering Based Local Outlier Factor), LOF (Local Outlier Factor), IF (Isolation Forest), LODA (Lightweight on-line detector of anomalies) 등 비지도 학습을 통해 결함 유형을 분류하고, XAI 기법 중 SHAP(Shapley Additive Explanations)와 Local-DIFFI(Local Depth-based Feature Importance for the Isolation Forest)를 이용해 모델이 제안한 특성 중요도 순위를 바탕으로 진단에 필요한 설명력을 부여해, 고장의 원인을 파악하는데 도움을 주는 방법론을 제안했다. 여기서 SHAP는 게임 이론에서 유래한 설명 가능한 인공지능 기법 중 하나로, shapley value를 기반으로 하며, 머신 러닝 모델의 각 특성이 모델의 예측에 어떤 영향을 미쳤는지를 정량적으로 평가하는 방법이다. 〈수식 5〉는 특정 특성의 기여도를 계산하기 위해 모든 가능한 특성 조합을 고려하는 수식으로, Φ_i는 특정 변수의 shapley value를 나타내고, 이 값은 i가 모델의 예측에 미치는 기여도를 나타낸다. S는 관심 변수가 제외된 특성 집합을 나타내고, F는 전체 특성 집합을 의미한다. $|S|,|F|$는 각각 집합 F,S의 크기이다. $F(S)$는 집합 S에 속하는 특성들만 사용했을 때 모델 출력을 나타낸다. 이 수식에 따라 특성 i가 추가될 때와 추가되지 않았을 때의 모델 출력 차이를 계산하고, 모든 특성 조합에 대해 평균해 최종적인 shapley value를 얻게 된다. 이러한 방식으로 각 특성의 기여도를 계산해 모델의 예측이 어떻게 이루어지는지 이해할 수 있게 도와준다.

$$\phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!}(f(S \cup \{i\}) - f(S))$$

〈수식 5〉 SHAP 계산 수식

Ⅴ. 유효수명 예측

유효수명 예측 (remaining useful life, RUL)은 기계 또는 구성 요소가 고장이 발생하기 전까지 기능을 수행할 수 있는 남은 수명을 예측하는 것을 의미한다[28]. RUL 예측은 수집된 데이터로부터 직접 유효수명 간의 관계를 학습해 예측하는 것과 건강지표[29] (health indicator, HI)를 기반으로 예측하는 방법이 존재한다. 직접 학습 방식은 센서 데이터와 같은 운영 데이터를 사용하여, 머신러닝 알고리즘을 통해 기계의 고장 시점까지의 시간을 직접 예측하는 것으로, 이 방법은 기계의 실제 운영 조건을 반영하여 보다 정밀한 예측이 가능하다. 반면, HI 기반 예측은 센서 데이터로부터 추출된 건강지표를 이용해 지표가 변화하는 패턴을 분석함으로써 RUL을 예측한다. 이 방법은 데이터의 전처리와 특징 추출이 중요하며, HI의 정확한 정의와 추출이 예측의 정확성에 크게 영향을 미친다.

수집된 데이터에서 직접 RUL을 예측하는 연구로 SVR (support vector regression)을 활용하기도 한다. SVR은 연속적인 값을 예측하는 모델로, 주어진 데이터 내에서 가능한 가장 작은 오류를 가지면서, 데이터 포인트들 사이의 마진을 최대화하는 함수를 찾는 것을 목표로 학습된다. 〈수식 6〉은 SVR 손실함수 수식으로 손실 함수 값을 최소화하게 학습된다. 수식에서 w는 모델의 가중치 벡터이고, b는 편향을 나타낸다. ξ_i, ξ_i^*는 각각 데이터 포인트 i가 설정된 마진 외부에 있을 때의 오차를 측정하며, 이 슬랙(slack) 변수를 통해 예측 오차를 허용하며, ξ_i는 실제 값보다 예측 값이 작을 때와 ξ_i^*는 실제 값 보다 예측 값이 클 때 발생하며 최종적으로 가중치 벡터의 크기와 슬랙 변수의 합을 최소화 되는 것을 목표로 한다.

khelif et al.[30]은 래퍼(wrapper) 선택 기법을 사용해 관련성 높은 센서를 선택하고 특징을 추출한 뒤, SVR-RUL (support vector regression) 모델을 기반으로 수집된 데이터의 특징과 RUL 간의 관계를 학습하는 방법

$$\min_{\omega,b,\xi,\xi^*} \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\xi_i + \xi_i^*)$$

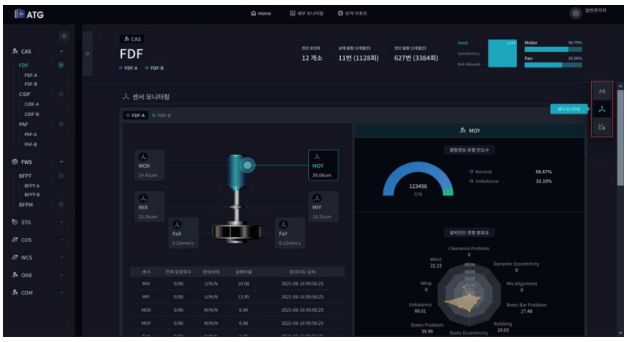
〈수식 6〉 SVR 손실 함수

론을 제안했다. 이 방법은 실패 임계값 설정 없이 SVR을 통해 센서 데이터로 부터 직접 RUL 예측이 가능하다. Xue et al.[31]은 RUL 예측을 위해 적응형 무향 칼만 필터(adaptive unscented kalman filter, AUKF)와 유전 알고리즘으로 최적화된 SVR을 결합한 방법론을 제안했다. 이 연구는 리튬 이온 배터리 RUL 예측에 관한 것으로, 배터리의 퇴화를 설명하기 위해 이중 지수 상태 공간 모델을 설정했다. 이는 배터리의 용량 변화를 설명하기 위해 설계된 수학적 표현으로, 수명 주기 동안 용량이 어떻게 감소하는지 시뮬레이션 할 수 있다. 이 모델을 설정한 이후, AUKF 알고리즘을 도입해 프로세스 노이즈 공분산과 관측 노이즈 공분산을 적응적으로 업데이트를 수행해 비선형 시스템의 상태를 추정한다. 이후 AUKF로부터 얻은 상태 추정치를 유전 알고리즘으로 최적화된 SVR의 입력으로 넣어 배터리 퇴화 패턴의 예측에 활용해 RUL을 예측하게 된다.

HI를 기반으로 RUL을 예측하는 연구로, Chen et al.[32]은 이차 함수 기반의 DCAE (deep convolutional autoencoder)를 사용해 원시 베어링 진동 신호로 부터 HI를 구축하는 방법을 제안했다. 이차 함수는 베어링의 성능 저하율이 시간에 따라 증가하는 경향을 모델링하기 위해 설계 되었으며, DCAE의 출력을 이용해 HI 변화 추세를 라벨링한다. 이후 GAU (gated attention unit) 모델을 사용해 RUL을 예측하는 방법을 제안했다. Ye et al.[33]은 MSCAE (Multi-Scale Convolutional Autoencoder)를 사용해 진동신호의 글로벌 및 로컬 특성을 포착해 성능 저하 특성을 추출해 HI를 구축하는 방법 제안했다. 이후 LSTM을 이용해 RUL을 예측하는 방법을 제안했다.

Ⅵ. 전망과 결론

기계 유지보수는 산업의 중요 요소로 자리 잡고 있으며, 유지보수 기술의 진화 또한 계속 진행되고 있다. 특히, 인공지능을 활용한 방법은 기계의 유지보수를 더욱 효율적이고 경제적으로 만들고 있다. 데이터 수집부터 유효수명 예측에 이르기까지, 각 단계에서 인공지능 기술을



〈그림 8〉 ATG사 진단 결과 예시[34]

〈그림 9〉 ATG사 설비 상태 현황 예시[34]

적용함으로써 예측 정확성을 높이고 유지보수 비용을 줄일 수 있게 되었다. 본 연구는 PHM의 각 단계에서 인공지능 기술이 어떻게 적용되고 있는지를 다루고 있으며, 이 과정에서 발생한 문제를 어떻게 접근해 문제를 해결하는지 접근 방향성을 볼 수 있었다.

현재 산업계에서는 AI를 활용한 다양한 PHM 시스템을 활용 중에 있으며 이를 통해 현재 기계 시설물의 상태와 이상 유무를 실시간으로 확인하면서 운영에 드는 비용을 줄이고 효율성을 증대시키고 있다. 또한 〈그림 9〉와 같이 수집된 데이터 기반 AI를 통해 시설물 결함 발생 또는 경고와 같이 관리자가 종합적인 관리가 수월하게 할 수 있는 다양한 제품들이 제안되고 있다.

하지만 기계의 구체적인 결함에 대한 원인 또는 영향을 주는 영역에 대한 설명이 가능한 모델에 대한 연구는 아직 부족한 것으로 보인다. 이 분야에 대한 연구가 활발히 진행된다면, 현장 관리자들은 기계 부품의 상태를 더 정확히 파악하여 신속한 대처와 효율적인 부품 조달 계획을 수립할 수 있을 것으로 보인다. 이를 통해 시간과 비용을 절약하고, 원인 분석을 바탕으로 시설물을 개선하여 보다

효율적인 유지보수 및 관리가 가능해질 것으로 기대된다.

참 고 문 헌

[1] Soualhi, Abdenour, et al. "PHM SURVEY: Implementation of prognostic methods for monitoring industrial systems." *Energies* 15.19 (2022): 6909.

[2] Homepage of Korea society for prognosis and health management. www.phm.or.kr. Accessed 18 Aug 2023

[3] Uhlemann, Thomas H-J., Christian Lehmann, and Rolf Steinhilper. "The digital twin: Realizing the cyber-physical production system for industry 4.0." *Procedia Cirp* 61 (2017): 335-340.

[4] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems* 27 (2014).

[5] Shao, Siyu, Pu Wang, and Ruqiang Yan. "Generative adversarial networks for data augmentation in machine fault diagnosis." *Computers in Industry* 106 (2019): 85-93.

[6] Liu, Wenbiao, et al. "Data augmentation method for fault diagnosis of mechanical equipment based on improved Wasserstein GAN." *2020 11th International Conference on Prognostics and System Health Management (PHM-2020 Jinan)*. IEEE, 2020.

[7] Atoui, Issam, et al. "Fault detection and diagnosis in rotating machinery by vibration monitoring using FFT and Wavelet techniques." *2013 8th international workshop on systems, signal processing and their applications (WoSSPA)*. IEEE, 2013.

[8] Cofre-Martel, Sergio, Enrique Lopez Droguett, and Mohammad Modarres. "Big machinery data preprocessing methodology for data-driven models in prognostics and health management." *Sensors* 21.20 (2021): 6841.

[9] Pandarakone, Shrinathan Esakimuthu, et al. "Deep neural network based bearing fault diagnosis of induction motor using fast Fourier transform analysis." *2018 IEEE energy conversion congress and exposition (ECCE)*. IEEE, 2018.

[10] 장계봉, and 조성배. "회전 기계 고장 진단을 위한 적대적 순환 일관성 유지 학습기반 교차 도메인 적응 방법." *정보과학회논문지* 49.7 (2022): 530-536.

[11] Ozdagli, Ali, and Xenofon Koutsoukos. "Domain adaptation for structural health monitoring." *Annual Conference of the PHM Society*. Vol. 12. No. 1. 2020.

[12] Nassif, Ali Bou, et al. "Machine learning for anomaly detection: A systematic review."*Ieee Access*9 (2021): 78658-78700.

[13] https://www.kdnuggets.com/2019/11/tips-class-imbalance-missing-labels.html

[14] Denkena, B., et al. "Data-based ensemble approach for semi-supervised anomaly detection in machine tool condition monitoring." *CIRP Journal of Manufacturing Science and Technology* 35 (2021): 795-802.

[15] Lv, Longlong, et al. "Research on Anomaly Detection Method of Non-Standard Machinery Production Process Based on Semi-Supervised Learning." *Journal of Networking and Network Applications* 3.2 (2023): 81-88.

[16] Yan, Shen, et al. "Hybrid robust convolutional autoencoder for unsupervised anomaly detection of machine tools under noises." *Robotics and Computer-Integrated Manufacturing* 79 (2023): 102441.

[17] Li, Zhiyuan, et al. "Unsupervised machine anomaly detection using autoencoder and temporal convolutional network." *IEEE Transactions on Instrumentation and Measurement* 71 (2022): 1-13.

[18] Mian, Zhibao, et al. "A literature review of fault diagnosis based on ensemble learning." *Engineering Applications of Artificial Intelligence* 127 (2024): 107357.

[19] Sahu, Atma Ram, Sanjay Kumar Palei, and Aishwarya Mishra. "Data-driven fault diagnosis approaches for industrial equipment: A review." *Expert Systems* 41.2 (2024): e13360.

[20] Li, Guoqiang, et al. "Self-supervised learning for intelligent fault diagnosis of rotating machinery with limited labeled data." *Applied Acoustics* 191 (2022): 108663.

[21] Liu, Ruonan, et al. "Artificial intelligence for fault diagnosis of rotating machinery: A review." *Mechanical Systems and Signal Processing* 108 (2018): 33-47.

[22] Cen, Jian, et al. "A review of data-driven machinery fault diagnosis using machine learning algorithms." *Journal of Vibration Engineering & Technologies* 10.7 (2022): 2481-2507.

[23] Vaswani, Ashish, et al. "Attention is all you need."*Advances in neural information processing systems*30 (2017).

[24] Wang, Zilin, et al. "Rolling Bearing Fault Diagnosis Using CNN-based Attention Modules and Gated Recurrent Unit." *2020 Global Reliability and Prognostics and Health Management (PHM-Shanghai)*. IEEE, 2020.

[25] Ge, Yang, Fusheng Zhang, and Yong Ren. "Adaptive fault diagnosis method for rotating machinery with unknown faults under multiple working conditions." *Journal of Manufacturing Systems* 63 (2022): 177-184.

