ISSN 1016-9288
제52권 2호
2025년 2월호

전자공학회지

The Magazine of the IEIE

vol.52. no.2

인공지능 · 신호처리 효율화 및 응용 기술

- 지능형 CCTV AI 신호처리 응용기술
- YOLO 아키텍처 기반의 자기증류를 활용한 객체 탐지 성능 향상
- AI의 메타인지 능력 개발
- Scientific LLMOps: 과학기술 분야 AI 개발의 중추
- 영상 초해상화 인공지능 모델의 양자화 기술
- 구조적 관점에서의 영상 초해상화 인공지능 모델의 경량화 기법 분석

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

AI-driven Latex 기반 논문 작성협업 솔루션 murfy (머피)를 소개합니다.

사용자는 Latex에 대한 지식이 없더라도 AI chat에게 테이블, 수식, 이미지, 알고리즘 등의 Latex 코드 작성 및 관련 참고 논문 추천 등 업무를 요청하면 됩니다. 머피를 사용하면 한 곳에서 논문 작성에 필요한 모든 과정들을 해결할 수 있기 때문에 논문 작성 소요시간을 대폭 줄여 연구에 더 몰두할 수 있습니다.

머피 웹사이트 | www.murfy.ai

매일경제

논문 작성시간 80% 줄여드려요

머피는 논문 작성에 특화된 대규모 언어모델로, 논문에 쓰였던 표현을 보다 더 적확한 표현으로 교정해준다. 또 문맥에 맞는 인용문을 자동 추천하는 등...

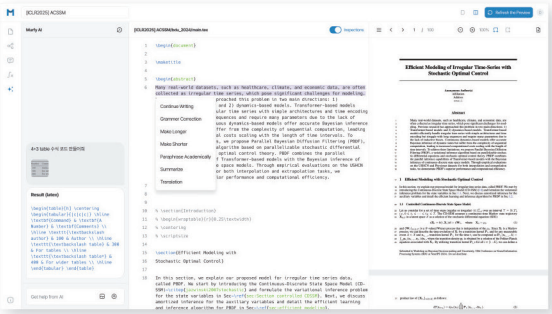
NATE

[Pangyo Interview] 머플, AI 기반 논문 작성 혁신 플랫폼 '머피'...학술 논문 작성의 새로운 기준 제시하는 게임 체인저

한눈에 보는 오늘 : 홈 - 뉴스 : 안수남 머플 대표 | 촬영-에이빙뉴스머플(Murple, 대표 안수남)이 연구자와 학생들을 위한 학술 논문 작성 방식을...

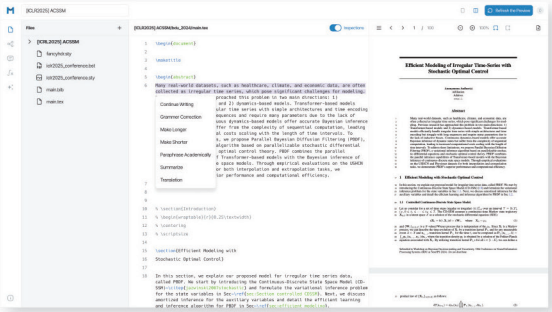
구글 검색창에 '논문작성' 혹은 '머플' 검색 시 나오는 보도자료 입니다.

한 곳에서 논문작성에 필요한 모든 과정들을 해결할 수 있고 논문 작성 소요시간을 대폭 줄여줍니다.



Latex를 몰라도 Murfy.AI가 원하는 코드를 대신 작성해드립니다.

테이블, 수식, 알고리즘의 Latex 작성하느라 힘드셨죠? 참고논문에서 원하는 테이블, 수식, 알고리즘 등을 화면 캡처해서 Murfy.AI 한테 요청하거나 프롬프트로 직접 요청하면 3초만에 원하는 Latex 코드를 생성해줍니다.

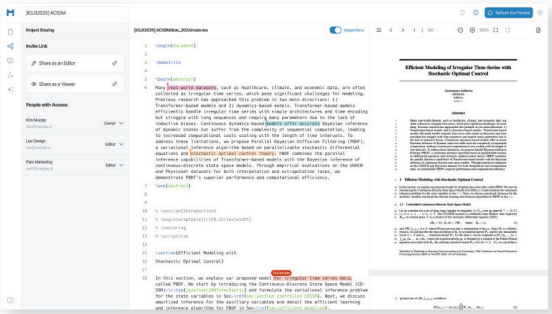


논문에 최적화 된 AI가 교정과 문법 오류 검수를 완벽하게 지원합니다.

480만 편의 논문 데이터를 바탕으로 학습된 AI 기능으로 문장 교정부터 문법 오류 검수, 영문 번역, 늘려쓰기, 간결하게 쓰기, 요약하기 와 Latex 문법 오류 검수 및 해결책 제시까지 완벽하게지원합니다.

자동 포맷 기능을 지원하며 실시간으로 출력본을 확인할 수 있습니다.

기존의 Word, 한글에서 직접 포맷팅 해야하는 번거로움 없이 자동 포맷을 제공 및 실시간으로 PDF(렌더링) 출력본을 스플릿 뷰로 확인 가능합니다.



수식 터보 기능을 제공하여 Latex 수식을 10배 빠르게 작성할 수 있습니다.

수식 터보 가이드에서 단축코드 사용 가이드를 제공하며 기존의 수식에서 가결한 수식 사용으로작성시간을 단축시킬 수 있습니다.

머피와 함께하는 기업과 대학원들을 소개합니다.

KAIST · POSTECH · 서울대학교 SEOUL NATIONAL UNIVERSITY · 중앙대학교 · NAVER · 고려대학교 KOREA UNIVERSITY · 광주과학기술원 Gwangju Institute of Science and Technology · 성균관대학교 SUNGKYUNKWAN UNIVERSITY(SKKU) · 연세대학교 YONSEI UNIVERSITY · Google

대표자 안수남

주소 경기 성남시 분당구 성남대로331번길 8, 킨스타워 20층 머플 주식회사

연락처 010 5379 3771

murfy

큐알코드(www.murfy.ai) 인식 후 구글 계정으로 로그인해주세요.

www.murfy.ai

머피 솔루션 도입 문의 혹은 무료 버전 체험 신청을 원하시면,

ethan.an@murple.ai로 문의 내용을 작성하여 보내주세요.

The Institute of Electronics and Information Engineers
Summer Annual Conference of IEIE, 2025

2025년도 하계종합 학술대회

2025년 6월 25일(수) ~ 27일(금) | 롯데호텔 제주 (중문)

논문모집 | 홈페이지(http://conf.theieie.org)

최우수/일반 논문

- 논문제출 : 2025년 5월 2일(금)
- 심사통보 : 2025년 5월 16일(금)
- 사전등록 : 2025년 6월 1일(일)

대 회 장	백광현 (중앙대학교)
담당부회장	김 훈 (인천대학교)
조직위원장	강명곤 (서울시립대학교)
TPC 위원장	김 현 (서울과학기술대학교)

발표분야 (학회 6개 소사이어티 + New Emerging Area)

발표분야(학회 6개 소사이어티)	연 구 회
통신 (Communication)	통신, 미래지능형네트워크, 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, 이미지센서
컴퓨터 (Computer)	융합컴퓨팅, 멀티미디어, 인공지능 신경망 및 퍼지시스템, M2M/IoT, 휴먼ICT, 인공지능 및 보안, 증강휴먼, AI응용, 블록체인
인공지능 신호처리 (AI Signal Processing)	영상처리, 음향 및 음성신호처리, 영상이해, 바이오영상신호처리, 딥러닝, 로봇지능
시스템 및 제어 (System and Control)	의용전자 및 생체공학, 제어계측, 회로 및 시스템, 전력전자, 지능로봇, 국방정보 및 제어, 자동차전자, 의료영상시스템, 스마트팩토리, 스마트 미터링
산업전자 (Industry Electronics)	산업전자제어, 임베디드시스템, 유비쿼터스 센서네트워크, 디지털통신시스템
New Emerging Area	의료, 에너지, Software, 바이오메디컬 연구회, 기타

행사문의 학회 사무국

- 전 화 (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

ITC-CSCC 2025

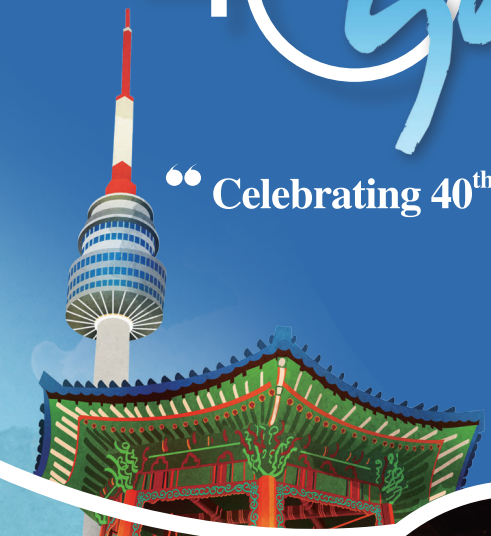
The 40th International Technical Conference on Circuits/ Systems, Computers, and Communications 2025

July 7 Mon – 10 Thu, 2025

Chung-Ang University, Seoul, South Korea

40th Seoul

"Celebrating 40th Anniversary!"