[26] Zhao, Yuanyuan, et al. "Novelty detection and fault diagnosis method for bearing faults based on the hybrid deep autoencoder network."*Electronics*12.13 (2023): 2826.

[27] Brito, Lucas C., et al. "An explainable artificial intelligence approach for unsupervised fault detection and diagnosis in rotating machinery." *Mechanical Systems and Signal Processing* 163 (2022): 108105.

[28] Wang, Youdao, Yifan Zhao, and Sri Addepalli. "Remaining useful life prediction using deep learning approaches: A review." *Procedia manufacturing* 49 (2020): 81-88.

[29] Lei, Yaguo, et al. "Machinery health prognostics: A systematic review from data acquisition to RUL prediction." *Mechanical systems and signal processing* 104 (2018): 799-834.

[30] Khelif, Racha, et al. "Direct remaining useful life estimation based on support vector regression." *IEEE Transactions on industrial electronics* 64.3 (2016): 2276-2285.

[31] Xue, Zhiwei, et al. "Remaining useful life prediction of lithium-ion batteries with adaptive unscented kalman filter and optimized support vector regression."*Neurocomputing*376 (2020): 95-102.

[32] Chen, Dingliang, et al. "Health indicator construction by quadratic function-based deep convolutional auto-encoder and its application into bearing RUL prediction." *ISA transactions* 114 (2021): 44-56.

[33] Ye, Zijian, et al. "Rolling bearing health indicator extraction and RUL prediction based on multi-scale convolutional autoencoder." *Applied Sciences* 12.11 (2022): 5747.

[34] https://www.atg.co.kr/

김관필

• 현 중앙대학교 컴퓨터공학과 석사과정

〈관심 분야〉
이상탐지, 유효수명 예측, 건강지표, 그래프 임베딩

김상범

• 1993년 동아대학교 전자공학과 학사
• 1995년 동아대학교 전자공학과 석사
• 1999년 동아대학교 전자공학과 박사
• 현 한국폴리텍대학 스마트소프트웨어과 교수

〈관심 분야〉
IOT, 비전처리, 빅데이터, 인공지능

THE INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS

논문지 논문목차

전자공학회 논문지 제 61권 5호 발행

통신 분야

[통신]

반도체 분야

[반도체 소자 및 재료]

인공지능 신호처리 분야

[영상 신호처리]

[영상 이해]

시스템 및 제어 분야

[제어계측]

[국방정보 및 제어]

국내외학술행사안내

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

〉〉2024년 6월

일 자	학술대회명	개최장소	홈페이지/연락처
06.02. - 06.05.	2024 IEEE Intelligent Vehicle Symposium (IV)	Jeju Island, Korea (South)	https://ieee-iv.org/
06.02. - 06.05.	2024 IEEE Electrical Insulation Conference (EIC)	Minneapolis, Minnesota, USA	https://ieee-eic.org/
06.02. - 06.05.	2024 IEEE 21st Biennial Conference on Electromagnetic Field Computation (CEFC)	Jeju, Korea (South)	http://cefc2024.com/index
06.02. - 06.06.	2024 36th International Symposium on Power Semiconductor Devices and ICs (ISPSD)	Bremen, Germany	https://www.ispsd2024.com/
06.03. - 06.05.	2024 11th International Workshop on Metrology for AeroSpace (MetroAeroSpace)	Lublin, Poland	https://www.metroaerospace.org/
06.03. - 06.05.	2024 Advances in Science and Engineering Technology International Conferences (ASET)	Abu Dhabi, United Arab Emirates	ASET 2024 - Higher Colleges of Technology (hct.ac.ae)
06.03. - 06.06.	2024 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)	Antwerp, Belgium	https://www.eucnc.eu/
06.03. - 06.06.	2024 1st International Conference on Smart Energy Systems and Artificial Intelligence (SESAI)	Mauritius	https://www.sesai2024.com/
06.03. - 06.06.	2024 IEEE 12th International Conference on Healthcare Informatics (ICHI)	Orlando, Florida, USA	https://ieeeichi2024.github.io/
06.03. - 06.06.	2024 IEEE International Interconnect Technology Conference (IITC)	San Jose, California, USA	https://iitc-conference.org/
06.03. - 06.06.	2024 IEEE 16th Workshop on Low Temperature Electronics (WOLTE)	Sardinia, Italy	https://wolte16.org/
06.03. - 06.06.	2024 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)	St. Petersburg, Russia	http://media-publisher.ru/en/about-weconf-2024/
06.03. - 06.06.	2024 11th International Conference on Electrical, Electronic and Computing Engineering (IcETRAN)	Nis, Serbia	https://www.etran.rs/2024/en/home-english/
06.04. - 06.06.	2024 IEEE International Conference on RFID (RFID)	Cambridge, Massachusetts, USA	https://2024.ieee-rfid.org/
06.04. - 06.07.	2024 International Conference on Unmanned Aircraft Systems (ICUAS)	Chania - Crete, Greece	https://uasconferences.com/2024_icuas/
06.04. - 06.07.	2024 IEEE 25th International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)	Perth, Australia	https://www.cse.unsw.edu.au/~wowmom/2024/
06.04. - 06.07.	2024 6th Global Power, Energy and Communication Conference (GPECOM)	Budapest, Hungary	https://gpecom.org/2024/
06.04. - 06.08.	2024 IEEE 21st International Conference on Software Architecture (ICSA)	Hyderabad, India	https://icsa-conferences.org/series/
06.05. - 06.06.	2024 1st International Conference on Production Technologies and Systems for E-Mobility (EPTS)	Bamberg, Germany	https://www.faps.fau.de/e_pts/
06.05. - 06.07.	2024 OPJU International Technology Conference (OTCON) on Smart Computing for Innovation and Advancement in Industry 4.0	Raigarh, India	https://opjuieeeotc.in/

일 자	학술대회명	개최장소	홈페이지/연락처
06.05. - 06.07.	2024 3rd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)	Event Format: Virtual	https://icaaic.com/
06.05. - 06.07.	2024 IEEE Gaming, Entertainment, and Media Conference (GEM)	Turin, Italy	https://www.ieee-gem2024.org/
06.06. - 06.07.	2024 9th International Conference on Integrated Circuits, Design, and Verification (ICDV)	Hanoi, Vietnam	https://icdv-conf.org/
06.06. - 06.07.	2024 International Conference on Smart Computing, IoT and Machine Learning (SIML)	Surakarta, Indonesia	https://siml.ums.ac.id/2024/
06.07. - 06.08.	2024 International Conference on Innovations and Challenges in Emerging Technologies (ICICET)	Nagpur, India	https://ghrce.raisoni.net/icicet-24/
06.07. - 06.09.	2024 6th World Symposium on Artificial Intelligence (WSAI)	Guangzhou, China	https://www.wsai.org/
06.09. - 06.12.	2024 19th Conference on PhD Research in Microelectronics and Electronics (PRIME)	Larnaca, Cyprus	https://cyprusconferences.org/prime2024/
06.09. - 06.13.	2024 IEEE IAS Pulp and Paper Industry Conference (PPIC)	Charleston, West Virginia, USA	https://www.pulppaper.org/
06.09. - 06.13.	2024 IEEE International Conference on Communications Workshops (ICC Workshops)	Denver, Colorado, USA	https://icc2024.ieee-icc.org/
06.09. - 06.13.	ICC 2024 - IEEE International Conference on Communications	Denver, Colorado, USA	https://icc2024.ieee-icc.org/
06.09. - 06.14.	2024 IEEE Magnetics Society Summer School (MSSS)	Taipei, Taiwan	https://sites.google.com/view/ieee-magsoc-summer-school-2024/
06.09. - 06.14.	2024 IEEE 52nd Photovoltaic Specialist Conference (PVSC)	Seattle, Washington, USA	https://ieee-pvsc.org/PVSC52/
06.10. - 06.12.	2024 IEEE 31st Symposium on Computer Arithmetic (ARITH)	Malaga, Spain	https://www.ac.uma.es/arith2024/index.html
06.10. - 06.12.	2024 20th International Conference on the European Energy Market (EEM)	Istanbul, Turkiye	https://eem24.khas.edu.tr/
06.10. - 06.13.	2024 IET 13th International Conference on Power Electronics, Machines and Drives (PEMD)	Nottingham, United Kingdom	https://pemd.theiet.org/?utm_source=ieee&utm_medium=referral&utm_campaign=pemd
06.11. - 06.13.	2024 IMAPS Nordic Conference on Microelectronics Packaging (NordPac)	Tampere, Finland	https://nordic.imapseurope.org/nordpac/
06.11. - 06.13.	2024 4th Interdisciplinary Conference on Electrics and Computer (INTCEC)	Chicago, Illinois, USA	https://www.intcec.org/
06.11. - 06.13.	2024 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)	Montreal, Quebec, Canada	https://imbioc-ieee.ece.mcgill.ca/
06.11. - 06.14.	2024 32nd Mediterranean Conference on Control and Automation (MED)	Chania - Crete, Greece	https://www.med-control.org/med2024/
06.11. - 06.14.	2024 13th Mediterranean Conference on Embedded Computing (MECO)	Budva, Montenegro	https://mecoconference.me/meco2024/
06.12. - 06.14.	2024 Picture Coding Symposium (PCS)	Taichung, Taiwan	https://2024.picturecodingsymposium.org/
06.12. - 06.14.	2024 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)	Chania, Greece	https://www.metrolivenv.org/
06.12. - 06.14.	2024 XVIII National Meeting on Optics and the IX Andean and Caribbean Conference on Optics and its Applications (ENO-CANCOA)	Cartagena, Colombia	https://www.utb.edu.co/eno-cancoa-2024/
06.12. - 06.15.	2024 23rd International Symposium on Electrical Apparatus and Technologies (SIELA)	Bourgas, Bulgaria	https://siela.tu-sofia.bg/
06.14. - 06.15.	2024 International Conference on Energy and Electrical Engineering (EEE)	Nanchang, China	https://www.iceeeconf.com/
06.14. - 06.16.	2024 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)	Xi'an, China	https://civemsa2024.ieee-ims.org/
06.14. - 06.16.	2024 6th International Conference on Computer Communication and the Internet (ICCCI)	Tokyo, Japan	https://www.iccci.org/

일 자	학술대회명	개최장소	홈페이지/연락처
06.16. - 06.18.	2024 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)	Washington, District of Columbia, USA	https://rfic-ieee.org/
06.16. - 06.19.	2024 22nd IEEE Interregional NEWCAS Conference (NEWCAS)	Sherbrooke, Quebec, Canada	https://newcas.org/
06.16. - 06.20.	2024 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)	Honolulu, Hawaii, USA	https://www.vlsisymposium.org/
06.16. - 06.20.	2024 IEEE International Conference on Plasma Science (ICOPS)	Beijing, China	https://www.icops2024.com/
06.16. - 06.21.	2024 IEEE/MTT-S International Microwave Symposium - IMS 2024	Washington, District of Columbia, USA	https://ims-ieee.org/
06.16. - 06.22.	2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	Seattle, Washington, USA	https://cvpr.thecvf.com/
06.17. - 06.19	2024 Silicon Valley Cybersecurity Conference (SVCC)	Seoul, Korea (South)	https://www.svcc-svcsi.org/
06.17. - 06.19	2024 European Conference on Integrated Optics (ECIO)	Aachen, Germany	https://www.ecio-conference.org/
06.17. - 06.19	2024 Italian Conference on Optics and Photonics (ICOP)	Firenze, Italy	https://www.icop2024.it/
06.17. - 06.19	2024 IEEE International Conference on Prognostics and Health Management (ICPHM)	Spokane, Washington, USA	https://phmconf.org/
06.17. - 06.20.	2024 International Conference on Intelligent Environments (IE)	Ljubljana, Slovenia	https://ie2024.ijs.si/
06.18. - 06.20.	2024 16th International Conference on Quality of Multimedia Experience (QoMEX)	Karlshamn, Sweden	https://qomex2024.itec.aau.at/
06.18. - 06.21.	2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE)	Ulsan, Korea (South)	https://www.ieee-ies.org/conferences
06.18. - 06.22.	2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)	Kamand, India	https://15icccnt.com/
06.19. - 06.21.	2024 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)	Wilmington, Delaware, USA	https://conferences.computer.org/chase2024/
06.19. - 06.21.	2024 IEEE/ACM 32nd International Symposium on Quality of Service (IWQoS)	Guangzhou, China	https://iwqos2024.ieee-iwqos.org/
06.19. - 06.21.	2024 5th Technology Innovation Management and Engineering Science International Conference (TIMES-iCON)	Bangkok, Thailand	https://timesicon2024.atdi.or.th/web/
06.19. - 06.21.	2024 IEEE Transportation Electrification Conference and Expo (ITEC)	Chicago, Illinois, USA	https://itec-conf.com/
06.19. - 06.21.	2024 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM)	Napoli, Italy	https://www.speedam.org/
06.19. - 06.21.	2024 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)	Toronto, Ontario, Canada	https://bmsb2024.com/
06.19. - 06.22.	2024 21st International Joint Conference on Computer Science and Software Engineering (JCSSE)	Phuket, Thailand	https://jcsse2024.computing.psu.ac.th/
06.20.	2024 V International Conference on Neural Networks and Neurotechnologies (NeuroNT)	Saint Petersburg, Russia	https://ieee.spb.etu.ru/en/home/konferencii/2024/v-international-conference-on-neural-networks-and-neurotechnologies-neuront
06.20. - 06.21.	2024 IEEE International Symposium on Robotic and Sensors Environments (ROSE)	Chemnitz, Germany	https://rose2024.ieee-ims.org/
06.20. - 06.23.	2024 International Conference on Meta Computing (ICMC)	Qingdao, China	http://meta-computing.info/index.html
06.21	2024 103rd ARFTG Microwave Measurement Conference (ARFTG)	Washington, District of Columbia, USA	https://arftg.org/
06.21. - 06.22.	2024 IEEE 3rd International Conference on Electrical Power and Energy Systems (ICEPES)	Bhopal, India	https://www.icepes2024-manit.in/
06.21. - 06.22.	2024 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)	KANNUR, India	https://gcek.ac.in/APCI2024/

일 자	학술대회명	개최장소	홈페이지/연락처
06.21. - 06.23.	2024 4th International Conference on Intelligent Technologies (CONIT)	Bangalore, India	http://inconf.in/
06.21. - 06.23.	2024 12th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC)	Chongqing, China	https://www.icwoc.org/
06.21. - 06.23.	2024 IEEE 4th International Conference on Software Engineering and Artificial Intelligence (SEAI)	Xiamen, China	https://www.seai.org/
06.21. - 06.23.	2024 IEEE Students Conference on Engineering and Systems (SCES)	Prayagraj, India	http://www.mnnit.ac.in/sces2024/
06.22	2024 4th International Conference on Code Quality (ICCQ)	Innopolis, Russia	https://www.iccq.ru/2024.html
06.23. - 06.26.	2024 IEEE 15th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)	Luxembourg, Luxembourg	https://www.pedg2024.lu/
06.23. - 06.27.	2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)	Brisbane, Australia	https://dsn2024uq.github.io/
06.23. - 06.27.	2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)	Brisbane, Australia	https://dsn2024uq.github.io/
06.23. - 06.27.	2024 54th Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S)	Brisbane, Australia	https://dsn2024uq.github.io/
06.23. - 06.27.	2024 61st ACM/IEEE Design Automation Conference (DAC)	San Francisco, California, USA	https://www.dac.com/
06.24. - 06.25.	2024 IEEE 32nd International Requirements Engineering Conference Workshops (REW)	Reykjavik, Iceland	https://conf.researchr.org/home/RE-2024
06.24. - 06.26.	2024 IEEE 18th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)	Gdynia, Poland	https://cpe-powereng2024.umg.edu.pl/
06.24. - 06.26.	2024 6th International Conference on Reconfigurable Mechanisms and Robots (ReMAR)	Chicago, Indiana, USA	https://web.cvent.com/event/bd1bd184-001d-4395-b097-51389324c186/summary
06.24. - 06.26.	2024 Device Research Conference (DRC)	College Park, Maryland, USA	https://2024.deviceresearchconference.org/
06.24. - 06.27.	2024 International Conference on Space Robotics (iSpaRo)	Luxembourg, Luxembourg	https://www.spacer.lu/isparo_24
06.24. - 06.27.	2024 21st International Conference on Ubiquitous Robots (UR)	New York, New York, USA	https://2024.ubiquitousrobots.org/
06.24. - 06.27.	2024 18th International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)	Auckland, New Zealand	https://www.pmaps2024.com/
06.24. - 06.27.	2024 25th IEEE International Conference on Mobile Data Management (MDM)	Brussels, Belgium	https://mdm2024.github.io/
06.24. - 06.27.	2024 14th European Conference on Magnetic Sensors and Actuators (EMSA)	Kosice, Slovakia	https://indico.science.upjs.sk/event/198/
06.24. - 06.27.	2024 IEEE Workshop on Control and Modeling for Power Electronics (COMPEL)	Lahore, Pakistan	https://ieee-compel.org/index.php
06.24. - 06.27.	2024 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)	Tbilisi, Georgia	https://blackseacom2024.ieee-blackseacom.org/
06.24. - 06.27.	2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)	Singapore, Singapore	https://events.vtsociety.org/vtc2024-spring/
06.24. - 06.28.	2024 IEEE 32nd International Requirements Engineering Conference (RE)	Reykjavik, Iceland	https://conf.researchr.org/home/RE-2024
06.24. - 06.28.	2024 IEEE 10th International Conference on Network Softwarization (NetSoft)	Saint Louis, Missouri, USA	https://netsoft2024.ieee-netsoft.org/
06.24. - 06.28.	2024 IEEE International Conference on Real-time Computing and Robotics (RCAR)	Alesund, Norway	http://www.ieee-rcar.org/
06.24. - 06.28.	2024 Joint Telematics Group/IEEE Information Theory Society Summer School in Information Theory, Signal Processing, Telecommunication, and Networking (JTG/ITSoc Summer School)	Hyderabad, India	https://ee.iith.ac.in/JTG2024/
06.24. - 06.28.	2024 IEEE International Conference on Engineering, Technology, and Innovation (ICE/ITMC)	Funchal, Portugal	https://mdtweek.digit-madeira.pt/ice/

일 자	학술대회명	개최장소	홈페이지/연락처
06.25. - 06.26.	2024 International e-Engineering Education Services Conference (e-Engineering)	Mostaganem, Algeria	https://e-engineering.org/
06.25. - 06.27.	2024 IEEE Conference on Artificial Intelligence (CAI)	Singapore, Singapore	https://ieeecai.org/2024/
06.25. - 06.27.	2024 International Conference on Localization and GNSS (ICL-GNSS)	Antwerp, Belgium	https://events.tuni.fi/icl-gnss2024/
06.25. - 06.27.	2024 IEEE 22nd Mediterranean Electrotechnical Conference (MELECON)	Porto, Portugal	https://2024.ieee-melecon.org/
06.25. - 06.28.	2024 9th International Conference on Smart and Sustainable Technologies (SpliTech)	Bol and Split, Croatia	https://2024.splitech.org/Home
06.26. - 06.28.	2024 XVI Congreso de Tecnología, Aprendizaje y Enseñanza de la Electrónica (XVI Technologies Applied to Electronics Teaching Conference) (TAEE)	Malaga, Spain	https://congresotaee.es/en/en-home/
06.26. - 06.28.	2024 IEEE International Workshop on Metrology for Automotive (MetroAutomotive)	Bologna, Italy	https://www.metroautomotive.org/
06.26. - 06.28.	2024 IEEE Sixth International Conference on Electrical, Computer and Communication Technologies (ICECCT)	Durg, India	https://icecct.com/
06.26. - 06.28.	2024 International Conference on System Science and Engineering (ICSSE)	Hsinchu, Taiwan	https://icsse2024.web.nycu.edu.tw/
06.26. - 06.28.	2024 13th International Conference on Modern Circuits and Systems Technologies (MOCAST)	Sofia, Bulgaria	http://mocast.eu/
06.26. - 06.28.	2024 IEEE International Symposium on Medical Measurements and Applications (MeMeA)	Eindhoven, Netherlands	https://memea2024.ieee-ims.org/
06.26. - 06.28.	2024 IEEE 37th International Symposium on Computer-Based Medical Systems (CBMS)	Guadalajara, Mexico	https://2024.cbms-conference.org/
06.26. - 06.29.	2024 IEEE Symposium on Computers and Communications (ISCC)	Paris, France	https://2023.ieee-iscc.org/
06.27. - 06.28.	2024 International Conference on Renewable Energies and Smart Technologies (REST)	Prishtina, Kosovo (UNMIK)	https://ic-rest.com/
06.27. - 06.28.	2024 IEEE Cloud Summit	Washington, District of Columbia, USA	https://www.ieeecloudsummit.org/
06.27. - 06.28.	2024 16th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)	Iasi, Romania	https://ecai.ro/
06.27. - 06.29.	2024 9th International Conference on Energy Efficiency and Agricultural Engineering (EE&AE)	Ruse, Bulgaria	https://eeae-conf.uni-ruse.bg/
06.28. - 06.29.	2024 IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)	Bangalore, India	https://iteics.in/index.php
06.28. - 06.29.	2024 IEEE LLM Aided Design Workshop (LAD)	San Jose, California, USA	https://www.islad.org/
06.28. - 06.30.	2024 IEEE 2nd International Conference on Image Processing and Computer Applications (ICIPCA)	Shenyang, China	http://www.icipca.org/
06.28. - 06.30.	2024 4th International Conference on Machine Learning and Intelligent Systems Engineering (MLISE)	Zhuhai, China	http://mlise.org/
06.28. - 06.30.	2024 9th International Conference on Information Science, Computer Technology and Transportation (ISCTT)	Mianyang, China	http://www.isctt.net/
06.28. - 07.01.	2024 7th International Conference on Electrical Engineering and Green Energy (CEEGE)	Los Angeles, California, USA	https://www.ceege.org/index.html
06.29.	2024 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)	Shah Alam, Malaysia	https://sites.google.com/view/i2cacis/home
06.29. - 07.02	2024 IEEE International Conference on Smart Computing (SMARTCOMP)	Osaka, Japan	https://smartcomp.w.waseda.jp/
06.29. - 07.03.	2024 ACM/IEEE 51st Annual International Symposium on Computer Architecture (ISCA)	Buenos Aires, Argentina	https://iscaconf.org/isca2024/
06.30. - 07.03	2024 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)	Tampere, Finland	https://2024.ieee-fleps.org/

일 자	학술대회명	개최장소	홈페이지/연락처
06.30. - 07.04	2024 IEEE Opto-Electronics and Communications Conference (OECC)	Melbourne, Australia	https://oecc2024.com/
06.30. - 07.04.	2024 IEEE 5th International Conference on Dielectrics (ICD)	Toulouse, France	https://ieee-icd.org/
06.30. - 07.05	2024 International Joint Conference on Neural Networks (IJCNN)	Yokohama, Japan	https://2024.ieeewcci.org/
06.30. - 07.05	2024 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)	Yokohama, Japan	https://2024.ieeewcci.org/
06.30. - 07.05	2024 IEEE Congress on Evolutionary Computation (CEC)	Yokohama, Japan	https://2024.ieeewcci.org/
06.30. - 07.05.	2024 IEEE World Congress on Computational Intelligence (WCCI)	Yokohama, Japan	https://2024.ieeewcci.org/
〉〉2024년 7월			
07.01. - 07.04.	2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)	Osaka, Japan	https://ieeecompsac.computer.org/2024/
07.01. - 07.03.	2024 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)	Vyborg, Russia	http://media-publisher.ru/en/about-synchroinfo-2024/
07.01. - 07.05.	2024 IEEE European School of Information Theory (ESIT)	Eindhoven, Netherlands	https://www.sps.tue.nl/ictlab/esit2024/
07.01. - 07.04.	2024 IEEE International Conference on Advanced Learning Technologies (ICALT)	Nicosia, North Cyprus, Cyprus	https://tc.computer.org/tclt/icalt-2024/
07.01. - 07.04.	2024 International Conference on Signal Processing and Communications (SPCOM)	Bangalore, India	https://ece.iisc.ac.in/~spcom/2024/
07.01. - 07.04.	2024 25th International Microwave and Radar Conference (MIKON)	Wroclaw, Poland	https://mrw2024.org/
07.01. - 07.03.	2024 59th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)	Sozopol, Bulgaria	https://icestconf.org/
07.01. - 07.03.	2024 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)	Knoxville, Tennessee, USA	https://www.ieee-isvlsi.org/ISVLSI_2024_Website/index.html
07.02 - 07.04.	2024 13th International Workshop on Robot Motion and Control (RoMoCo)	Pozna□, Poland	https://romoco.put.poznan.pl/
07.02 - 07.05.	2024 Fifteenth International Conference on Ubiquitous and Future Networks (ICUFN)	Budapest, Hungary	https://icufn.org/
07.02 - 07.04.	2024 International Radar Symposium (IRS)	Wroclaw, Poland	https://mrw2024.org/irs-2024-2/
07.02 - 07.05.	2024 31st International Workshop on Active-Matrix Flatpanel Displays and Devices (AM-FPD)	Kyoto, Japan	http://www.amfpd.jp/
07.02 - 07.06.	2024 9th International Youth Conference on Energy (IYCE)	Colmar, France	https://www.iyce-conf.org/
07.02 - 07.05.	2024 IEEE International Symposium on Measurements & Networking (M&N)	Rome, Italy	https://www.ing.uniroma1.it/en
07.02 - 07.03.	2024 2nd International Conference on Fundamental, Applied Sciences and Technology (ICoFAST)	Mukalla, Yemen	https://hu.edu.ye/icofast/
07.03. - 07.05.	2024 IEEE 30th International Symposium on On-Line Testing and Robust System Design (IOLTS)	Rennes, France	https://orion.polito.it/iolts/
07.04. - 07.06.	2024 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)	BALI, Indonesia	https://iaict.org/
07.05. - 07.07.	2024 International Conference on Ubiquitous Communication (Ucom)	Xi'an, China	https://www.ucom.org.cn/index.html
07.05.	2024 8th International Young Engineers Forum on Electrical and Computer Engineering (YEF-ECE)	Caparica / Lisbon, Portugal	https://yef-ece.deec.fct.unl.pt/
07.06. - 07.08.	2024 International Conference on Advanced Robotics and Mechatronics (ICARM)	Tokyo, Japan	http://www.ieee-arm.org/