TOPICS

The conference is open to researchers from all regions of the world. Participation from Asia Pacific region is particularly encouraged. Proposals for special sessions are welcome. Papers with original works in all aspects of Circuits/Systems, Computers and Communications are invited. Topics include, but not limited to, the followings:

Circuits & Systems

- Semiconductor Devices & Technology
- Computer Aided Design
- Power Electronics & Circuits
- Intelligent Transportation Systems & Technology
- Analog Circuits
- RF Circuits
- Linear / Nonlinear Systems
- Modern Control
- Medical Electronics & Circuits
- Neural Networks
- VLSI Design
- Verification & Testing
- Sensors & Related Circuits

Computers

- Artificial Neural Network
- Biocomputing
- Computer Systems & Applications
- Human-Computer Interaction
- Internet Technology & Applications
- Metaverse
- Multimedia Service & Technology
- Security
- Watermarking

Image/Signal Processing

- Artificial Intelligence
- Computer Vision
- Image Coding & Analysis
- Image Detection & Recognition
- Image Processing
- Motion Analysis
- Natural Language Processing
- Object Extraction & Technology

Communications

- Signal Processing
- Antenna & Wave Propagation
- Network Management & Design
- Optical Communications & Components
- Circuits & Components for Communications
- Radar / Remote Sensing
- IP Networks & QoS
- MIMO & Space-Time Codes
- Ubiquitous Networks
- Multimedia Communications
- UWB
- Mobile & Wireless Communications
- Visual Communications
- Future Internet Architectures

Important Date

February	March 31, 2025	May 15, 2025	May 31, 2025	July 7-10, 2025
Full-Paper Submission Begin	Full-Paper Submission Deadline	Notification of Paper Acceptance	Camera-Ready Paper Submission Deadline	Conference Date

SUBMISSION OF PAPERS

Prospective authors are invited to submit original papers (1~6 pages) of either MS Word or PDF format written in English. Paper submission procedures are available at www.itc-cscc2025.org.

Contact Point

• E-mail : inter@theieie.org • Phone : +82-2-553-0255(Ext. 4) • www.itc-cscc2025.org.

IEIE IEEE

www.theieie.org / www.kiee.or.kr

ICS 2025
정보 및 제어 심포지엄

2025 Information and Control Symposium

2025. 4. 23(수) ~ 25(금)

전북대학교 국제컨벤션센터

대한전자공학회와 대한전기학회가 공동으로 주관하는 2025년도 정보 및 제어 심포지엄 (ICS 2025)이 2025년 4월 23일(수)~25일(금)에 개최됩니다. 정보 및 제어 분야의 최신 연구 결과가 발표되는 본 학술회의는 정보 제어분야의 특별강연과 초청논문을 포함한 논문발표 등 다양한 학술행사로 구성되며, 학계 및 산업계의 전문가들의 교류의 장이 될 것입니다. 정보 및 제어 분야의 현재를 진단하고 미래를 전망할 수 있는 ICS 2025에 여러분을 초대합니다.

대한전자공학회 시스템 및 제어 소사이어티 회장 **이 덕 진**
대한전기학회 정보 및 제어 부문회 회장 **정 종 진**

주요일정

- 논문제출마감 : **2025. 3. 21(금)**
- 논문심사통보 : **2025. 3. 31(월)**
- 사전등록마감 : **2025. 4. 11(금)**
- 제출처 : **http://ics.theieie.org**

문의처

대한전자공학회 사무국 배지영 부장
전화 : 02-553-0255(내선 2번)
이메일 : conf@theieie.or.kr

논문발표분야

- Aerospace Control & Application
- AI Autonomous Manufacturing
- Artificial Intelligence
- Audio, Image, Video Signal Processing
- Automotive Electronics
- Biomedical Systems, Bio Infomax
- Communication and Signal Processing
- Control Systems
- Embedded System
- Human-Computer Interaction
- Industrial Artificial Intelligence
- Industrial Digital Transformation
- Intelligent Control
- Mechatronics
- Mobile Systems
- Navigation
- Pattern Recognition
- Power Electronics
- Robotics
- Security & Content Protection
- Smart Car
- Sensor and Measurement
- Ship Information Technology
- Telematics
- Ubiquitous Network

대한전자공학회 시스템 및 제어 소사이어티 IEIE

대한전기학회 정보 및 제어 부문회
THE KOREAN INSTITUTE OF ELECTRICAL ENGINEERS

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

SAE MOBILUS™ 에서 세계 최고 Automotive Engineering 정보를 만나보세요.

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

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

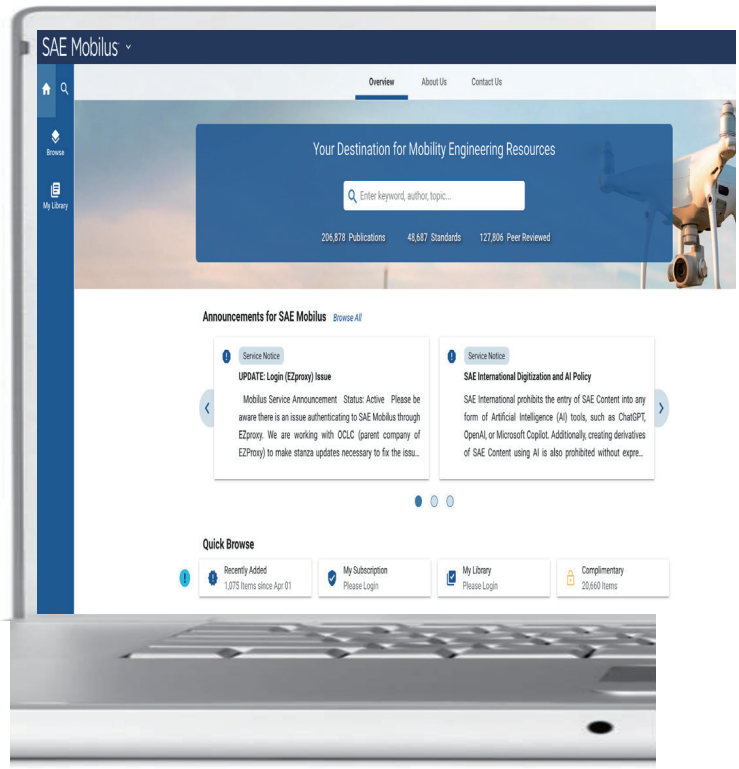
SAE MOBILUS™에서 다음과 같은 다양한 주제 분야에 대한 자료를 확인할 수 있습니다.

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

SAE MOBILOUS™ 에서 항공 우주 분야 R&D에 필요한 표준 및 규격 그리고 다양한 자료를 검색 및 이용할 수 있습니다.

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

http://saemobilus.sae.org



SAE INTERNATIONAL

Authorized Dealer in Korea

키티스 産學研情報(株)
KITIS Info. & Co., Ltd.
Tel. 02.3474.5290 Web. www.kitis.co.kr

IEEE VISION | INNOVATION CHALLENGES SUMMIT
An IEEE Summit and Honors Ceremony

2025 VIC Summit & Honors Ceremony Gala

글로벌 공학 및 기술 혁신을 만나는 자리, 2025년 IEEE VIC Summit & Honors Ceremony Gala가 일본 도쿄에서 개최됩니다!

기술, 공학, 과학 분야의 성과를 기리는 이번 행사에서는
IEEE Medal of Honor를 포함한 다양한 시상이 이루어지며,
최신 기술 동향 공유와 글로벌 네트워킹의 기회를 제공합니다.
2025년 IEEE 어워즈는 기술 발전과 협력의 가치를 기념하며,
업계 전문가들이 교류하고 영감을 얻는 자리가 될 것입니다.
KITIS 고객 여러분을 혁신의 현장에 초대합니다.

- 일자: 2025년 4월 23일(수) - 4월24일(목)
- 장소: 일본 힐튼 도쿄 오다이바 / Hilton Tokyo Odaiba (주소: 1 Chome-9-1 Daiba, Minato City, Tokyo 135-8625, Japan / Phone: +81 3-5500-5500)
- 주최 및 후원: IEEE & KITIS

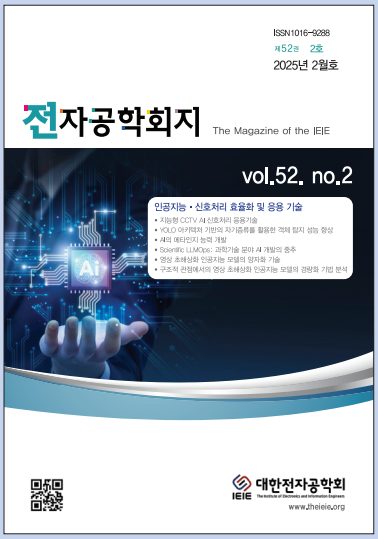
*참석을 원하시는 엔지니어, 교수님, 연구원께서는 당사 이희진 과장(hjlee@kitis.co.kr/02-3474-5290)에게 연락주시면, IEEE와 협의 후 참석가능 여부를 안내해 드리겠습니다.

IEEE | Authorized Dealer

키티스 産學研情報(株)
KITIS Info. & Co., Ltd.