일 자	학술대회명	개최장소	홈페이지/연락처
07.06. - 07.07.	2024 IEEE Symposium on Industrial Electronics & Applications (ISIEA)	Kuala Lumpur, Malaysia	https://www.isiea.asia/
07.07 - 07.12.	IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium	Athens, Greece	https://www.2024.ieeeigarss.org/
07.07 - 07.13.	2024 IEEE World Congress on Services (SERVICES)	Shenzhen, China	https://conferences.computer.org/services/2024/
07.07.	2024 IEEE International Symposium on Information Theory Workshops (ISIT-W)	Athens, Greece	https://2024.ieee-isit.org/home
07.07 - 07.13.	2024 IEEE 17th International Conference on Cloud Computing (CLOUD)	Shenzhen, China	https://cloud.conferences.computer.org/2024/
07.07 - 07.13.	2024 IEEE International Conference on Edge Computing and Communications (EDGE)	Shenzhen, China	https://edge.conferences.computer.org/2024/
07.07 - 07.13.	2024 IEEE International Conference on Digital Health (ICDH)	Shenzhen, China	https://icdh.conferences.computer.org/2024/
07.07 - 07.13.	2024 IEEE International Conference on Web Services (ICWS)	Shenzhen, China	https://icws.conferences.computer.org/2024/
07.07 - 07.13.	2024 IEEE International Conference on Quantum Software (QSW)	Shenzhen, China	https://qsw.conferences.computer.org/2024/
07.07 - 07.13.	2024 IEEE International Conference on Software Services Engineering (SSE)	Shenzhen, China	https://sse.conferences.computer.org/2024/
07.08. - 07.12.	2024 IEEE International Symposium on Information Theory (ISIT)	Athens, Greece	https://2024.ieee-isit.org/home
07.08. - 07.11.	2024 27th International Conference on Information Fusion (FUSION)	Venice, Italy	https://fusion2024.org/
07.08. - 07.12.	2024 IEEE 9th European Symposium on Security and Privacy (EuroS&P)	Vienna, Austria	https://eurosp2024.ieee-security.org/
07.08. - 07.11.	2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)	Madrid, Spain	https://meditcom2024.ieee-meditcom.org/
07.08. - 07.12.	2024 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)	Vienna, Austria	https://eurosp2024.ieee-security.org/workshops.html
07.08. - 07.12.	2024 Conference on Precision Electromagnetic Measurements (CPEM)	Denver, Colorado, USA	https://ncsli.org/?
07.08. - 07.11.	2024 IEEE 13th Sensor Array and Multichannel Signal Processing Workshop (SAM)	Corvallis, Oregon, USA	https://2024.ieeesam.org/
07.08. - 07.12.	2024 IEEE 37th Computer Security Foundations Symposium (CSF)	Enschede, Netherlands	https://csf2024.ieee-security.org/
07.08. - 07.11.	2024 16th International Conference on Human System Interaction (HSI)	Paris, France	https://hsi2024.welcometohsi.org/
07.08. - 07.11.	2024 IEEE 24th International Conference on Nanotechnology (NANO)	Gijon, Spain	https://2024.ieeenano.org/
07.08. - 07.12.	2024 17th International Conference on Scintillating Materials and their Applications (SCINT)	Milan, Italy	https://www.scint2024.com/
07.09. - 07.11.	2024 IEEE International Conference on Microwaves, Communications, Antennas, Biomedical Engineering and Electronic Systems (COMCAS)	Tel Aviv, Israel	https://www.comcas.org/
07.09. - 07.13	2024 United States National Committee of URSI National Radio Science Meeting (USNC-URSI NRSM)	Boulder, Colorado, USA	https://www.nrsmboulder.org/
07.09. - 07.11.	2024 International Conference on Broadband Communications for Next Generation Networks and Multimedia Applications (CoBCom)	Graz, Austria	https://www.cobcom.tugraz.at/wordpress/
07.09. - 07.12	2024 IEEE/IAS Industrial and Commercial Power System Asia (I&CPS Asia)	Pattaya, Thailand	https://ieee-icps.com/2024/index.html
07.09. - 07.11.	2024 International Conference on Consumer Electronics - Taiwan (ICCE-Taiwan)	Taichung, Taiwan	https://www.icce-tw.org/

일 자	학술대회명	개최장소	홈페이지/연락처
07.09. - 07.12	2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)	Boston, Massachusetts, USA	https://www.aim2024.org/
07.10. - 07.11.	2024 IEEE 30th International Symposium on Local and Metropolitan Area Networks (LANMAN)	Boston, Massachusetts, USA	https://lanman2024.ieee-lanman.org/
07.10. - 07.12.	2024 American Control Conference (ACC)	Toronto, Ontario, Canada	https://acc2024.a2c2.org/
07.11. - 07.13.	2024 IEEE 7th International Conference on Advanced Technologies, Signal and Image Processing (ATSIP)	Sousse, Tunisia	http://www.lab-atms.com/atsip_2024.php
07.12. - 07.14.	2024 2nd World Conference on Communication & Computing (WCONF)	RAIPUR, India	https://wconf.in/
07.12. - 07.14.	2024 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)	Bangalore, India	https://ieee-conecct.org/
07.12. - 07.14.	2024 9th International Conference on Signal and Image Processing (ICSIP)	Nanjing, China	https://www.icsip.org/
07.12. - 07.14.	2024 7th International Conference on Computer Information Science and Application Technology (CISAT)	Hangzhou, China	http://www.cisat.org/
07.12. - 07.13.	2024 4th International Conference of Science and Information Technology in Smart Administration (ICSINTESA)	Balikpapan, Indonesia	https://icsintesa.universitasmulia.ac.id/
07.13. - 07.20.	2024 IEEE International Symposium on Antennas and Propagation and INC/USNC□URSI Radio Science Meeting (AP-S/INC-USNC-URSI)	Firenze, Italy	https://2024.apsursi.org/
07.13. - 07.14.	2024 Intelligent Methods, Systems, and Applications (IMSA)	Giza, Egypt	http://imsa.msa.edu.eg/
07.14. - 07.19.	2024 IEEE INC-USNC-URSI Radio Science Meeting (Joint with AP-S Symposium)	Florence, Italy	https://2024.apsursi.org/
07.14. - 07.17.	2024 IEEE International Professional Communication Conference (ProComm)	Pittsburgh, Pennsylvania, USA	https://procomm.ieee.org/
07.14. - 07.17.	2024 19th International Symposium on Wireless Communication Systems (ISWCS)	Rio de Janeiro, Brazil	http://iswcs2024.usuarios.rdc.puc-rio.br/
07.14. - 07.23	2024 11th International Conference on Machine Intelligence Theory and Applications (MiTA)	Melbourne, Australia	https://mita2024.com/
07.14. - 07.16	2024 3rd International Conference on Energy and Electrical Power Systems (ICEEPS)	Guangzhou, China	http://www.iceeps.net/
07.15. - 07.19.	2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)	Orlando, Florida, USA	https://embc.embs.org/2024/
07.15. - 07.18.	2024 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)	Singapore, Singapore	https://ipfa-ieee.org/2024/
07.15. - 07.17.	2024 IEEE Photonics Society Summer Topicals Meeting Series (SUM)	Bridgetown, Barbados	https://www.ieee-sum.org/
07.15. - 07.19.	2024 IEEE International Conference on Multimedia and Expo (ICME)	Niagara Falls, Ontario, Canada	https://2024.ieeeicme.org/
07.15. - 07.18.	NAECON 2024 - IEEE National Aerospace and Electronics Conference	Dayton, Ohio, USA	http://www.naecon.org/
07.15. - 07.18.	2024 IEEE International Congress On Intelligent and Service-Oriented Systems Engineering (CISOSE)	Shanghai, China	https://ieee-cisose-congress.org/
07.15. - 07.17.	2024 9th International Conference on Image, Vision and Computing (ICIVC)	Suzhou, China	https://www.icivc.org/
07.15. - 07.18.	2024 IEEE Intelligent Mobile Computing	Shanghai, China	https://ieee-cisose-congress.org/
07.15. - 07.18.	2024 IEEE International Conference on Service-Oriented System Engineering (SOSE)	Shanghai, China	https://ieeesose.com/
07.15. - 07.18.	2024 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)	Shanghai, China	https://ieeedapps.com/
07.15. - 07.19.	2024 IEEE Space Computing Conference (SCC)	Mountain View, California, USA	https://smcit-scc.space/

일 자	학술대회명	개최장소	홈페이지/연락처
07.15. - 07.19.	2024 IEEE 10th International Conference on Space Mission Challenges for Information Technology (SMC-IT)	Mountain View, California, USA	https://smcit-scc.space/
07.15. - 07.16.	2024 IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)	Niagara Falls, Ontario, Canada	https://zinc.cse.buffalo.edu/ubmdfl/AVSS2024/index.php
07.15. - 07.18.	2024 IEEE International Conference on Joint Cloud Computing (JCC)	Shanghai, China	http://www.ieeejcc.org/
07.15. - 07.18.	2024 IEEE International Conference on Artificial Intelligence Testing (AITest)	Shanghai, China	https://ieeeaitest.com/
07.16. - 07.18.	2024 IEEE/ACIS 9th International Conference on Big Data, Cloud Computing, and Data Science (BCD)	Kitakyushu, Japan	https://acisinternational.org/conferences/bcd-2024/
07.17. - 07.19.	2024 14th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP)	Rome, Italy	https://comlab.uniroma3.it/CSNDSP2024.php
07.17. - 07.19.	2024 International Conference on Computer, Information and Telecommunication Systems (CITS)	Girona, Spain	https://cits.udg.edu/
07.17. - 07.19.	2024 International Symposium on Computer Science and Educational Technology (ISCSET)	Lauta/Laubusch, Germany	https://www.iscset-conf.org/
07.18. - 07.19.	2024 International Conference on Computational Intelligence for Green and Sustainable Technologies (ICCIGST)	Vijayawada, India	https://www.iccigst.com/
07.18. - 07.20.	2024 9th International Conference on Automation, Control and Robotics Engineering (CACRE)	Jeju Island, Korea (South)	https://cacre.org/
07.18. - 07.19.	2024 IEEE Technology and Engineering Management Society (TEMSCON LATAM)	Panama, Panama	https://attend.ieee.org/temscon-2024/
07.18. - 07.19.	2024 IEEE International Future Energy Challenge (IFEC)	Austin, Texas, USA	http://energychallenge.weebly.com/
07.20. - 07.22.	2024 7th Asia Conference on Energy and Electrical Engineering (ACEEE)	Chengdu, China	https://www.aceee.net/index.html
07.20. - 07.21.	2024 IEEE Symposium on Wireless Technology & Applications (ISWTA)	Kuala Lumpur, Malaysia	https://www.ieee-iswta.org/
07.21. - 07.25.	2024 IEEE Power & Energy Society General Meeting (PESGM)	Seattle, Washington, USA	https://pes-gm.org/
07.22. - 07.26.	2024 IEEE Nuclear and Space Radiation Effects Conference (NSREC)	Ottawa, Ontario, Canada	https://www.nsrec.com/
07.22. - 07.27.	2024 IEEE Radiation Effects Data Workshop (REDW) (in conjunction with 2024 NSREC)	Ottawa, Ontario, Canada	https://www.nsrec.com/
07.22 - 07.24.	2024 IEEE International Conference on Computational Photography (ICCP)	Lausanne, Switzerland	http://iccp2024.iccp-conference.org/
07.22. - 07.25.	2024 International Telecommunications Conference (ITC-Egypt)	Cairo, Egypt	https://itc-egypt-adc.org/
07.22. - 07.24.	2024 IEEE Workshop on Complexity in Engineering (COMPENG)	Florence, Italy	ttps://compeng2024.ino.cnr.it
07.22. - 07.24.	2024 IEEE 25th International Conference on High Performance Switching and Routing (HPSR)	Pisa, Italy	https://hpsr2024.ieee-hpsr.org/
07.23. - 07.25.	2024 IEEE Sensors Applications Symposium (SAS)	Caserta, Italy	https://2024.sensorapps.org/
07.23. - 07.26.	2024 IEEE 44th International Conference on Distributed Computing Systems (ICDCS)	Jersey City, New Jersey, USA	https://icdcs2024.icdcs.org/
07.24. - 07.26.	2024 IEEE 35th International Conference on Application-specific Systems, Architectures and Processors (ASAP)	Hong Kong, Hong Kong	https://www.asap2024.org/
07.24. - 07.27.	2024 Second International Conference on Advances in Information Technology (ICAIT)	Chikkamagaluru, Karnataka, India	https://icait.in/
07.25. - 07.27.	2024 International Conference on Electrical, Computer and Energy Technologies (ICECET)	Sydney, Australia	https://www.icecet.com/2024
07.25. - 07.26.	2024 7th International Conference on Green Technology and Sustainable Development (GTSD)	Ho Chi Minh City, Vietnam	https://gtsd2024.hcmute.edu.vn/
07.25. - 07.27.	2024 10th International Conference on Smart Computing and Communication (ICSCC)	Bali, Indonesia	https://icscc.undiknas.ac.id/

일 자	학술대회명	개최장소	홈페이지/연락처
07.26. - 07.27.	2024 Asia Pacific Conference on Innovation in Technology (APCIT)	MYSORE, India	https://apcit.in/
07.26. - 07.27.	2024 International Conference on Data Science and Network Security (ICDSNS)	Tiptur, India	https://icdsns.co.in/
07.26. - 07.28.	2024 IEEE 6th International Conference on Power, Intelligent Computing and Systems (ICPICS)	Shenyang, China	http://www.icpics.org/
07.27. - 07.28.	2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC)	Gwalior, India	https://scrs.in/conference/aic2024
07.28. - 08.01.	2024 International Conference on Optical MEMS and Nanophotonics (OMN)	San Sebastian, Spain	https://www.nanogune.eu/en/nanogune/newsroom/events/international-conference-optical-mems-nanophotonics-omn-2024
07.28. - 07.30.	2024 IEEE BioSensors Conference (BioSensors)	Cambridge, United Kingdom	https://2024.ieee-biosensors.org/
07.28. - 07.31.	2024 Academic Conference of China Instrument and Control Society (ACCIS)	Chengdu, China	http://meeting.cis.org.cn/EN/
07.29. - 07.31.	2024 33rd International Conference on Computer Communications and Networks (ICCCN)	Kailua-Kona, Hawaii, USA	http://www.icccn.org/icccn24/
07.29. - 07.31.	2024 IEEE International Conference on Omni-layer Intelligent Systems (COINS)	London, United Kingdom	https://coinsconf.com/
07.29. - 08.02.	2024 IEEE International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (3M-NANO)	Zhongshan, China	http://www.3m-nano.org/2024/main/index.asp
07.29. - 08.01.	2024 International Symposium on Educational Technology (ISET)	Macau, Macao	https://hksmic.org.hk/iset/2024/
07.29. - 07.31.	2024 International Conference on Computing, Internet of Things and Microwave Systems (ICCIMS)	Gatineau, Quebec, Canada	https://www.iccims.net/
07.30. - 08.02.	2024 IEEE East Asian School of Information Theory (EASIT)	Kanagawa, Japan	https://easit2024.github.io/
07.31. - 08.03.	2024 IEEE 4th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)	Hyderabad, India	https://sefet.in/
07.31. - 08.02.	2024 IEEE 7th International Conference on Electronic Information and Communication Technology (ICEICT)	Xi'an, China	http://www.iceict.org/ICEICT2024/
〉〉2024년 8월			
08.01. - 08.02.	2024 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)	Virtual	https://icabcd.org/2024/
08.02. - 08.04.	2024 First International Conference on Pioneering Developments in Computer Science & Digital Technologies (IC2SDT)	Delhi, India	https://www.ic2sdt.com/
08.02. - 08.03.	2024 Control Instrumentation System Conference (CISCON)	Manipal, India	https://conference.manipal.edu/ciscon/
08.04. - 08.09.	2024 Conference on Lasers and Electro-Optics Pacific Rim (CLEO-PR)	Incheon, Korea (South)	http://www.cleopr2024.org/
08.04. - 08.07.	2024 IEEE International Communications Energy Conference (INTELEC)	Bengaluru, India	https://www.intelec2024.org/
08.04. - 08.07.	2024 IEEE 14th International Conference on the Properties and Applications of Dielectric Materials (ICPADM)	Phuket, Thailand	https://icpadm2024.kmitl.ac.th/
08.04. - 08.07.	2024 IEEE International Conference on Mechatronics and Automation (ICMA)	Tianjin, China	http://2024.ieee-icma.org/
08.05. - 08.09.	2024 IEEE International Symposium on Electromagnetic Compatibility, Signal & Power Integrity (EMC+SIPI)	Phoenix, Arizona, USA	https://www.emc2024.org/
08.05. - 08.08.	2024 IEEE 19th Conference on Industrial Electronics and Applications (ICIEA)	Kristiansand, Norway	https://www.ieeeiciea.org/2024/
08.05. - 08.08.	2024 IEEE Conference on Games (CoG)	Milan, Italy	https://2024.ieee-cog.org/
08.05. - 08.07.	2023 IEEE 3rd German Education Conference (GECon)	Munich, Germany	https://gecon2024.org/