CONTENTS

제52권 2호 (2025년 2월)



※ 학회지 2월호 표지 (vol 52. No 2)

- 사무국 편집담당
  이 안 순 부장
  TEL : (02)553-0255(내선 6)
  FAX : (02)562-4753
- 학회 홈페이지
  http://www.theieie.org

학회소식

특집 : 인공지능 · 신호처리 효율화 및 응용 기술

회원광장

정보교차로

대한전자공학회
IEIE The Institute of Electronics and Information Engineers

2025년도 임원 및 각 위원회 위원

회 장	백 광 현 (중앙대학교 교수)
수석부회장	김 종 옥 (고려대학교 교수) – 총괄 / ITC-CSCC
고 문	권 오 경 (한양대학교 석좌교수)
	김 진 상 (경희대학교 총장)
	방 승 찬 (한국전자통신연구원 원장)
	신 희 동 (한국전자기술연구원 원장)
	오 상 록 (한국과학기술연구원 원장)
	천 경 준 (씨젠 회장)
감 사	이 재 훈 (유정시스템 대표이사)
부 회 장	강 문 식 (강릉원주대학교 교수) – 지부 / 교육 총괄 / 표준화
	권 호 열 (강원대학교 교수) – 대외협력 / 기술정책
	김 훈 (인천대학교 교수) – 하계 / 추계 / 국문논문
	노 태 문 (한국전자통신연구원 연구전문위원) – 산학연
	송 병 철 (인하대학교 교수) – ICEIC / ITC-CSCC / ICCE-Asia / SPC
	이 승 호 (한밭대학교 교수) – 학술대회
	정 영 모 (한성대학교 교수) – 홍보 · 정보화 / 회지편집
소사이어티 회장	고 병 철 (계명대학교 교수) – 인공지능신호처리소사이어티
	신 오 순 (숭실대학교 교수) – 통신소사이어티
	이 덕 진 (전북대학교 교수) – 시스템및제어소사이어티
협동부회장	강 성 원 (한국전자통신연구원 부원장)
	김 동 식 (인하공업전문대학 교수)
	노 승 원 (LG이노텍 CTO)
	심 동 규 (광운대학교 교수)
	오 윤 제 (정보통신기획평가원 PM)
	유 창 동 (KAIST 교수)
	이 병 선 (김포대학교 교수)
	이 서 규 (픽셀플러스 대표이사)
	이 장 규 (텔레칩스 대표이사)
	전 선 익 (파이낸셜뉴스 부회장)
	조 진 우 (한국전자기술연구원 부원장)
상 임 이 사	강 명 곤 (서울시립대학교 교수) – 하계 총괄 1
	강 태 욱 (성균관대학교 교수) – 사업
	구 민 석 (서울시립대학교 교수) – 국제협력
	김 윤 (서울시립대학교 교수) – 총무
	김 현 (서울과학기술대학교 교수) – 하계 총괄 2
	김 민 휘 (중앙대학교 교수) – 차세대 리더육성
	김 수 연 (동국대학교 교수) – 회원, 여성
	김 용 신 (고려대학교 교수) – 홍보 · 정보화 총괄 / 신기술
	김 원 준 (건국대학교 교수) – SPC
	김 태 인 (인하대학교 교수) – 회원 · 여성
	동 성 수 (용인예술과학대학교 교수) – 교육
	변 대 석 (삼성전자 마스터) – 추계 총괄 2
	서 병 석 (상지대학교 교수) – 추계
	손 교 민 (삼성전자 마스터) – 사업
	신 창 환 (고려대학교 교수) – 하계
	유 찬 세 (한국전자기술연구원 수석연구원) – 사업 총괄
	이 형 민 (고려대학교 교수) – 추계
	임 홍 기 (인하대학교 교수) – ICCE-Asia
	전 세 영 (서울대학교 교수) – 기획 · 대외협력
	정 일 권 (한국전자통신연구원 본부장) – 사업
	조 성 현 (한양대학교 교수) – 추계 총괄 1
	최 우 영 (서울대학교 교수) – JSTS
	추 민 성 (한양대학교 교수) – 차세대 리더육성
	황 원 준 (아주대학교 교수) – 회지편집 총괄
산업체이사	고 용 남 (하나마이크론 CTO)
	김 현 수 (삼성전자 상무)
	우 정 호 (비전넥스트 대표이사)
	이 도 훈 (국가보안기술연구소 연구위원)
	김 기 남 (한국공학한림원 회장)
	박 성 욱 (차세대지능형반도체사업단 이사장)
	선우명훈 (아주대학교 교수 / IEEE CASS 회장)
	안 승 권 (연암공과대학교 총장)
	전 영 현 (삼성전자 부회장)
	최 창 식 (DB하이텍 부회장)
	전 병 우 (성균관대학교 교수)
	구 본 태 (한국전자통신연구원 본부장) – 산학연
	김 병 훈 (LG전자 부사장/CTO) – 산업체
	노 원 우 (연세대학교 교수) – 신기술 / 차세대 리더육성
	류 수 정 (서울대학교 교수) – 회원 · 여성 총괄
	이 강 윤 (성균관대학교 교수) – 사업 / 기술워크숍 / 대외협력
	인 치 호 (세명대학교 교수) – 대외협력
	황 인 철 (강원대학교 교수) – 소사이어티 네트워크 협력
	고 정 환 (인하공업전문대학 교수) – 산업전자소사이어티
	이 덕 기 (연암공과대학교 교수) – 컴퓨터소사이어티
	장 성 진 (삼성전자 상근고문) – 반도체소사이어티
	김 강 태 (삼성전자 부사장)
	김 형 준 (한국과학기술연구원 소장)
	배 순 민 (케이티 소장)
	엄 낙 웅 (한국전자통신연구원 연구전문위원)
	원 제 형 (도쿄일렉트론코리아 대표이사)
	이 광 엽 (서경대학교 교수)
	이 상 훈 (웨이브피아 대표이사)
	이 용 욱 (한화시스템 부사장)
	이 재 관 (한국자동차연구원 소장)
	정 준 (쏠리드 대표이사)
	최 병 호 (한국전자기술연구원 소장)
	강 석 주 (서강대학교 교수) – 기획 · 대외협력 총괄 / ITC-CSCC
	고 형 호 (충남대학교 교수) – 기획 · 대외협력
	권 혁 인 (중앙대학교 교수) – 총무
	김 혁 (서울시립대학교 교수) – 교육
	김 동 현 (연세대학교 교수) – 바이오메티컬연구회
	김 성 우 (서울대학교 교수) – 신기술 / 홍보 · 정보화
	김 영 로 (명지전문대학 교수) – 사업
	김 원 종 (한국전자통신연구원 책임연구원) – 표준화
	김 익 균 (한국전자통신연구원 부장) – 하계
	김 학 구 (중앙대학교 교수) – 추계
	박 철 수 (광운대학교 교수) – SPC 총괄
	변 영 재 (UNIST 교수) – 사업
	선 우 경 (서울대학교 교수) – 회지편집
	송 준 영 (인천대학교 교수) – 하계
	연 규 봉 (한국자동차연구원 수석연구원) – 하계
	이 채 은 (한양대학교 교수) – 사업
	이 후 진 (한성대학교 교수) – 논문편집
	장 익 준 (경희대학교 교수) – 논문편집 총괄
	정 성 엽 (고려대학교 교수) – 차세대 리더육성
	조 성 재 (이화여자대학교 교수) – 신기술 총괄 / 홍보 · 정보화
	최 강 선 (한국기술교육대학교 교수) – 논문편집
	최 재 혁 (서울대학교 교수) – ICCE-Asia 총괄
	한 재 호 (고려대학교 교수) – 재무
	김 동 현 (ICTK 대표이사)
	오 의 열 (LG디스플레이 연구위원)
	윤 영 권 (삼성전자 마스터)
	이 수 민 (한국센서연구소 대표이사)