일 자	학술대회명	개최장소	홈페이지/연락처
08.05. - 08.08.	2024 IEEE Sixth International Conference on DC Microgrids (ICDCM)	Columbia, South Carolina, USA	https://attend.ieee.org/icdcm-2024/
08.05. - 08.07.	2024 IEEE 35th Magnetic Recording Conference (TMRC)	Berkeley, California, USA	https://sites.google.com/andrew.cmu.edu/tmrc2024/home
08.06. - 08.09.	2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)	Kingston, Ontario, Canada	https://ccece2024.ieee.ca/
08.06. - 08.07.	2024 4th International Conference on Emerging Smart Technologies and Applications (eSmarTA)	Sana'a, Yemen	https://esmarta.yostr.org/
08.07. - 08.09.	2024 IEEE 12th International Conference on Serious Games and Applications for Health (SeGAH)	Funchal, Portugal	https://www.segah.org/2024/
08.07. - 08.08.	2024 International Conference on Future Technologies for Smart Society (ICFTSS)	Kuala Lumpur, Malaysia	https://icftss.org/
08.07. - 08.09.	2024 IEEE/CIC International Conference on Communications in China (ICCC)	Hangzhou, China	https://iccc2024.ieee-iccc.org/
08.07. - 08.09.	2024 IEEE 7th International Conference on Multimedia Information Processing and Retrieval (MIPR)	San Jose, California, USA	https://sites.google.com/view/mipr2024
08.07. - 08.09.	2024 IEEE International Conference on Information Reuse and Integration for Data Science (IRI)	San Jose, California, USA	https://homepages.uc.edu/~niunn/IRI24/
08.07. - 08.08.	2024 12th International Conference on Information and Communication Technology (ICoICT)	Bandung, Indonesia	https://www.icoict.org/
08.07. - 08.09.	2024 IEEE/CIC International Conference on Communications in China (ICCC Workshops)	Hangzhou, China	https://iccc2024.ieee-iccc.org/
08.07. - 08.09.	2024 7th International Conference on Intelligent Robotics and Control Engineering (IRCE)	Xi'an, China	https://www.irce.org/
08.08. - 08.09.	2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT)	Kollam, India	Kollam, India
08.08. - 08.10.	2024 First International Conference on Electronics, Communication and Signal Processing (ICECSP)	New Delhi, India	https://www.icecsp.in/
08.08. - 08.09.	2024 International Conference on HVDC (HVDC)	Urumqi, China	https://hvdc2024.csee.org.cn/
08.09. - 08.10.	2024 Second International Conference on Networks, Multimedia and Information Technology (NMITCON)	Bengaluru, India	https://www.nmit.ac.in/nmitcon/
08.09. - 08.11.	2024 IEEE 9th International Conference on Computational Intelligence and Applications (ICCIA)	Haikou, China	https://www.iccia.org/
08.09. - 08.11.	2024 6th International Conference on Internet of Things, Automation and Artificial Intelligence (IoTAAI)	Guangzhou, China	http://www.ic-iotaai.com/
08.09. - 08.11.	2024 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE International Conference on Robotics, Automation and Mechatronics (RAM)	Hangzhou, China	https://www.cis-ram.org/2024/
08.10. - 08.11.	2024 International Conference on Control, Computing, Communication and Materials (ICCCCM)	Prayagraj, India	https://www.united.ac.in/i4cm2024/
08.11. - 08.14.	2024 IEEE 67th International Midwest Symposium on Circuits and Systems (MWSCAS)	Springfield, Massachusetts, USA	https://www.mwscas2024.org/
08.13. - 08.14.	2024 9th International Conference on Mechatronics Engineering (ICOM)	Kuala Lumpur, Malaysia	https://conference.iium.edu.my/icom/index.php/home/
08.13. - 08.16.	2024 IEEE International Symposium on Consumer Technology (ISCT)	Kuta, Bali, Indonesia	https://isct.ctsoc.id/
08.14. - 08.16.	2024 IEEE Research and Applications of Photonics in Defense Conference (RAPID)	Miramar Beach, Florida, USA	https://ieee-rapid.org/
08.15. - 08.17.	2024 7th International Conference on Pattern Recognition and Artificial Intelligence (PRAI)	Hangzhou, China	https://www.prai.net/
08.16. - 08.18.	2024 7th International Conference on Education, Network and Information Technology (ICENIT)	Dalian, China	http://www.icenit.org/
08.17.	2024 IEEE 15th Control and System Graduate Research Colloquium (ICSGRC)	SHAH ALAM, Malaysia	https://sites.google.com/view/icsgrc/home

일 자	학술대회명	개최장소	홈페이지/연락처
08.18. - 08.20.	2024 IEEE 22nd International Conference on Industrial Informatics (INDIN)	Beijing, China	https://indin2024.ieee-ies.org/index.html
08.18. - 08.22.	2024 IEEE International Conference on High Voltage Engineering and Applications (ICHVE)	Berlin, Germany	https://www.ichve2024.org/
08.19. - 08.22.	2024 IEEE International Conference on Blockchain (Blockchain)	Copenhagen, Denmark	https://ieee-cybermatics.org/2024/blockchain/
08.19. - 08.22.	2024 IEEE International Conferences on Internet of Things (iThings) and IEEE Green Computing & Communications (GreenCom) and IEEE Cyber, Physical & Social Computing (CPSCom) and IEEE Smart Data (SmartData) and IEEE Congress on Cybermatics (Cybermatics)	Copenhagen, Denmark	https://ieee-cybermatics.org/2024/
08.19. - 08.24.	2024 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)	Bali, Indonesia	http://www.icspcc.org/
08.20. - 08.23.	2024 21st International Bhurban Conference on Applied Sciences and Technology (IBCAST)	Murree, Pakistan	https://www.ibcast.org.pk/
08.21. - 08.23.	2024 IEEE Global Blockchain Conference (GBC)	Shanghai, China	http://www.ieee-gbc.org/
08.21. - 08.23.	2024 IEEE Conference on Control Technology and Applications (CCTA)	Newcastle upon Tyne, United Kingdom	https://ccta2024.ieeecss.org/
08.21. - 08.24.	2024 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)	Victoria, British Columbia, Canada	https://onlineacademiccommunity.uvic.ca/pacrim2024/#:~:text=In%202024%2C%20it%20will%20be,scenic%20cities%20in%20the%20world.
08.21. - 08.23.	2024 IEEE 30th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA)	Sokcho, Korea (South)	https://rtcsa2024.github.io/
08.22. - 08.24.	2024 International Conference on Advanced Robotics and Intelligent Systems (ARIS)	Taipei, Taiwan	https://aris2024.info/
08.22. - 08.23.	2024 2nd International Conference on Networking, Embedded and Wireless Systems (ICNEWS)	Bangalore, India	https://icnews2024.com/
08.22. - 08.24.	2024 IEEE 5th India Council International Subsections Conference (INDISCON)	Chandigarh, India	https://ieeeindiscon.org/
08.22. - 08.24.	2024 International Conference on Emerging Techniques in Computational Intelligence (ICETCI)	Hyderabad, India	https://ietcint.com/user/home
08.23. - 08.24.	2024 International Conference on Intelligent Algorithms for Computational Intelligence Systems (IACIS)	Hassan, India	https://iacis.co.in/index.php
08.23. - 08.25.	2024 12th International Conference on Traffic and Logistic Engineering (ICTLE)	Macau, China	https://www.ictle.org/
08.23. - 08.25.	2024 5th International Conference on Control, Robotics and Intelligent System (CCRIS)	Macao, Macao	http://www.ccris-conf.net/
08.23. - 08.24.	2024 IEEE 14th International Conference on Control System, Computing and Engineering (ICCSCE)	Penang, Malaysia	http://acscrg.com/iccsce2024/
08.24. - 08.25.	2024 16th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)	Hangzhou, China	http://ihmsc.zju.edu.cn/
08.25. - 08.27.	2024 IEEE Hot Chips 36 Symposium (HCS)	Stanford, California, USA	https://www.hotchips.org/
08.26. - 08.29.	2024 IEEE AUTOTESTCON	National Harbor, Maryland, USA	https://2024.autotestcon.com/
08.26. - 08.30.	2024 33rd IEEE International Conference on Robot and Human Interactive Communication (ROMAN)	Pasadena, California, USA	https://www.ro-man2024.org/overview/welcome
08.26. - 08.28.	2024 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)	Kota Kinabalu, Malaysia	Kota Kinabalu, Malaysia
08.26. - 08.28.	2024 IEEE 3rd Conference on Information Technology and Data Science (CITDS)	Debrecen, Hungary	https://konferencia.unideb.hu/en/announcement-3rd-conference-information-technology-and-data-science
08.27. - 08.29.	2024 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)	Natal, Brazil	https://cmte.ieee.org/cis-bbtc/
08.27. - 08.29.	2024 IEEE 10th International Conference on Photonics (ICP)	Pulau Pinang, Malaysia	https://icp2024.photonicssociety.org.my/

일 자	학술대회명	개최장소	홈페이지/연락처
08.27. - 08.28.	2024 IEEE SustainTech Expo: Technology Solutions for a Sustainable Future (SustainTech)	San Diego, California, USA	https://ieeesustaintechexpo.org/
08.27. - 08.30.	2024 28th International Conference on Methods and Models in Automation and Robotics (MMAR)	Poland	https://mmar.edu.pl/
08.28. - 09.01.	2024 IEEE 20th International Conference on Automation Science and Engineering (CASE)	Bari, Italy	https://2024.ieeecase.org/
08.28. - 08.30.	2024 Advanced Topics on Measurement and Simulation (ATOMS)	Constanta, Romania	http://atoms2024.org/
08.28. - 08.30.	2024 29th International Conference on Automation and Computing (ICAC)	Sunderland, United Kingdom	https://cacsuk.co.uk/icac/
08.29. - 08.31.	2024 IEEE 2nd International Conference on Sensors, Electronics and Computer Engineering (ICSECE)	Jinzhou, China	http://www.icsece.org/
08.29. - 08.31.	2024 IEEE 12th International Conference on Intelligent Systems (IS)	Varna, Bulgaria	https://www.ieee-is.org/
08.30. - 08.31.	2024 IEEE 4th Mysore Sub Section International Conference (MysuruCon)	Mysuru, India	https://vvietmysurucon2024.in/
〉〉2024년 9월			
09.01. - 09.04.	2024 10th IEEE RAS/EMBS International Conference for Biomedical Robotics and Biomechatronics (BioRob)	Heidelberg, Germany	https://www.biorob2024.org/
09.01. - 09.04.	2024 International Conference on Electrical Machines (ICEM)	Torino, Italy	https://www.symposium.it/en/events/2024/26th-international-conference-on-electrical-machines-icem-2024
09.01. - 09.06.	2024 49th International Conference on Infrared, Millimeter, and Terahertz Waves (IRMMW-THz)	Perth, Australia	https://www.irmmw-thz.org/
09.02. - 09.05.	2024 IEEE 35th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)	Valencia, Spain	https://pimrc2024.ieee-pimrc.org/
09.02. - 09.05.	2024 International Symposium on Electromagnetic Compatibility – EMC Europe	Brugge, Belgium	https://emceurope2024.org/
09.02. - 09.06.	2024 IEEE Symposium on Visual Languages and Human-Centric Computing (VL/HCC)	Liverpool, United Kingdom	https://conf.researchr.org/home/vlhcc-2024
09.02. - 09.06.	2024 International Conference on Electromagnetics in Advanced Applications (ICEAA)	Lisbon, Portugal	https://www.iceaa-offshore.org/
09.02. - 09.06.	2024 IEEE-APS Topical Conference on Antennas and Propagation in Wireless Communications (APWC)	Lisbon, Portugal	https://www.iceaa-offshore.org/
09.02. - 09.05.	2024 11th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)	Malmö, Sweden	https://emergingtechnet.org/IOTSMS2024/index.php
09.02. - 09.05.	2024 9th International Conference on Fog and Mobile Edge Computing (FMEC)	Malmö, Sweden	https://emergingtechnet.org/FMEC2024/index.php
09.02. - 09.06.	2024 Energy Conversion Congress & Expo Europe (ECCE Europe)	Darmstadt, Germany	https://www.ecce-europe.org/
09.02. - 09.06.	2024 8th International Symposium on Instrumentation Systems, Circuits and Transducers (INSCIT)	Joao Pessoa, Brazil	https://chiponthecliffs2024.cear.ufpb.br/inscit
09.02. - 09.04.	2024 IEEE International Conference on Cyber Security and Resilience (CSR)	London, United Kingdom	https://www.ieee-csr.org/
09.03. - 09.05.	2024 International Conference on Diagnostics in Electrical Engineering (Diagnostika)	Pilsen, Czech Republic	https://www.diagnostika.zcu.cz/
09.03. - 09.05.	2024 Open Source Modelling and Simulation of Energy Systems (OSMSES)	Vienna, Austria	https://www.osmses2024.org/home
09.03. - 09.05.	2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA)	Malaysia	https://sps.ieeemy.org/icsipa-2024/
09.04. - 09.05.	2024 International Conference on Applied Electronics (AE)	Virtual	https://www.appel.zcu.cz/
09.04. - 09.06.	2024 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)	Craiova, Romania	https://dcti.ucv.ro/inista2024/

일 자	학술대회명	개최장소	홈페이지/연락처
09.05. - 09.06.	2024 International Conference on Information Technology Research and Innovation (ICITRI)	Jakarta, Indonesia	https://icitri.nusamandiri.ac.id/
09.06. - 09.07.	2024 Asian Conference on Intelligent Technologies (ACOIT)	KOLAR, India	https://www.acoit.in/
09.06. - 09.07.	2024 IEEE International Conference on Public Key Infrastructure and its Applications (PKIA)	Bangalore, India	https://pkiindia.in/pkia/
09.07.	2024 IEEE International Conference on Agrosystem Engineering, Technology & Applications (AGRETA)	Kuala Lumpur, Malaysia	https://www.agreta.org/
09.07. - 09.11.	2024 XXXIV International Scientific Symposium Metrology and Metrology Assurance (MMA)	Sozopol, Bulgaria	https://metrology-bg.org/
09.08. - 09.11.	2024 19th Conference on Computer Science and Intelligence Systems (FedCSIS)	Belgrade, Serbia	https://2024.fedcsis.org/
09.08. - 09.11.	2024 12th European Workshop on Visual Information Processing (EUVIP)	Geneva, Switzerland	https://www.euvip2024.org/
09.08. - 09.14.	2024 International Russian Automation Conference (RusAutoCon)	Sochi, Russia	https://rusautocon.org/index-eng.html
09.09. - 09.13.	2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)	Lucca, Italy	https://spawc2024.org/
09.09. - 09.14.	2024 Eighteenth International Congress on Artificial Materials for Novel Wave Phenomena (Metamaterials)	Chania, Greece	https://congress.metamorphose-vi.org/
09.09. - 09.12.	2024 IEEE International Workshop Technical Committee on Communications Quality and Reliability (CQR)	Seattle, Washington, USA	https://cqr2023.ieee-cqr.org/
09.09. - 09.12.	2024 IEEE European Solid-State Electronics Research Conference (ESSERC)	Bruges, Belgium	https://www.esserc2024.org/
09.09. - 09.12.	2024 18th International Workshop on Acoustic Signal Enhancement (IWAENC)	Aalborg, Denmark	https://iwaenc2024.org/
09.09. - 09.11.	2024 ACM/IEEE 6th Symposium on Machine Learning for CAD (MLCAD)	Salt Lake City (Snowbird), Utah, USA	https://mlcad-workshop.org/
09.09. - 09.10.	2024 IEEE International Conference on Advanced Telecommunication and Networking Technologies (ATNT)	Johor Bahru, Malaysia	http://ke.utm.my/atnt2024/
09.10. - 09.13.	2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)	Padova, Italy	https://2024.ieee-etfa.org/
09.10. - 09.12.	2024 International Conference on Smart Energy Systems and Technologies (SEST)	Torino, Italy	https://sest2024.polito.it/
09.10. - 09.11.	2024 IEEE International Conference on Advanced Power Engineering and Energy (APEE)	Johor Bahru, Malaysia	https://attend.ieee.org/apee-2024/
09.11. - 09.13.	2024 IEEE 10th Electronics System-Integration Technology Conference (ESTC)	Berlin, Germany	https://www.estc-conference.net/
09.11. - 09.14.	2024 IEEE ANDESCON	Cusco, Peru	https://attend.ieee.org/andescon/
09.12. - 09.13.	2024 2nd International Conference on Technology Innovation and Its Applications (ICTIIA)	Medan, Indonesia	https://ictiia.uph.edu/
09.12. - 09.13.	2024 International Conference on Signal Processing and Advance Research in Computing (SPARC)	LUCKNOW, India	https://www.amity.edu/lucknow/sparc-2024/
09.15. - 09.20.	2024 IEEE International Conference on Quantum Computing and Engineering (QCE)	Montreal, Quebec, Canada	https://qce.quantum.ieee.org/2024/
09.15. - 09.18	2024 IEEE International Flexible Electronics Technology Conference (IFETC)	Bologna, Italy	http://www.ifetc.org/
09.15. - 09.18	2024 IEEE International Joint Conference on Biometrics (IJCB)	Buffalo, New York, USA	https://attend.ieee.org/ijcb-2024/
09.16. - 09.19.	2024 IEEE IAS Petroleum and Chemical Industry Technical Conference (PCIC)	Orlando, Florida, USA	https://ieeepcic.com/
09.16. - 09.20.	2024 IEEE International Conference on Autonomic Computing and Self-Organizing Systems (ACSOS)	Aarhus, Denmark	https://2024.acsos.org/
09.16. - 09.18.	2024 International Symposium ELMAR	Zadar, Croatia	https://www.elmar-zadar.org/2024/

일 자	학술대회명	개최장소	홈페이지/연락처
09.16. - 09.18.	2024 IEEE Workshop on Wide Bandgap Power Devices and Applications in Europe (WiPDA Europe)	Cardiff, United Kingdom	https://wipda-europe.org/
09.17. - 09.20.	2024 International Topical Meeting on Microwave Photonics (MWP)	Pisa, Italy	https://mwp2024.org/
09.17. - 09.20.	2024 4th Intelligent Cybersecurity Conference (ICSC)	Valencia, Spain	https://icsc-conference.org/2024/index.php
09.17. - 09.19.	2024 XXXIII International Scientific Conference Electronics (ET)	Sozopol, Bulgari	https://e-university.tu-sofia.bg/e-conf/?konf=24
09.18. - 09.20.	2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI)	Milano, Italy	https://2024.ieee-rtsi.org/
09.18. - 09.20.	2024 IEEE Biennial Congress of Argentina (ARGENCON)	San Nicolás de los Arroyos, Argentina	https://attend.ieee.org/argencon-2024/san-nicolas/
09.18. - 09.20.	2024 IEEE 14th International Workshop on Applied Measurements for Power Systems (AMPS)	Caserta, Italy	https://amps2024.ieee-ims.org/
09.18. - 09.20.	2024 7th International Conference on Contemporary Computing and Informatics (IC3I)	Greater Noida, India	https://amity.edu/ic3i2024/
09.19. - 09.21.	2024 IEEE 22nd Jubilee International Symposium on Intelligent Systems and Informatics (SISY)	Pula, Croatia	https://conf.uni-obuda.hu/sisy2024/
09.19. - 09.21.	2024 14th International Conference on Advanced Computer Information Technologies (ACIT)	Ceske Budejovice, Czech Republic	https://acit.wunu.edu.ua/
09.20. - 09.22.	2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)	Chongqing, China	http://www.itnec.org/
09.20. - 09.22.	2024 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES)	KOTTAYAM, India	https://ieeespices2024.iiitkottayam.ac.in/
09.20. - 09.21.	2024 International Conference on Distributed Systems, Computer Networks and Cybersecurity (ICDSCNC)	Bengaluru, India	https://icdscnc.co.in/
09.20. - 09.23.	2024 6th International Conference on Circuits and Systems (ICCS)	Chengdu, China	https://www.iccs.org/
09.20. - 09.22.	2024 7th International Conference on Mechatronics, Robotics and Automation (ICMRA)	Wuhan, China	https://www.icmra.org/index.html
09.20. - 09.22.	2024 3rd International Conference on Artificial Intelligence and Computer Information Technology (AICIT)	Yichang, China	http://www.2024aicit.com/
09.21. - 09.22.	2024 International Symposium on Parallel Computing and Distributed Systems (PCDS)	Singapore, Singapore	https://www.pcds.cc/
09.21. - 09.22.	2024 IEEE North Karnataka Subsection Flagship International Conference (NKCon)	Bagalkote, India	https://ieeenkcon2024.org/
09.22. - 09.25.	2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP)	London, United Kingdom	https://2024.ieeemlsp.org/
09.22. - 09.26.	2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS)	Taipei, Taiwan	https://2024.ieee-uffc-js.org/
09.22. - 09.27.	2024 ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems (MODELS)	Linz, Austria	https://conf.researchr.org/home/models-2024
09.22. - 09.26.	2024 10th Euro-Asian Pulsed Power Conference, 25th International Conference on High-Power Particle Beams and 20th International Symposium on Electromagnetic Launch Technology (EAPPC/BEAMS/EML)	Amsterdam, Netherlands	https://www.eappc-beams-eml2024.org/
09.22. - 09.25	2024 IEEE 12th Asia-Pacific Conference on Antennas and Propagation (APCAP)	Nanjing, China	http://www.em-conf.com/apcap2024/
09.23. - 09.26.	OCEANS 2024 - Halifax	Halifax, Nova Scotia, Canada	https://halifax24.oceansconference.org/contact/
09.23. - 09.25.	2024 IEEE 21st International Conference on Mobile Ad-Hoc and Smart Systems (MASS)	Seoul, Korea (South)	https://sites.google.com/view/ieee-mass-2024/
09.23. - 09.24.	2024 19th European Microwave Integrated Circuits Conference (EuMIC)	Paris, France	https://www.eumweek.com/conferences/eumic.html
09.23. - 09.27.	2024 IEEE High Performance Extreme Computing Conference (HPEC)	Virtual	https://ieee-hpec.org/

일 자	학술대회명	개최장소	홈페이지/연락처
09.24. - 09.27.	2024 IEEE International Conference on Cluster Computing (CLUSTER)	Kobe, Japan	https://clustercomp.org/committee/
09.24. - 09.27.	2024 IEEE International Conference on Cloud Engineering (IC2E)	Paphos, Cyprus	https://conferences.computer.org/IC2E/2024/index.html
09.24. - 09.27.	2024 International Conference on Intelligent Computing, Communication, Networking and Services (ICCNS)	DUBROVNIK, Croatia	https://iccns-conference.org/2024/
09.24. - 09.27.	2024 Fifth International Conference on Intelligent Data Science Technologies and Applications (IDSTA)	DUBROVNIK, Croatia	https://idsta-conference.org/2024/
09.24. - 09.27.	2024 IEEE International Conference on Cluster Computing Workshops (CLUSTER Workshops)	Kobe, Japan	https://clustercomp.org/committee/
09.24. - 09.26.	2024 54th European Microwave Conference (EuMC)	Paris, France	https://www.eumweek.com/conferences/eumc.html
09.25. - 09.27.	2024 IEEE 18th International Conference on Application of Information and Communication Technologies (AICT)	Turin, Italy	https://www.aict.info/?csc=2024
09.25. - 09.27.	2024 International Conference of the Biometrics Special Interest Group (BIOSIG)	Darmstadt, Germany	https://biosig.de/
09.25. - 09.27.	2024 Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)	Poznan, Poland	https://spaconference.org.pl/
09.25. - 09.27.	2024 30th International Workshop on Thermal Investigations of ICs and Systems (THERMINIC)	Toulouse, France	https://therminic2024.eu/
09.25. - 09.26.	2024 2nd International Conference on Integrated Systems in Medical Technologies (ISMT)	Erlangen, Germany	https://www.faps.fau.de/ismt/
09.25. - 09.27.	2024 7th International Conference on Renewable Energy and Power Engineering (REPE)	Beijing, China	https://www.repe.net/
09.25. - 09.27.	2024 IEEE VII Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC)	David, Panama	https://redamitic.utp.ac.pa/amitic2024/
09.25. - 09.27.	2024 21st European Radar Conference (EuRAD)	Paris, France	https://www.eumweek.com/conferences/eurad.html
09.26. - 09.27.	2024 International Conference on Actual Problems of Electron Devices Engineering (APEDE)	Saratov, Russia	http://apede.sstu.ru/
09.26. - 09.28.	2024 2nd International Conference on Cyber Physical Systems, Power Electronics and Electric Vehicles (ICPEEV)	Hyderabad, India	https://www.mahindrauniversity.edu.in/icpeev-2024/
09.27. - 09.29.	2024 International Ural Conference on Electrical Power Engineering (UralCon)	Magnitogorsk, Russia	https://uralcon.su-ieee.ru/index-eng.html
09.27. - 09.29.	2024 IEEE 7th International Conference on Information Systems and Computer Aided Education (ICISCAE)	Dalian, China	http://www.iciscae.org/
09.27. - 09.29.	2024 IEEE Region 10 Symposium (TENSYMP)	New Delhi, India	https://ieeedelhi-tensymp2024.org/
09.27. - 09.28.	2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)	Indore, India	https://aitr.ac.in/conference/
09.28. - 09.30.	2024 IEEE International Conference on Cognitive Computing and Complex Data (ICCD)	Qinzhou City, China	http://www.icc-cs.cn/
09.28. - 09.29.	2024 4th International Conference on Computer, Communication, Control & Information Technology (C3IT)	Hooghly, India	https://icic3-2023.aot.edu.in/
09.29. - 10.04.	2024 Embedded Systems Week (ESWEEK)	Raleigh, North Carolina, US	https://esweek.org/
09.29. - 10.04.	2024 International Conference on Compilers, Architecture, and Synthesis for Embedded Systems (CASES)	Raleigh, North Carolina, USA	https://esweek.org/cases/
09.29. - 10.04.	2024 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)	Raleigh, North Carolina, USA	https://esweek.org/codes-isss/
09.29. - 10.04.	2024 International Conference on Embedded Software (EMSOFT)	Raleigh, North Carolina, USA	https://esweek.org/emosft/
09.29. - 09.30.	2024 2nd International Conference on Sustaining Heritage: Embracing Technological Advancements (ICSH)	Virtual	https://icsh.uob.edu.bh/