	이 수 인 (텔레칩스 상무)
	조 영 민 (SkyMirr CEO)
	최 성 민 (해치텍 대표이사)
	함 철 희 (삼성전자 마스터)
이 사	궁 재 하 (고려대학교 교수) – 홍보 · 정보화
	권 기 룡 (부경대학교 교수) – 산학연
	김 동 순 (세종대학교 교수) – 하계
	김 병 수 (한국전자기술연구원 센터장) – 하계 / ICCE-Asia
	김 선 욱 (고려대학교 교수) – 교육
	김 소 영 (성균관대학교 교수) – JSTS
	김 재 욱 (한국과학기술연구원 선임연구원) – 하계
	김 지 훈 (이화여자대학교 교수) – ICCE-Asia
	류 현 석 (서울대학교 교수) – 교육
	박 재 영 (광운대학교 교수) – 표준화
	서 민 재 (서울시립대학교 교수) – 기획 / 차세대리더 / 하계 / ICCE-Asia
	안 광 호 (한국전자기술연구원 본부장) – 대외협력
	안 성 수 (명지전문대학 교수) – 회원
	양 현 종 (서울대학교 교수) – ICCE-Asia
	유 담 (서울대학교 교수) – 하계
	윤 명 국 (이화여자대학교 교수) – 신기술
	윤 종 윤 (파두 사장) – 교육
	이 상 만 (고려대학교 교수) – 산학연
	이 철 호 (서울대학교 교수) – 신기술
	임 기 택 (한국산업기술기획평가원 PD) – 대외협력
	정 해 준 (경희대학교 교수) – 국문논문
	제 민 규 (KAIST 교수) – 사업
	채 관 엽 (삼성전자 마스터) – 대외협력
	최 재 용 (가천대학교 교수) – 국문논문
	하 정 우 (네이버 센터장) – 신기술
	한 동 윤 (네이버 리더) – 학회지
	허 용 석 (아주대학교 교수) – 학회지
	황 진 영 (한국항공대학교 교수) – 국제협력
협 동 이 사	권 경 하 (KAIST 교수) – 사업
	김 가 인 (DGIST 교수) – ICCE-Asia / 차세대리더
	김 기 현 (전북대학교 교수) – 전북지부
	김 대 영 (순천향대학교 교수) – 호서지부
	김 성 준 (동국대학교 교수) – 국제협력
	김 영 성 (인하대학교 교수) – 학회지
	김 형 진 (한양대학교 교수) – 하계 / ITC-CSCC
	민 경 식 (국민대학교 교수) – 하계 / JSTS
	박 익 준 (중부대학교 교수) – 회원/여성
	박 혜 림 (서울과학기술대학교 교수) – 회원/여성
	배 준 호 (가천대학교 교수) – 표준화
	송 익 현 (한양대학교 교수) – 국제협력
	신 세 운 (UNIST 교수) – 사업
	심 원 보 (서울과학기술대학교 교수) – 하계
	유 동 훈 (디사일로 CRO) – 회원 · 여성
	이 권 형 (LG전자 책임연구원) – 하계
	이 상 현 (아주대학교 교수) – 홍보 · 정보화
	이 재 규 (삼성전자 마스터) – 산학연
	이 철 (동국대학교 교수) – 학회지
	임 영 현 (경희대학교 교수) – 국문논문
	정 소 이 (아주대학교 교수) – 하계
	정 이 품 (연세대학교 교수) – 차세대 리더
	진 경 환 (고려대학교 교수) – 기획
	채 주 형 (광운대학교 교수) – ICCE-Asia / 차세대리더
	함 범 섭 (연세대학교 교수) – 추계
	정 민 수 (라온텍 CTO)
	천 이 우 (넥스트칩 연구소장)
	한 은 혜 (에스에스앤씨 대표이사)
	홍 국 태 (LX세미콘 연구위원)
	권 구 덕 (강원대학교 교수) – 기획
	권 기 원 (성균관대학교 교수) – 표준화
	김 민 규 (LG이노텍 상무) – 산학연
	김 상 완 (서강대학교 교수) – 기획
	김 성 동 (서울과학기술대학교 교수) – 표준화
	김 유 철 (LG AI연구원 부문장) – 신기술
	김 중 헌 (고려대학교 교수) – 사업
	류 은 석 (성균관대학교 교수) – 사업
	박 동 욱 (서울시립대학교 교수) – 하계
	배 현 철 (한국전자통신연구원 책임연구원) – 산학연
	심 용 (중앙대학교 교수) – 국제협력
	안 상 철 (한국과학기술연구원 책임연구원) – 신기술
	안 호 균 (한국전자통신연구원 책임연구원) – 사업
	오 윤 호 (고려대학교 교수) – 신기술
	유 현 우 (성균관대학교 교수) – 홍보 · 정보화
	윤 상 훈 (한국산업기술기획평가원 PD) – 산학연
	이 구 순 (파이낸셜뉴스 이사) – 홍보 · 정보화
	이 정 원 (서울대학교 교수) – 회원/여성
	이 태 동 (국제대학교 교수) – 홍보 · 정보화
	전 동 석 (서울대학교 교수) – 하계 / 국문논문
	정 희 철 (경북대학교 교수) – 학회지
	조 성 인 (동국대학교 교수) – 하계
	최 광 성 (한국전자통신연구원 실장) – 대외협력
	최 정 욱 (한양대학교 교수) – 국문논문
	하 태 준 (광운대학교 교수) – 추계
	한 태 희 (성균관대학교 교수) – 국문논문
	허 재 두 (한국전자통신연구원 연구전문위원) – 사업
	황 태 호 (한국전자기술연구원 본부장) – 하계
	권 종 원 (한국산업기술시험원 센터장) – 추계
	김 건 우 (목포대학교 교수) – 광주.전남지부
	김 다 완 (한국교통대학교 교수)
	김 도 훈 (한국전자통신연구원 박사) – 표준화
	김 승 환 (한국전자통신연구원 책임연구원) – 산학연
	김 정 석 (가천대학교 교수) – 홍보 · 정보화
	류 성 주 (서강대학교 교수) – 하계
	박 관 서 (연세대학교 교수) – ICCE-Asia / 차세대리더 / 기획
	박 정 환 (경희대학교 교수) – 국문논문
	배 준 성 (강원대학교 교수) – 국제협력
	백 지 선 (부산대학교 교수) – ICCE-Asia
	송 철 (DGIST 교수) – 바이오메디컬연구회
	신 재 영 (한국교통대학교 교수) – 교육
	유 경 창 (삼성전자 수석연구원) – 회원
	이 규 호 (UNIST 교수) – 차세대 리더 / 하계
	이 명 재 (연세대학교 교수) – 신기술
	이 신 형 (서울시립대학교 교수) – 회원 · 여성
	이 정 석 (인하공업전문대학 교수) – 학회지
	임 매 순 (한국과학기술연구원 책임연구원) – 사업
	정 동 윤 (한국전자통신연구원 센터장) – 산학연
	정 완 영 (KAIST 교수) – 홍보 · 정보화
	정 한 울 (광운대학교 교수) – 신기술
	차 철 웅 (한국전자기술연구원 센터장) – 표준화
	최 욱 (인천대학교 교수) – 학회지
	홍 성 완 (서강대학교 교수) – 하계 / 대외협력

지부장 명단

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

위원회 명단

자문위원회

위 원 장	박 성 한 (한양대학교 명예교수)		
부 위 원 장	이 진 구 (동국대학교 명예교수)		
위 원	고 성 제 (고려대학교 교수)	공 준 진 (웨이브피아 CSO)	구 용 서 (단국대학교 석좌교수)
	김 도 현 (국민대학교 명예교수)	김 성 대 (KAIST 명예교수)	김 수 중 (경북대학교 명예교수)
	김 재 희 (연세대학교 명예교수)	나 정 웅 (KAIST 명예교수)	문 영 식 (한림성심대학교 총장)
	박 규 태 (연세대학교 명예교수)	박 진 옥 (육군사관학교 명예교수)	박 항 구 (소암시스텔 회장)
	백 준 기 (중앙대학교 교수)	서 승 우 (서울대학교 교수)	성 굉 모 (서울대학교 명예교수)
	이 문 기 (연세대학교 명예교수)	이 상 설 (한양대학교 명예교수)	이 재 홍 (서울대학교 명예교수)
	이 혁 재 (서울대학교 교수)	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 석 주 (서강대학교 교수)		
부 위 원 장	전 세 영 (서울대학교 교수)	고 형 호 (충남대학교 교수)	
위 원	권 구 덕 (강원대학교 교수)	김 상 완 (서강대학교 교수)	박 관 서 (연세대학교 교수)
	서 민 재 (서울시립대학교 교수)	진 경 환 (고려대학교 교수)	

학술연구위원회 – 하계

위 원 장	강 명 곤 (서울시립대학교 교수)		
TPC위원장	김 현 (서울과학기술대학교 교수)		
위 원	김 동 순 (세종대학교 교수)	김 병 수 (한국전자기술연구원 센터장)	김 익 균 (한국전자통신연구원 부장)
	김 형 진 (한양대학교 교수)	민 경 식 (국민대학교 교수)	박 동 욱 (서울시립대학교 교수)
	서 경 원 (서울과학기술대학교 교수)	서 민 재 (서울시립대학교 교수)	서 승 호 (고려대학교 교수)
	송 준 영 (인천대학교 교수)	신 창 환 (고려대학교 교수)	심 원 보 (서울과학기술대학교 교수)
	연 규 봉 (한국자동차연구원 수석연구원)	유 담 (서울대학교 교수)	이 규 호 (연세대학교 교수)
	이 권 형(LG전자 책임연구원)	이 승 호 (한밭대학교 교수)	인 치 호 (세명대학교 교수)
	정 소 이 (아주대학교 교수)	조 성 인 (동국대학교 교수)	최 승 규 (경희대학교 교수)
	홍 성 완 (서강대학교 교수)		

학술연구위원회 – 추계

위 원 장	조 성 현 (한양대학교 교수)		
TPC위원장	변 대 석 (삼성전자 마스터)		
위 원	강 문 식 (강릉원주대학교 교수)	구 본 태 (한국전자통신연구원 본부장)	권 종 원 (한국산업기술시험원 센터장)
	김 중 헌 (고려대학교 교수)	김 학 구 (중앙대학교 교수)	노 태 문 (한국전자통신연구원 연구전문위원)
	서 병 석 (상지대학교 교수)	이 승 호 (한밭대학교 교수)	이 형 민 (고려대학교 교수)
	추 민 성 (한양대학교 교수)	하 태 준 (광운대학교 교수)	함 범 섭 (연세대학교 교수)
	황 인 철 (강원대학교 교수)		