일 자	학술대회명	개최장소	홈페이지/연락처
09.29. - 10.02.	2024 IEEE International Semiconductor Laser Conference (ISLC)	Orlando, Florida, USA	https://ieee-islc.org/
09.30. - 10.03	2024 IEEE 21st International Power Electronics and Motion Control Conference (PEMC)	Pilsen, Czech Republic	https://ieee-pemc2024.org/
09.30. - 10.02.	2024 Artificial Intelligence x Humanities, Education, and Art (AIxHEART)	Laguna Hills, California, USA	https://www.aixheart.org/
09.30. - 10.03.	2024 37th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)	Manaus, Brazil	https://sibgrapi2024.uea.edu.br/
09.30. - 10.02.	2024 Conference on AI, Science, Engineering, and Technology (AIxSET)	Laguna Hills, California, USA	https://www.aixset.org/
09.30. - 10.04.	2024 IEEE SmartBlock4Africa	Accra, Ghana	https://ieee.vvu.edu.gh/
09.30. - 10.02.	2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)	Erlangen, Germany	https://internetofsounds.net/is2_2024/
〉〉2024년 10월			
10.01. - 10.03.	2024 IEEE 12th Region 10 Humanitarian Technology Conference (R10-HTC)	Kuala Lumpur, Malaysia	https://ieeemy.org/htc2024/
10.01. - 10.03.	2024 International Conference on Advances in Electrical and Communication Technologies (ICAECOT)	Setif, Algeria	https://icaecot24.univ-setif.dz/
10.02. - 10.04.	2024 19th Biennial Baltic Electronics Conference (BEC)	Tallinn, Estonia	https://taltech.ee/en/bec2024
10.02. - 10.04.	2024 15th International Conference on Network of the Future (NoF)	Castelldefels, Spain	https://nof.dnac.org/
10.02. - 10.04.	2024 IEEE 26th International Workshop on Multimedia Signal Processing (MMSP)	West Lafayette, Indiana, USA	https://attend.ieee.org/mmsp-2024/
10.03. - 10.05.	2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)	Kirtipur, Nepal	http://i-smac.org/
10.03. - 10.04.	2024 International Symposium on Accreditation of Engineering and Computing Education (ICACIT)	Bogota, Colombia	http://icacit.org.pe/simposio/
10.03. - 10.04.	2024 22nd ACM-IEEE International Symposium on Formal Methods and Models for System Design (MEMOCODE)	Raleigh, North Carolina, USA	https://memocode2024.github.io/index.html
10.04. - 10.06.	2024 Second International Conference on Microwave, Antenna and Communication (MAC)	Dehradun, India	https://www.mac2024.org/
10.04. - 10.06.	2024 5th IEEE Global Conference for Advancement in Technology (GCAT)	Bangalore, India	http://globeconf.org/
10.06. - 10.11.	2024 IEEE International Conference on Software Maintenance and Evolution (ICSME)	Flagstaff, Arizona, USA	https://conf.researchr.org/home/icsme-2024
10.06. - 10.10.	2024 IEEE 69th Holm Conference on Electrical Contacts (HOLM)	Annapolis, Maryland, USA	https://ieee-holm.org/
10.06. - 10.09.	2024 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)	Auburn, Alabama, USA	https://ceidp.org/
10.06. - 10.09.	2024 IEEE 33rd Conference on Electrical Performance of Electronic Packaging and Systems (EPEPS)	Toronto, Ontario, Canada	https://attend.ieee.org/epeps-2024/
10.06. - 10.10.	2024 IEEE 11th International Conference on Data Science and Advanced Analytics (DSAA)	San Diego, California, USA	http://dsaa2024.dsaa.co/
10.07. - 10.10.	2024 IEEE International Conference on Systems, Man, and Cybernetics (SMC)	Kuching, Malaysia	https://www.ieeesmc2024.org/#/home
10.07. - 10.10.	2024 IEEE Marine Imaging Workshop (MIW)	Monterey, California, USA	https://miw2024.org/
10.07. - 10.10.	2024 IEEE Secure Development Conference (SecDev)	Pittsburgh, Pennsylvania, USA	https://secdev.ieee.org/2024/home
10.07. - 10.10.	2024 IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)	Kharkiv, Ukraine	http://khpiweek.ieee.org.ua/

일 자	학술대회명	개최장소	홈페이지/연락처
10.07. - 10.10.	2024 World Technology Summit (WTS)	San Jose, California, USA	https://www.afternic.com/forsale/wts.org?utm_source=TDFS_DASLNC&utm_medium=DASLNC&utm_campaign=TDFS_DASLNC&traffic_type=TDFS_DASLNC&traffic_id=daslnc&
10.07. - 10.10.	2024 IEEE 100th Vehicular Technology Conference (VTC2024-Fall)	Washington, District of Columbia, USA	https://events.vtsociety.org/vtc2024-fall/
10.07. - 10.10.	2024 IEEE Vehicle Power and Propulsion Conference (VPPC)	Washington, District of Columbia, USA	http://www.events.vtsociety.org/vppc2024/
10.08. - 10.10.	2024 IEEE 49th Conference on Local Computer Networks (LCN)	Normandy, France	https://www.ieeelcn.org/
10.08. - 10.10.	2024 IEEE International Conference and Expo on Real Time Communications at IIT (RTC)	Chicago, Illinois, USA	https://www.rtc-conference.com/2024/
10.09. - 10.11	2024 20th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)	Paris, France	http://www.wimob.org/wimob2024/
10.09. - 10.11	2024 IEEE Conference on Antenna Measurements and Applications (CAMA)	Da Nang, Vietnam	https://2024ieeecama.org/
10.09. - 10.11	2024 International Semiconductor Conference (CAS)	Sinaia, Romania	https://www.imt.ro/cas/
10.10. - 10.12.	2024 28th International Conference on System Theory, Control and Computing (ICSTCC)	Sinaia, Romania	https://icstcc2024.ace.ucv.ro/
10.10. - 10.12.	2024 IEEE 65th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)	Riga, Latvia	http://www.conference.rtu.lv/
10.10. - 10.12.	2024 International Conference Automatics and Informatics (ICAI)	Varna, Bulgaria	http://www.icai-conf.org/
10.10. - 10.12.	2024 IEEE International Conference on Artificial Intelligence & Green Energy (ICAIGE)	Yasmine Hammamet, Tunisia	https://icaige.tn/
10.11. - 10.13.	2024 IEEE International Conference on e-Business Engineering (ICEBE)	Shanghai, China	https://conferences.computer.org/icebe/2024/index.htm
10.11. - 10.13.	2024 Global Reliability and Prognostics and Health Management Conference (PHM-Beijing)	Beijing, China	https://www.icphm.org/
10.13. - 10.18.	2024 IEEE Visualization and Visual Analytics (VIS)	Tampa Bay, Florida, USA	https://ieeevis.org/year/2024/welcome
10.13. - 10.16.	2024 IEEE Frontiers in Education Conference (FIE)	Washington, District of Columbia, USA	https://2024.fie-conference.org/
10.14. - 10.18.	2024 6th International Conference on Process Mining (ICPM)	Kgs. Lyngby, Denmark	https://icpmconference.org/2024/
10.14. - 10.16.	2024 IEEE International Conference on Design, Test and Technology of Integrated Systems (DTTIS)	Aix-EN-PROVENCE, France	https://dttis2024.org/
10.14. - 10.17.	2024 14th International Conference on Indoor Positioning and Indoor Navigation (IPIN)	Kowloon, Hong Kong	http://ipin-conference.org/2024/
10.14. - 10.15.	2024 IEEE Digital Platforms and Societal Harms	Washington, District of Columbia, USA	https://www.computer.org/conferences/dpsh-2024
10.15. - 10.18.	2024 Formal Methods in Computer-Aided Design (FMCAD)	Prague, Czech Republic	https://fmcad.org/FMCAD24/
10.15. - 10.18.	2024 21st International Conference on Harmonics and Quality of Power (ICHQP)	Chengdu, China	https://www.ichqp2024.org/
10.16 - 10.18.	2024 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)	Beijing, China	http://ccci.udg.edu/
10.16 - 10.18.	2024 International Conference on Smart Systems and Technologies (SST)	Osijek, Croatia	https://sst-conference.org/
10.16 - 10.18.	2024 New Trends in Signal Processing (NTSP)	Demanovska Dolina, Slovakia	https://ntsp.sk/2024/home/
10.16 - 10.18.	2024 15th International Conference on Information and Communication Technology Convergence (ICTC)	Jeju Island, Korea (South)	https://ictc.org/
10.16 - 10.18.	2024 IEEE 40th International Electronics Manufacturing Technology (IEMT)	Penang, Malaysia	https://iemt.com.my/

일 자	학술대회명	개최장소	홈페이지/연락처
10.16 - 10.18.	2024 International Conference on Engineering Management of Communication and Technology (EMCTECH)	Vienna, Austria	http://media-publisher.eu/conference-emctech/call-for-papers/
10.16 - 10.18.	2024 12th Electrical Power, Electronics, Communications, Controls and Informatics Seminar (EECCIS)	Malang, Indonesia	https://eeccis.ub.ac.id/
10.17. - 10.18.	2024 10th International Conference on Optimization and Applications (ICOA)	Almeria, Spain	http://w3.ual.es/Congresos/icoa/
10.17. - 10.18.	2024 International Conference on Electrical Engineering and Photonics (EExPolytech)	Saint Petersburg, Russia	https://eexpolytech.spbstu.ru/
10.17. - 10.19.	2024 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS)	Pune, India	https://icbds.ieeepunesection.org/
10.17. - 10.19.	2024 International Conference on Advanced Technologies for Communications (ATC)	Ho Chi Minh City, Vietnam	https://atc-conf.org/
10.18. - 10.20.	2024 2nd International Conference on Advanced Electrical Equipment and Reliable Operation (IAEERO)	Beijing, China	http://www.aeero.cn/
10.18. - 10.20.	2024 IEEE International Conference on Unmanned Systems (ICUS)	Nanjing, China	http://icus.c2.org.cn/
10.18. - 10.20.	2024 10th Asia Conference on Mechanical Engineering and Aerospace Engineering (MEAE)	Taichang, China	https://www.meae.org/index.html
10.18. - 10.19.	2024 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)	Mangalore, India	http://www.ieee-discover.org/
10.18. - 10.20.	2024 6th International Communication Engineering and Cloud Computing Conference (CECCC)	Chengdu, China	https://www.ceccc.org/
10.18. - 10.20.	2024 International Conference on Networking, Sensing and Control (ICNSC)	Hangzhou, China	http://www.icnsc2024.cn/
10.18. - 10.19.	2024 First International Conference on Software, Systems and Information Technology (SSITCON)	Tumkur, India	https://ssitcon.in/
10.19. - 10.20.	2024 IEEE 12th International Conference on Computer Science and Network Technology (ICCSNT)	Dalian, China	http://www.iccsnt.org/ICCSNT2024/
10.19. - 10.20.	2024 IEEE International Conference on Computer Vision and Machine Intelligence (CVMI)	Prayagraj, India	https://cvmi2024.iiita.ac.in/
10.20. - 10.24.	2024 IEEE Industry Applications Society Annual Meeting (IAS)	Phoenix, Arizona, USA	https://ias.ieee.org/
10.20. - 10.23.	2024 IEEE SENSORS	Kobe, Japan	https://2024.ieee-sensorsconference.org/
10.20. - 10.24.	2024 10th International Conference on Condition Monitoring and Diagnosis (CMD)	Gangneung, Korea (South)	https://www.cmd2024.org/
10.20. - 10.24.	2024 IEEE Energy Conversion Congress and Exposition (ECCE)	Phoenix, Arizona, USA	https://www.ieee-ecce.org/2024/
10.20. - 10.23.	2024 IEEE International Conference on Automation/XXVI Congress of the Chilean Association of Automatic Control (ICA-ACCA)	Santiago, Chile	https://ieee-ica-acca-2024.cl/
10.21. - 10.24.	2024 IEEE Nanotechnology Materials and Devices Conference (NMDC)	Salt Lake City, Utah, USA	https://ieeenmdc.org/nmdc-2024/
10.21. - 10.25.	2024 International Radar Conference (RADAR)	Rennes, France	https://conference-radar.org/
10.21. - 10.25.	2024 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)	Bellevue, Washington, USA	https://www.ismar.net/
10.21. - 10.23.	2024 3rd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)	Limassol, Cyprus	https://synergy-med.org/
10.21. - 10.23.	2024 International Conference on Sustainable Energy: Energy Transition and Net-Zero Climate Future (ICUE)	Pattaya City, Thailand	https://icue2024.ait.ac.th/
10.21. - 10.24.	2024 17th International Workshop on Variable Structure Systems (VSS)	Abu Dhabi, United Arab Emirates	https://vss2024.net/
10.21. - 10.23.	2024 IEEE International Automated Vehicle Validation Conference (IAVVC)	Pittsburgh, Pennsylvania, USA	http://2024.iavvc.org/