논문편집위원회

위 원 장	장 익 준 (경희대학교 교수)		
위 원	박 정 환 (경희대학교 교수)	이 후 진 (한성대학교 교수)	임 영 현 (경희대학교 교수)
	전 동 석 (서울대학교 교수)	정 해 준 (경희대학교 교수)	최 강 선 (한국기술교육대학교 교수)
	최 재 용 (가천대학교 교수)	최 정 욱 (한양대학교 교수)	한 태 희 (성균관대학교 교수)

국제협력위원회

위 원 장	구 민 석 (서울시립대학교 교수)		
위 원	김 성 준 (동국대학교 교수)	배 준 성 (강원대학교 교수)	송 익 현 (한양대학교 교수)
	심 용 (중앙대학교 교수)	황 진 영 (한국항공대학교 교수)	

회원관리위원회

위 원 장	류 수 정 (서울대학교 교수)		
위 원	김 수 연 (동국대학교 교수)	김 태 인 (인하대학교 교수)	박 익 준 (중부대학교 교수)
	박 혜 림 (서울과학기술대학교 교수)	안 성 수 (명지전문대학 교수)	유 경 창 (삼성전자 수석연구원)
	유 동 훈 (디사일로 CRO)	이 신 형 (서울시립대학교 교수)	이 정 아 (AURA CEO)
	이 정 원 (서울대학교 교수)	임 정 연 (SK텔레콤 담당)	

회지편집위원회

위 원 장	황 원 준 (아주대학교 교수)		
부 위 원 장	선 우 경 (서울대학교 교수)		
위 원	김 영 성 (인하대학교 교수)	이 철 (동국대학교 교수)	정 희 철 (경북대학교 교수)
	최 욱 (인천대학교 교수)	한 동 윤 (네이버 리더)	허 용 석 (아주대학교 교수)

사업위원회

위 원 장	유 찬 세 (한국전자기술연구원 수석연구원)		
위 원	감 익 균 (한국전자통신연구원 본부장)	강 태 욱 (성균관대학교 교수)	권 경 하 (KAIST 교수)
	김 영 로 (명지전문대학 교수)	김 중 헌 (고려대학교 교수)	류 은 석 (성균관대학교 교수)
	변 영 재 (UNIST 교수)	선우명훈 (아주대학교 교수)	손 교 민 (삼성전자 마스터)
	신 세 운 (UNIST 교수)	안 호 균 (한국전자통신연구원 책임연구원)	이 채 은 (한양대학교 교수)
	임 매 순 (한국과학기술연구원 책임연구원)	정 일 권 (한국전자통신연구원 본부장)	제 민 규 (KAIST 교수)
	허 재 두 (한국전자통신연구원 연구전문위원)		

교육연구위원회

위 원 장	강 문 식 (강릉원주대학교 교수)		
위 원	김 혁 (서울시립대학교 교수)	동 성 수 (용인예술과학대학교 교수)	류 현 석 (서울대학교 교수)
	박 영 우 (TEL 부사장)	박 혜 림 (서울과학기술대학교 교수)	변 대 석 (삼성전자 교수)
	변 영 재 (UNIST 교수)	신 재 영(한국교통대학교 교수)	윤 종 윤 (Fadu 사장)
	이 신 형 (서울시립대학교 교수)	이 영 택 (ASML 전무)	이 후 진 (한성대학교 교수)
	정 성 엽 (고려대학교 교수)	정 용 규 (을지대학교 교수)	

홍보/정보화위원회

위 원 장	김 용 신 (고려대학교 교수)		
위 원	궁 재 하 (고려대학교 교수)	김 성 우 (서울대학교 교수)	김 정 석 (가천대학교 교수)
	유 현 우 (성균관대학교 교수)	이 상 현 (아주대학교 교수)	

표준화위원회

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

신기술위원회

위 원 장	노 원 우 (연세대학교 교수)		
부 위 원 장	김 용 신 (고려대학교 교수)	조 성 재 (이화여자대학교 교수)	
위 원	김 성 우 (서울대학교 교수)	이 철 호 (서울대학교 교수)	오 윤 호 (고려대학교 교수)
	윤 명 국 (이화여자대학교 교수)	이 명 재 (연세대학교 교수)	정 한 울 (광운대학교 교수)

선거관리위원회

위 원 장	전 홍 태 (중앙대학교 교수)		
위 원	강 명 곤 (서울시립대학교 교수)	강 석 주 (서강대학교 교수)	권 혁 인 (중앙대학교 교수)
	김 윤 (서울시립대학교 교수)	류 수 정 (서울대학교 교수)	한 재 호 (고려대학교 교수)

포상위원회

위 원 장	공 준 진 (웨이브피아 CSO)		
위 원	강 석 주 (서강대학교 교수)	강 명 곤 (서울시립대학교 교수)	김 종 옥 (고려대학교 교수)
	장 익 준 (경희대학교 교수)	한 재 호 (고려대학교 교수)	
위원 및 간사겸임	권 혁 인 (중앙대학교 교수)		

인사위원회

위 원 장	백 광 현 (중앙대학교 교수)		
위 원	김 종 옥 (고려대학교 교수)	강 석 주 (서강대학교 교수)	권 혁 인 (중앙대학교 교수)
	조 성 재 (이화여자대학교 교수)	한 재 호 (고려대학교 교수)	

차세대 리더육성위원회

위 원 장	노 원 우 (연세대학교 교수)		
부 위 원 장	김 민 휘 (중앙대학교 교수)	정 성 엽 (고려대학교 교수)	추 민 성 (한양대학교 교수)
위 원	김 가 인 (DGIST 교수)	박 관 서 (연세대학교 교수)	서 민 재 (서울시립대학교 교수)
	이 규 호 (UNIST 교수)	정 이 품 (연세대학교 교수)	채 주 형 (광운대학교 교수)

통신소사이어티

회장	신오순 (숭실대학교 교수)		
부회장	김선용 (건국대학교 교수)	김재현 (아주대학교 교수)	김진영 (광운대학교 교수)
	김 훈 (인천대학교 교수)	오정근 (ATNS 대표이사)	윤석현 (단국대학교 교수)
	이인규 (고려대학교 교수)	이재진 (숭실대학교 교수)	이정우 (중앙대학교 교수)
	최천원 (단국대학교 교수)	허 준 (고려대학교 교수)	
감사	유명식 (숭실대학교 교수)	이흥노 (광주과학기술원 교수)	
협동부회장	김병남 (에이스테크놀로지 연구소장)	김연은 (브로던 대표이사)	김영한 (숭실대학교 교수)
	김용석 (답스 대표이사)	방승찬 (한국전자통신연구원 원장)	연철흠 (LGT 상무)
	이재훈 (유정시스템 대표이사)	장병수 (이노벨류네트웍스 부사장)	정현규 (한국전자통신연구원 부장)
	조인호 (에이스테크놀로지 박사)		
이사	김광순 (연세대학교 교수)	김성훈 (한국전자통신연구원 박사)	김정호 (이화여자대학교 교수)
	김준수 (한국공학대학교 교수)	서철헌 (숭실대학교 교수)	성원진 (서강대학교 교수)
	손인수 (동국대학교 교수)	신수용 (금오공과대학교 교수)	신요안 (숭실대학교 교수)
	윤종호 (한국항공대학교 교수)	윤지훈 (서울과학기술대학교 교수)	이예훈 (서울과학기술대학교 교수)
	이재훈 (동국대학교 교수)	이종호 (숭실대학교 교수)	이호경 (홍익대학교 교수)
	임종태 (홍익대학교 교수)	장석호 (건국대학교 교수)	조성래 (중앙대학교 교수)
	조성현 (한양대학교 교수)	최지웅 (대구경북과학기술원 교수)	최진식 (한양대학교 교수)
	허서원 (홍익대학교 교수)	홍인기 (경희대학교 교수)	황승훈 (동국대학교 교수)
	황인태 (전남대학교 교수)		
간사	김중헌 (고려대학교 교수)	정소이 (아주대학교 교수)	
연구회위원장	백상헌 (고려대학교 교수) – 통신	윤상민 (국민대학교 교수) – 지능형네트워크	
	이용식 (연세대학교 교수) – 마이크로파 및 전파전파	이철기 (아주대학교 교수) – ITS	
	김강욱 (경북대학교 교수) – 군사전자	허재두 (한국전자통신연구원 연구전문위원) – 무선 PAN/BAN	

반도체소사이어티

회장	장성진 (삼성전자 상근고문)		
자문위원	공준진 (웨이브피아 CSO, 본회명예회장)	권오경 (한양대학교 교수, 반소전임회장)	김영환 (POSTECH 교수)
	김재석 (연세대학교 교수)	김진상 (경희대학교 총장, 반소전임회장)	김희석 (청주대학교 교수)
	박홍준 (POSTECH 교수)	선우명훈 (아주대학교 교수, 반소전임회장)	소경신 (Synopsys 사장)
	손보익	신윤승 (반소전임회장)	신현철 (한양대학교 교수, 반소전임회장)
	우남성 (반소전임회장)	이승훈 (서강대학교 교수)	이윤종 (KAIST 교수)
	임신일 (서경대학교 교수)	임형규 (반소전임회장)	전영현 (삼성전자 부회장, 반소전임회장)
	정연모 (경희대학교 교수)	정항근 (전북대학교 교수)	조경순 (한국외국어대학교 교수)
	조상복 (울산대학교 교수)	조중휘 (인천대학교 교수, 반소전임회장)	최기영 (서울대학교 교수)
	최승종 (웰랑 대표)	허 염 (실리콘마이터스 회장)	
감사	이강윤 (성균관대학교 교수)	이광엽 (서경대학교 교수)	
부회장	고대협 (LX세미콘 연구소장)	김동규 (한양대학교 교수)	박영우 (EXICON 사장)
	안기현 (한국반도체산업협회 전무)	이한호 (인하대학교 교수)	이희덕 (충남대학교 교수)
	최중호 (서울시립대학교 교수, 수석부회장)	황상준 (삼성전자 부사장)	
총무이사	강명곤 (서울시립대학교 교수)	고형호 (충남대학교 교수)	김지훈 (이화여자대학교 교수)
	김 현 (서울과학기술대학교 교수)	류현석 (서울대학교 교수)	박종선 (고려대학교 교수)
	배승준 (삼성전자 부사장)	윤찬호 (삼성전자 마스터)	추민성 (한양대학교 교수)
편집이사	노정진 (한양대학교 교수)	유창식 (삼성전자 부사장)	조성재 (이화여자대학교 교수)
	한태희 (성균관대학교 교수)		
학술이사	강진구 (인하대학교 교수)	김철우 (고려대학교 교수)	범진욱 (서강대학교 교수)
	변영재 (UNIST 교수)	송민규 (동국대학교 교수)	이병훈 (POSTECH 교수)
	이승호 (한밭대학교 교수)	이혁재 (서울대학교 교수)	이희덕 (충남대학교 교수)
	인치호 (세명대학교 교수)	정진균 (전북대학교 교수)	차호영 (홍익대학교 교수)
	최우영 (연세대학교 교수)	최창환 (한양대학교 교수)	
사업이사	강운병 (삼성전자 마스터)	공배선 (성균관대학교 교수)	공정택 (성균관대학교 교수)
	김동순 (세종대학교 교수)	김소영 (성균관대학교 교수)	김시호 (연세대학교 교수)
	김용석 (성균관대학교 교수)	김원종 (한국전자통신연구원 책임연구원)	김종선 (홍익대학교 교수)
	나영선 (LX세미콘 상무)	백광현 (중앙대학교 교수)	변대석 (삼성전자 마스터)
	손교민 (삼성전자 마스터)	송용호 (삼성전자 부사장)	엄낙웅 (한국전자통신연구원 연구위원)
	오정우 (연세대학교 교수)	이상훈 (웨이브피아 대표)	조태제 (DGIST 교수)
	최규명 (서울대학교 교수)	최병호 (한국전자기술연구원 본부장)	최윤경 (고려대학교 교수)
	최준림 (경북대학교 교수)		
재무이사	권기원 (성균관대학교 교수)	이성수 (숭실대학교 교수)	
산학이사	김경수 (넥스트칩 대표)	김동현 (ICTK 사장)	김보은 (라온텍 사장)
	김준석 (ADT 사장)	나준호 (실리콘웍스 전무)	손재철 (가천대학교 교수)
	송태훈 (휴인스 사장)	신용석 (케이던스코리아 사장)	오문욱 (삼성전자 부사장)
	이도영 (옵토레인 사장)	이장규 (텔레칩스 대표)	
회원이사	노원우 (연세대학교 교수)	문 용 (숭실대학교 교수)	
연구회위원장	조일환 (명지대학교 교수) – 반도체소자 및 재료	김상인 (아주대학교 교수) – 광파 및 양자전자	
	이강윤 (성균관대학교 교수) – SoC설계	황인철 (강원대학교 교수) – RF집적회로	
	정원영 (강운공업 본부장) – PCB&Package	김익균 (한국전자통신연구원 부장) – 정보보안시스템	
	장익준 (경희대학교 교수) – 내방사선 반도체 설계 및 소자	김한구 (EESOP 대표) – ESD/EOS & Latchup	
	노원우 (연세대학교 교수) – 인 메모리 컴퓨팅	송민규 (동국대학교 교수) – 이미지센서	
협동위원	강석형 (POSTECH 교수)	궁재하 (고려대학교 교수)	권구덕 (강원대학교 교수)
	권영수 (한국전자통신연구원 책임연구원)	김민준 (한국외국어대학교 교수)	김병수 (한국전자기술연구원 센터장)
	김수연 (동국대학교 교수)	김영민 (홍익대학교 교수)	김재욱 (한국과학기술연구원 그룹장)
	김태환 (한국항공대학교 교수)	류성주 (서강대학교 교수)	류시강 (한국항공대학교 교수)
	박관서 (연세대학교 교수)	박성정 (건국대학교 교수)	박인호 (인하대학교 교수)
	박준은 (성균관대학교 교수)	방혁준 (쿤텍 대표)	백지선 (부산대학교 교수)
	서민재 (서울시립대학교 교수)	송준영 (인천대학교 교수)	양준성 (연세대학교 교수)
	오윤호 (성균관대학교 교수)	유호영 (충남대학교 교수)	윤명국 (이화여자대학교 교수)
	이영주 (POSTECH 교수)	이우주 (중앙대학교 교수)	이윤명 (성균관대학교 교수)
	이주열 (가천대학교 교수)	이형민 (고려대학교 교수)	임채강 (한국외국어대학교 교수)
	전동석 (서울대학교 교수)	전성훈 (삼성전자 상무)	전재훈 (인하대학교 교수)
	전천후 (단국대학교 교수)	정무경 (SK 사피온 CTO)	정윤호 (한국항공대학교 교수)
	제민규 (KAIST 교수)	창동진 (충남대학교 교수)	채주형 (광운대학교 교수)
	채형일 (건국대학교 교수)	최우영 (서울대학교 교수)	최재혁 (서울대학교 교수)
	최재혁 (성균관대학교 교수)	한정환 (충남대학교 교수)	

컴퓨터소사이어티

회장	이덕기 (연암공과대학교 교수)		
명예회장	강문식 (강릉원주대학교 교수)	김승천 (한성대학교 교수)	김형중 (고려대학교 교수)
	박인정 (단국대학교 명예교수)	박춘명 (한국교통대학교 교수)	신인철 (단국대학교 명예교수)
	안병구 (홍익대학교 교수)	안현식 (동명대학교 교수)	이규대 (공주대학교 교수)
	정용규 (을지대학교 교수)	최용수 (신한대학교 교수)	허 영 (스마트의료기기산업진흥재단 부이사장)
	홍유식 (상지대학교 교수)	황성운 (가천대학교 교수)	
자문위원	권호열 (강원대학교 교수)	남상엽 (국제대학교 교수)	정교일 (조선대학교 교수)
감사	박수현 (국민대학교 교수)	변영재 (UNIST 교수)	
부회장	김도현 (제주대학교 교수)	박성욱 (강릉원주대학교 교수)	심정연 (강남대학교 교수)
	윤은준 (경일대학교 교수)	조민호 (고려대학교 교수)	
협동부회장	강상욱 (상명대학교 교수)	김병서 (홍익대학교 교수)	김영학 (산업기술평가관리원 본부장)
	우운택 (KAIST 교수)	유성철 (쌍용정보통신 본부장)	이기영 (인천대학교 교수)
	이후진 (한성대학교 교수)	정은성 (홍익대학교 교수)	황인정 (명지병원 책임연구원)
총무이사	심규성 (한경대학교 교수)		
재무이사	박영훈 (숙명여자대학교 교수)		
홍보이사	오희석 (한성대학교 교수)		
편집이사	강병권 (순천향대학교 교수)	기장근 (공주대학교 교수)	김선욱 (고려대학교 교수)
	김정욱 (경희대학교 교수)	김진홍 (배재대학교 교수)	김천식 (세종대학교 교수)
	김한울 (서울과학기술대학교 교수)	노소영 (월송출판 대표)	윤상훈 (한국전자기술연구원 책임연구원)
	이문구 (김포대학교 교수)	이민호 (경북대학교 교수)	이세호 (전북대학교 교수)
	이주형 (가천대학교 교수)	이찬수 (영남대학교 교수)	정혜명 (김포대학교 교수)
	한규필 (금오공과대학교 교수)		
학술이사	고한얼 (고려대학교 교수)	김대홍 (을지대학교 교수)	김명선 (한성대학교 교수)
	김종윤 (경동대학교 교수)	김형진 (전북대학교 교수)	김홍균 (다스파워 이사)
	백광현 (중앙대학교 교수)	서민석 (고려대학교 교수)	신건윤 (한경대학교 교수)
	이정선 (을지대학교 교수)	이충규 (조선대학교 교수)	임경원 (대림대학교 교수)
	임재균 (명지병원 소장)	진 훈 (안양대학교 교수)	최한용 (신한대학교 교수)
	한상민 (순천향대학교 교수)	한영선 (부경대학교 교수)	한태화 (연세대의료원 팀장)
	황석중 (SK텔레콤 박사)	황재정 (군산대학교 교수)	
사업이사	박승창 (유오씨 사장)	차시호 (청운대학교 교수)	황진영 (한국항공대학교 교수)
산학이사	권태윤 (네이버시스템즈 대표이사)	김대휘 (한국정보통신 대표)	김성길 (우주텔레콤 이사)
	김용수 (비온시이노베이터 대표이사)	김은영 (투와이시스템즈 이사)	김진수 (지엔소프트 대표이사)
	김효선 (연세세브란스병원 연구원)	서봉상 (올포랜드 이사)	송치봉 (웨이버스 이사)
	신동희 (엠티데이타 부장)	오승훈 (주얼린 대표)	이명옥 (진우ATS 팀장)
	이재홍 (유비벨록스모바일 대표)	이학준 (이노지에스코리아 연구소장)	임준섭 (대신정보통신 차장)
	전한수 (세림TSG 전무)	조병순 (CNCinstrument 사장)	조병영 (태진인포텍 전무)