일 자	학술대회명	개최장소	홈페이지/연락처
10.21. - 10.23.	2024 IEEE 29th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)	Athens, Greece	https://camad2024.ieee-camad.org/
10.22. - 10.25.	2024 IEEE International Symposium on Phased Array Systems and Technology (ARRAY)	Boston, Massachusetts, USA	https://www.ieee-array.org/
10.22. - 10.25.	2024 IEEE 17th International Conference on Solid-State & Integrated Circuit Technology (ICSICT)	Zhuhai, China	https://www.icsict.com/
10.22. - 10.23.	2024 DGON Inertial Sensors and Applications (ISA)	Braunschweig, Germany	http://www.tu-braunschweig.de/iff/dgon-iss-inertial-sensors-and-systems
10.22. - 10.23.	2024 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFEr)	Hoboken, New Jersey, USA	https://www.stevens.edu/page-right-nav/cifer-conference-2024
10.22. - 10.25.	2024 International Symposium on Networks, Computers and Communications (ISNCC)	Washington DC, District of Columbia, USA	https://www.isncc-conf.org/
10.22. - 10.24.	2024 IEEE Future Networks World Forum (FNWF)	Dubai, United Arab Emirates	https://fnwf2024.ieee.org/
10.23. - 10.24.	2024 Sixth International Conference on Intelligent Computing in Data Sciences (ICDS)	Marrakech, Morocco	http://www.researchnetwork.ma/icds2024/index.html
10.23. - 10.25.	2024 16th International Conference on Information Technology and Electrical Engineering (ICITEE)	Bali, Indonesia	http://icitee.ugm.ac.id/2024/
10.23. - 10.26.	2024 14th International Symposium on Antennas, Propagation and EM Theory (ISAPE)	Hefei, China	https://www.rpsoc.cn/isape2024/index.php
10.23. - 10.25.	2024 IEEE 14th International Symposium on Industrial Embedded Systems (SIES)	Chengdu, China	https://www.ieee-sies.org/
10.23. - 10.25.	2024 9th International Engineering, Sciences and Technology Conference (IESTEC)	Panama City, Panama	https://congreso.utp.ac.pa/
10.23. - 10.25.	2024 IEEE 6th International Conference on Civil Aviation Safety and Information Technology (ICCASIT)	Hangzhou, China	http://www.iccasit.org/
10.23. - 10.25.	2024 21st International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)	Mexico City, Mexico	https://cce.cinvestav.mx/
10.23. - 10.25.	2024 5th International Conference on Biomedical Engineering (IBIOMED)	Bali, Indonesia	http://ibiomed.ugm.ac.id/2024/
10.23. - 10.26.	2024 IEEE Global Humanitarian Technology Conference (GHTC)	Radnor, Pennsylvania, USA	https://ieeeghtc.org/
10.24. - 10.26.	2024 IEEE Biomedical Circuits and Systems Conference (BioCAS)	Xi'an, China	http://2024.ieee-biocas.org/
10.24. - 10.26.	2024 12th International Conference on Internet of Everything, Microwave, Embedded, Communication and Networks (IEMECON)	Jaipur, India	https://iemecon.uem.edu.in/
10.24. - 10.26.	2024 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)	Guangzhou, China	https://cyberc.org/
10.24. - 10.26.	2024 International Conference on Applied and Theoretical Electricity (ICATE)	Craiova, Romania	https://elth.ucv.ro/icate/icate2024/
10.24. - 10.25.	2024 First International Conference on Innovations in Communications, Electrical and Computer Engineering (ICICEC)	Davangere, India	https://icicec.com/
10.24. - 10.25.	2024 International Conference on Computing, Sciences and Communications (ICCSC)	Ghaziabad, India	http://iccsc.info/
10.25. - 10.26	2024 33rd Wireless and Optical Communications Conference (WOCC)	Hsinchu, Taiwan	https://www.wocc.org/wocc2024/
10.25. - 10.27	2024 IEEE 24th International Conference on Communication Technology (ICCT)	Chengdu, China	https://www.icct2024.com/
10.25. - 11.02.	2024 IEEE Nuclear Science Symposium (NSS), Medical Imaging Conference (MIC) and Room Temperature Semiconductor Detector Conference (RTSD)	Tampa, Florida, USA	https://nssmic.ieee.org/2024/
10.25. - 10.26	2024 Global Conference on Communications and Information Technologies (GCCIT)	BANGALORE, India	https://gccit.in/index.php

일 자	학술대회명	개최장소	홈페이지/연락처
10.25. - 10.27	2024 IEEE PES 16th Asia-Pacific Power and Energy Engineering Conference (APPEEC)	Nanjing, China	https://ieee-appeec.tiemeeting.com/
10.25. - 10.26	2024 International Conference on Artificial Intelligence, Metaverse and Cybersecurity (ICAMAC)	Dubai, United Arab Emirates	https://www.icamac.com/
10.25. - 10.27	2024 8th International Conference on Smart Grid and Smart Cities (ICSGSC)	Shanghai, China	https://www.icsgsc.org/
10.26. - 10.27.	2024 International Conference on Innovations in Science, Engineering and Technology (ICISET)	Chittagong, Bangladesh	https://iciset.iiuc.ac.bd/
10.26. - 10.28.	2024 4th International Conference on Artificial Intelligence and Signal Processing (AISP)	VIJAYAWADA, India	https://www.aisp.in/
10.26. - 10.28.	2024 9th International Conference on Computer Science and Engineering (UBMK)	Antalya, Turkiye	https://ubmk.org.tr/en/
10.27. - 10.30.	2024 IEEE International Conference on Image Processing (ICIP)	Abu Dhabi, United Arab Emirates	https://2024.ieeeicip.org/
10.27. - 10.30.	2024 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS)	Fort Lauderdale, Florida, USA	https://bcicts.org/
10.27. - 10.30.	2024 58th Asilomar Conference on Signals, Systems, and Computers	Pacific Grove, California, USA	https://www.asilomarsscconf.org/
10.27. - 11.01.	2024 39th IEEE/ACM International Conference on Automated Software Engineering (ASE)	Sacramento, California, USA	https://conf.researchr.org/home/ase-2024
10.28. - 11.01.	MILCOM 2024 - 2024 IEEE Military Communications Conference (MILCOM)	Washington, District of Columbia, USA	https://milcom2024.ieee-milcom.org/
10.28. - 10.31.	2024 IEEE 32nd International Conference on Network Protocols (ICNP)	Charleroi, Belgium	https://icnp24.cs.ucr.edu/
10.28. - 10.31.	2024 IEEE 17th International Conference on Signal Processing (ICSP)	Suzhou, China	http://www.icsp-ieee.org/
10.28. - 10.30.	2024 IEEE 10th International Conference on Collaboration and Internet Computing (CIC)	Washington, District of Columbia, USA	http://www.sis.pitt.edu/lersais/conference/cic/2024/
10.28. - 10.31.	2024 IEEE 6th International Conference on Cognitive Machine Intelligence (CogMI)	Washington, District of Columbia, USA	http://www.sis.pitt.edu/lersais/conference/cogmi/2024/
10.28. - 10.31.	2024 IEEE 35th International Symposium on Software Reliability Engineering (ISSRE)	Tsukuba, Japan	https://issre.github.io/2024/
10.28. - 10.30.	2024 IEEE 6th International Conference on Trust, Privacy and Security in Intelligent Systems, and Applications (TPS-ISA)	Washington, District of Columbia, USA	http://www.sis.pitt.edu/lersais/conference/tps/2024/
10.29. - 11.01.	2024 IEEE International Smart Cities Conference (ISC2)	Pattaya, Thailand	https://www.isc2-2024.com/

The Magazine of the IEIE

특별회원사 명단

회원사	대표자	주 소	전 화	홈페이지
(주)디비하이텍	조기석	경기도 부천시 수도로 90(도당동)	032-680-4700	www.dbhitek.com
(주)레티널	김재혁	경기도 안양시 동안구 부림로170번지 41-10, 4층	02-6959-7007	https://letinar.com
(주)마르시스	박용규	서울시 강남구 언주로 85길 7	02-3445-3999	http://www.marusys.com
(주)세미파이브	조명현	경기도 성남시 분당구 양현로 322, 코리아디자인센터 2층		http://www.semifive.com
(주)센서위드유	이윤식	울산광역시 울주군 언양읍 유니스트길 50, 106동 501-4호	052-912-4282	http://www.sensorwyou.com
(주)에스비솔루션	변영재	울산광역시 울주군 언양읍 유니스트길 50, 106동 401-3호	052-217-7343	http://www.sb-solutions.co.kr
(주)와이솔	김지호	경기도 오산시 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	김영섭	경기도 성남시 분당구 정자동 206	031-727-0114	http://www.kt.com
LG이노텍㈜	문혁수	서울시 강서구 마곡중앙10로 30	02-3777-1114	www.lginnotek.com
LG전자㈜	조주완	서울시 영등포구 여의도동 30	02-3777-1114	http://www.lge.co.kr
LIG넥스원	김지찬	서울시 서초구 강남대로 369(서초동, 나라빌딩)	02-1644-2005	http://www.lignex1.com
LPKF Laser&Electronics	이용상,벤델레괴츠마티아스	경기도 안양시 동안구 흥안대로 427번길	031-689-3660	www.lpkf.com/kr
SK텔레콤㈜	유영상	서울시 중구 을지로65(을지로2가) SK T-타워	02-2121-2114	http://www.sktelecom.com
SK하이닉스㈜	곽노정	경기도 이천시 부발읍 아미리 산 136-1	031-630-4114	http://www.skhynix.com
네이버㈜	최수연	경기도 성남시 분당구 불정로 6 (정자동 그린팩토리)	031-784-2560	https://www.navercorp.com
누리미디어	최순일	서울시 영등포구 선유로 63, 4층(문래동 6가)	02-710-5300	http://www.nurimedia.co.kr
대덕전자(주)	신영환	경기도 안산시 단원구 강촌로230 (목내동 475)	031-8040-8000	http://www.daeduck.com
대전테크노파크	김우연	대전시 유성구 테크로9로	042-930-4300	www.djtp.or.kr
도쿄일렉트론코리아(주)	원제형	경기도 화성시 장안면 장안공단 6길 51	031-260-5000	https://www.tel.com
(주)동인시스템	곽동달	부산광역시 해운대구 센텀북대로 60,	051-787-7288	http://www.donginsm.com
리얼텍코리아 주식회사	팅치창	서울시 서초구 사임당로 18, 석오빌딩 5층	070-4120-7966	www.realtek.cpm/en
비젼테크	이원복	대전 유성구 테크노2로 187, 미건테크노월드2차 1층 118호	042-934-0236	http://www.visiontechkorea.com
삼성전자㈜	한종희, 경계현	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㈜	최기환	충청남도 천안시 서북구 3공단1로 20(백석동)	041-629-7480	http://www.steco.co.kr
에스에스앤씨㈜	한은혜	서울시 영등포구 당산로171, 1301	02-6925-2550	http://www.secnc.co.kr
에어스메디컬	이진구	서울시 관악구 남부순환로 1838	070-7777-3186	www.airsmed.com
오토아이티㈜	정명환	대구시 수성구 알파시티1로 117	053-795-6303	www.auto-it.co.kr
유정시스템㈜	이재훈	서울시 구로구 디지털로26길 110	02-852-8721	www.yjsys.co.kr
정보통신정책연구원	배경율	충북 진천군 덕산읍 정통로 18	043-531-4389	www.kisdi.re.kr
㈜LX세미콘	이윤태	대전시 유성구 탑립동 707	042-712-7700	www.lxsemicon.com

회원사	대표자	주 소	전 화	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
HL만도㈜	조성현	경기도 평택시 포승읍 하만호길 32	02-6244-2114	https://www.hlmando.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 벌말로48(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
㈜쏠리드	정준, 이승희	경기도 성남시 분당구 판교역로 220 쏠리드스페이스	031-627-6000	http://www.st.co.kr
㈜유니트론텍	남궁 선	서울시 강남구 영동대로 638(삼도빌딩) 9층	02-573-6800	http://unitrontech.com
㈜코클리어닷에이아이	한윤창	서울시 강남구 봉은사로 51길 26		www.cochl.ai
㈜크레셈	오상민	대전시 유성구 대덕대로 582, 4층 402호(도룡동, 옥토빌딩)	031-427-3445	http://wwww.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	강석진	경상남도 진주시 동진로 430	055-751-9380	www.kosmes.or.kr
케이케이테크(주)	김경하	경기도 안성시 대덕면 무능로132	031-678-1586	http://www.k-ktech.co.kr
코어인사이트㈜	유용훈	경기도 성남시 중원구 갈마치로 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
한국전자기술연구원	신희동	경기도 성남시 분당구 새나리로 25 (야탑동)	031-789-7740	http://www.keti.re.kr
한국전자통신연구원	방승찬	대전시 유성구 가정로 218	042-860-6114	http://www.etri.re.kr
한화시스템㈜	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㈜	정의선, 이규석	서울시 강남구 테헤란로 203	02-2018-5114	http://www.mobis.co.kr
현대자동차㈜	정의선, 장재훈, 이동석	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	생년월일	년 월 일
	취득년월	년 월	지도교수	
현 근무처 (또는 연락처)	주 소			(우편번호 :)
	전화번호		FAX번호	
학위논문 제목	국 문			
	영 문			
KEY WORD				

국문 초록(요약) : 1000자 이내

보내실 곳 _ 06130
서울특별시 강남구 테헤란로 7길 22(역삼동, 과학기술회관 제1관 907호)
사무국 회지담당자앞
E-mail : ieie@theieie.org
TEL : (02)553-0255(내선 6번) FAX : (02)562-4753

대한전자공학회
IEIE The Institute of Electronics and Information Engineers

전자공학회지 〈월간〉 제51권 제5호(통권 제480호) The Magazine of the IEIE

2024년 5월 20일 인쇄
2024년 5월 25일 발행

발행및 편집인 **(사) 대한전자공학회** **회장 이 충 용**
인쇄인 **한림원(주)** **대표 김 흥 중**
발행인 **사단법인 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)562-4753
E-mail : ieie@theieie.org
Homepage : http://www.theieie.org

2024년도 회비납부 안내

1. 회비의 납부 및 유효기간

2024년도 회원 연회비는 2023년과 동일함을 알려드리며, 아직 2024년도 회비를 납부하지 않으신 회원님께서는 납부하여 주시기 바라며, 연회비의 유효기간은 회비를 납부한 당해연도에 한합니다.

◆ 2024년도 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 : 2024년 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

위의 **QR코드**를 찍어보세요
SK하이닉스가 만드는 **미래 AI세상**이 펼쳐집니다

AI와 미래를 잇는 코드
SK하이닉스로부터

미래 AI 인프라의 핵심, 글로벌 No.1 AI 메모리 기업

We Do Technology | SK hynix