연구회위원장

- 황성운 (가천대학교 교수) – 인공지능 및 보안
- 윤은준 (경일대학교 교수) – 융합컴퓨팅
- 정은성 (홍익대학교 교수) – 블록체인
- 김명선 (한성대학교 교수) – AI 응용
- 한태화 (연세대의료원 팀장) – AAST
- 김도현 (제주대학교 교수) – M2M/IoT
- 우운택 (KAIST 교수) – 증강휴먼
- 이민호 (경북대학교 교수) – 인공지능/신경망/퍼지
- 진 훈 (안양대학교 교수) – 휴먼ICT

인공지능 신호처리소사이어티

회장	고병철 (계명대학교 교수)		
자문위원	김정태 (이화여자대학교 교수)	김종옥 (고려대학교 교수)	김창익 (KAIST 교수)
	김홍국 (광주과학기술원 교수)	박종일 (한양대학교 교수)	송병철 (인하대학교 교수)
	심동규 (광운대학교 교수)	이영렬 (세종대학교 교수)	전병우 (성균관대학교 교수)
	조남익 (서울대학교 교수)	홍민철 (숭실대학교 교수)	
부회장	강석주 (서강대학교 교수)	김승룡 (고려대학교 교수)	황원준 (아주대학교 교수)
협동부회장	강경진 (LG전자 연구위원)	권기룡 (부경대학교 교수)	김남수 (서울대학교 교수)
	김진웅 (한국전자통신연구원 그룹장)	김창수 (고려대학교 교수)	백준기 (중앙대학교 교수)
	예종철 (KAIST 교수)	윤재웅 (LG전자 연구위원)	이병욱 (이화여자대학교 교수)
	이찬수 (영남대학교 교수)	지인호 (홍익대학교 교수)	최강선 (한국기술교육대학교 교수)
	최병호 (한국전자기술연구원 센터장)	하정우 (네이버 Head)	한재준 (삼성전자 마스터)
감사	홍성은 (성균관대학교 교수)	이종택 (경북대학교 교수)	
총무	정승원 (고려대학교 교수)		
이사	강수명 (계명대학교 교수)	고영준 (충남대학교 교수)	고종환 (성균관대학교 교수)
	고현석 (한양대학교 교수)	곽노준 (서울대학교 교수)	곽수영 (한밭대학교 교수)
	곽수하 (POSTECH 교수)	구본학 (LG전자 연구원)	김동국 (전남대학교 교수)
	김상원 (한국전자통신연구원 연구원)	김원준 (건국대학교 교수)	김준모 (KAIST 교수)
	김진규 (고려대학교 교수)	김태욱 (한양대학교 교수)	김태현 (한양대학교 교수)
	김학구 (중앙대학교 교수)	김한울 (서울과학기술대학교 교수)	김휘용 (경희대학교 교수)
	남우정 (경북대학교 교수)	문경식 (DGIST 교수)	민동보 (이화여자대학교 교수)
	박상현 (DGIST 교수)	박영경 (이화여자대학교 교수)	박은병 (성균관대학교 교수)
	박인규 (인하대학교 교수)	박진선 (부산대학교 교수)	박철수 (광운대학교 교수)
	백종덕 (연세대학교 교수)	손진희 (광주과학기술원 교수)	신종원 (광주과학기술원 교수)
	심재영 (UNIST 교수)	심현정 (KAIST 교수)	엄찬호 (중앙대학교 교수)
	오병태 (한국항공대학교 교수)	우성민 (한국기술교육대학교 교수)	우운택 (KAIST 교수)
	유윤규 (두산인프라코어 연구원)	유재준 (UNIST 교수)	유종빈 (아주대학교 교수)
	유현우 (UNIST 교수)	윤정인 (한화디펜스 연구원)	이 철 (동국대학교 교수)
	이경오 (한국전자통신연구원 선임연구원)	이금하 (KT 연구원)	이범식 (조선대학교 교수)
	이상철 (인하대학교 교수)	이윤구 (광운대학교 교수)	이채은 (한양대학교 교수)
	임성빈 (고려대학교 교수)	임성훈 (DGIST 교수)	장준혁 (한양대학교 교수)
	전세영 (서울대학교 교수)	정교일 (조선대학교 교수)	정영주 (숙명여자대학교 교수)
	정희철 (경북대학교 교수)	조동현 (충남대학교 교수)	조성인 (동국대학교 교수)
	조성현 (POSTECH 교수)	차영수 (고려대학교 교수)	최 욱 (인천대학교 교수)
	최동걸 (한밭대학교 교수)	최성준 (고려대학교 교수)	최종원 (중앙대학교 교수)
	최종현 (연세대학교 교수)	최해철 (한밭대학교 교수)	한보형 (서울대학교 교수)
	한재호 (고려대학교 교수)	함범섭 (연세대학교 교수)	황효석 (경희대학교 교수)
협동이사	강현수 (충북대학교 교수)	곽노준 (서울대학교 교수)	곽진태 (고려대학교 교수)
	구형일 (아주대학교 교수)	권구락 (조선대학교 교수)	권준석 (중앙대학교 교수)
	김광주 (한국전자통신연구원 연구원)	김기백 (숭실대학교 교수)	김동현 (연세대학교 교수)
	김상효 (성균관대학교 교수)	김선주 (연세대학교 교수)	김용환 (한국전자기술연구원 수석)
	김원하 (경희대학교 교수)	김재곤 (한국항공대학교 교수)	김정진 (계명대학교 교수)
	김종민 (강원대학교 교수)	김중욱 (항공우주연구원 연구원)	김진우 (경성대학교 교수)
	김해광 (세종대학교 교수)	남승우 (한국전자통신연구원 연구원)	박구만 (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 연구원)	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 (계명대학교 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)
연구회위원장	황 원 준 (아주대학교 교수) – 딥러닝		강 제 원 (이화여자대학교 교수) – 영상처리
	장 길 진 (경북대학교 교수) – 음향 및 음성신호처리		이 덕 우 (계명대학교 교수) – 영상이해
	이 종 호 (서울대학교 교수) – 바이오영상신호처리		김 성 우 (서울대학교 교수) – 로봇지능

시스템 및 제어소사이어티

회 장	이 덕 진 (전북대학교 교수)		
명 예 회 장	김 덕 원 (연세대학교 교수)	김 희 식 (서울시립대학교 교수)	박 종 국 (경희대학교 교수)
	서 일 홍 (한양대학교 교수)	오 상 록 (한국과학기술연구원 원장)	오 승 록 (단국대학교 교수)
	오 창 현 (고려대학교 교수)	유 정 봉 (공주대학교 교수)	정 길 도 (전북대학교 교수)
	주 영 복 (한국기술교육대학교 교수)	허 경 무 (단국대학교 교수)	
부 회 장	권 종 원 (한국산업기술시험원 센터장)		
감 사	김 영 진 (한국생산기술연구원 수석연구원)	김 영 철 (군산대학교 교수)	
총무이사 (겸)	김 기 연 (한국산업기술시험원 선임연구원)		
편집/학술이사	강 성 복 (한국생산기술연구원 수석연구원)	서 영 석 (영남대학교 교수)	이 왕 상 (경상대학교 교수)
	임 대 영 (한국산업기술시험원 선임연구원)		
홍 보 이 사	김 현 갑 (B&P인스트루먼트 연구소장)	박 덕 우 (한국기계전기전자시험연구원 센터장)	최 현 택 (한국해양과학기술원 책임연구원)
산 학 연 이 사	김 광 식 (소올아이티 전무)	문 태 주 (부일하우징 대표)	신 대 현 (대연씨앤아이 대표)
	유 병 길 (현대선기 대표)		
회 원 이 사	권 오 민 (충북대학교 교수)	김 수 찬 (한경대학교 교수)	김 종 만 (전남도립대학교 교수)
	김 호 철 (을지대학교 교수)	남 기 창 (동국대학교 교수)	류 지 형 (한국전자통신연구원 박사)
	문 정 호 (강릉원주대학교 교수)	박 명 진 (경희대학교 교수)	박 재 병 (전북대학교 교수)
	변 영 재 (UNIST 교수)	송 철 규 (전북대학교 교수)	여 희 주 (대진대학교 교수)
	유 재 현 (한경대학교 교수)	이 수 열 (경희대학교 교수)	이 용 귀 (한국전자통신연구원 책임연구원)
	이 태 희 (전북대학교 교수)	이 학 성 (세종대학교 교수)	정 재 훈 (동국대학교 교수)
	최 수 범 (한국과학기술정보연구원 연구원)	최 우 영 (전북대학교 교수)	한 아 (한국산업기술시험원 선임연구원)
연구회위원장	권 종 원 (한국산업기술시험원 센터장) – 인공지능 전환(AX)		남 기 창 (동국대학교 교수) – 의용전자 및 생체공학
	연 규 봉 (한국자동차연구원 수석연구원) – 자동차전자		오 창 현 (고려대학교 교수) – 의료영상시스템
	이 석 재 (대구보건대학교 교수) – 국방정보및제어		이 성 준 (한양대학교 교수) – 회로 및 시스템
	정 범 진 (서울과학기술대학교 교수) – 스마트미터링		정 재 훈 (동국대학교 교수) – 지능로봇
	한 수 희 (POSTECH 교수) – 제어계측		

산업전자소사이어티

회 장	고 정 환 (인하공업전문대학 교수)		
명 예 회 장	김 은 원 (대림대학교 교수)		
자 문 위 원	김 대 휘 (한국정보기술 대표이사)	원 영 진 (부천대학교 교수)	윤 기 방 (인천대학교 명예교수)
	이 상 회 (동서울대학교 교수)	이 원 석 (동양미래대학교 명예교수)	조 규 남 (로봇신문사 대표이사)
	한 성 준 (대흥정보 고문)		
수석부회장	동 성 수 (용인예술과학대학교 교수)		
상 임 이 사	구 자 일 (인하공업전문대학 교수)	김 상 범 (한국폴리텍대학 교수)	김 영 로 (명지전문대학 교수)
	박 병 섭 (인하공업전문대학 교수)	서 병 석 (상지대학교 교수)	서 춘 원 (K-My지능정보기술 대표이사)
	안 성 수 (명지전문대학 교수)	윤 중 현 (조선이공대학교 교수)	이 정 석 (인하공업전문대학 교수)
	이 태 동 (국제대학교 교수)	조 도 현 (인하공업전문대학 교수)	
협동상임이사	강 현 석 (로보웰코리아 대표이사)	권 오 병 (진인프라 전무이사)	김 세 종 (SJ정보통신 이사)
	김 윤 철 (대흥정보 상무이사)	김 정 석 (오디에이테크놀로지 대표이사)	김 젬 마 (엠티데이타 상무이사)
	김 진 선 (청파이엠티 대표이사)	박 현 영 (진우에이티에스 전무이사)	서 봉 상 (명화지리정보 대표이사)
	서 승 현 (글로벌텔레콤 대표이사)	성 재 용 (오픈링크시스템 대표이사)	송 관 식 (아이씨티웨이 상무이사)
	송 광 헌 (복두출판사 대표이사)	송 치 봉 (웨이버스 대표이사)	유 성 철 (쌍용정보통신 상무이사)
	이 영 준 (투비콤 대표이사)	장 대 현 (LG U+ 상무이사)	전 한 수 (세림TSG 전무이사)
	조 병 영 (태진티엔에스 대표이사)	조 한 일 (하이제이컨설팅 대표이사)	최 석 우 (대보정보통신 상무이사)
이 사	강 희 훈 (여주대학교 교수)	권 오 상 (경기과학기술대학교 교수)	권 윤 중 (세명대학교 교수)
	김 현 (부천대학교 교수)	김 남 섭 (서일대학교 교수)	김 대 순 (전주비전대학교 교수)
	김 백 기 (강릉원주대학교 교수)	김 용 민 (충청대학교 교수)	김 태 원 (상지대학교 교수)
	김 필 중 (조선이공대학교 교수)	성 홍 석 (부천대학교 교수)	안 태 원 (동양미래대학교 교수)
	엄 우 용 (인하공업전문대학 교수)	우 찬 일 (서일대학교 교수)	원 우 연 (한국폴리텍대학 교수)
	이 문 구 (김포대학교 교수)	이 상 철 (인천재능대학교 교수)	이 성 재 (대림대학교 교수)
	이 시 현 (동서울대학교 교수)	이 용 구 (한림성심대학교 교수)	이 주 연 (전주비전대학교 교수)
	장 기 동 (동양미래대학교 교수)	최 의 선 (한국폴리텍대학 교수)	
협 동 이 사	고 강 일 (이지테크 대표이사)	곽 정 희 (한국정보기술 이사)	김 응 연 (인터그래텍 대표이사)
	김 철 룡 (LIG 책임연구원)	변 용 필 (진인프라 전무이사)	신 동 희 (엠티데이타 상무이사)
	신 우 현 (시티랩스 이사)	오 재 곤 (한국정보기술 상무이사)	이 경 원 (동해종합기술공사 상무이사)
	이 성 대 (엔클라우드 상무이사)	이 승 민 (튠시스템 대표이사)	이 승 태 (하나텍시스템 이사)
	이 용 우 (쌍용정보통신 상무이사)	이 지 학 (송암시스콤 대표이사)	임 준 섭 (대신정보통신 차장)
	장 기 웅 (나날에스엠아이 부장)	장 철 (우송정보기술 대표이사)	정 민 우 (한국정보기술 상무이사)
	정 용 훈 (롯데정보통신 선임)	한 찬 석 (동해종합기술공사 부사장)	
감 사	김 동 식 (인하공업전문대학 교수)	김 영 선 (대림대학교 교수)	

제24대 평의원 명단

강 명 곤 (서울시립대학교 교수)
강 석 판 (LG전자 상무)
강 제 원 (이화여자대학교 교수)
고 성 제 (고려대학교 명예교수)
고 진 환 (경상대학교 교수)
공 배 선 (성균관대학교 교수)
구 민 석 (서울시립대학교 교수)
궁 재 하 (고려대학교 교수)
권 구 덕 (강원대학교 교수)
권 태 수 (서울과학기술대학교 교수)
김 강 태 (삼성전자 부사장)
김 기 현 (전북대학교 교수)
김 도 훈 (한국전자통신연구원 박사)
김 동 현 (ICTK 대표이사)
김 민 휘 (중앙대학교 교수)
김 상 태 (전남테크노파크 본부장)
김 성 대 (한국과학기술원 명예교수)
김 수 중 (경북대학교 명예교수)
김 시 호 (연세대학교 교수)
김 영 민 (홍익대학교 교수)
김 영 재 (해동과학문화재단 이사장)
김 원 종 (한국전자통신연구원 책임연구원)
김 은 원 (대림대학교 교수)
김 재 현 (아주대학교 교수)
김 종 옥 (고려대학교 교수)
김 진 상 (경희대학교 교수)
김 철 우 (고려대학교 교수)
김 한 구 (EESOP 대표)
김　현 (부천대학교 교수)
김 형 진 (한양대학교 교수)
나 정 웅 (한국과학기술원 명예교수)
남 일 구 (부산대학교 교수)
노 정 진 (한양대학교 교수)
류 성 주 (서강대학교 교수)
문 상 미 (나사렛대학교 교수)
민 경 식 (국민대학교 교수)
박 규 태 (연세대학교 명예교수)
박 성 정 (건국대학교 교수)
박 영 훈 (숙명여자대학교 교수)
박 종 일 (한양대학교 교수)
박 철 수 (광운대학교 교수)
배 순 민 (케이티 Lab장)
백 광 현 (중앙대학교 교수)
범 진 욱 (서강대학교 교수)
서 민 재 (서울시립대학교 교수)
서 승 우 (서울대학교 교수)
서 지 원 (연세대학교 교수)
성 굉 모 (서울대학교 명예교수)
송 민 규 (동국대학교 교수)
송 익 현 (한양대학교 교수)
신 세 운 (포항공과대학교 교수)
신 창 환 (고려대학교 교수)
심　용 (중앙대학교 교수)
안 상 철 (한국과학기술연구원 책임연구원)
안 호 균 (한국전자통신연구원 책임연구원)
연 규 봉 (한국자동차연구원 수석연구원)
오 태 현 (포항공과대학교 교수)

강 문 식 (강릉원주대학교 교수)
강 성 원 (한국전자통신연구원 부원장)
고 병 철 (계명대학교 교수)
고 용 남 (하나마이크론 전무)
고 형 호 (충남대학교 교수)
공 준 진 (웨이브피아 CSO)
구 본 태 (한국전자통신연구원 본부장)
권 건 우 (홍익대학교 교수)
권 기 룡 (부경대학교 교수)
권 혁 인 (중앙대학교 교수)
김 경 기 (대구대학교 교수)
김 도 현 (제주대학교 교수)
김 동 규 (한양대학교 교수)
김 동 현 (연세대학교 교수)
김 병 서 (홍익대학교 교수)
김 선 용 (건국대학교 교수)
김 성 우 (서울대학교 교수)
김 승 천 (한성대학교 교수)
김 영 권 (후레대학교 명예총장)
김 영 민 (서울대학교 교수)
김 영 한 (UC San Diego/가우스랩스 대표이사)
김 원 준 (건국대학교 교수)
김 익 균 (한국전자통신연구원 부장)
김 재 희 (연세대학교 명예교수)
김 중 헌 (고려대학교 교수)
김 진 영 (광운대학교 교수)
김 태 욱 (연세대학교 교수)
김　혁 (서울시립대학교 교수)
김 현 수 (삼성전자 상무)
김 형 탁 (홍익대학교 교수)
남 광 희 (포항공과대학교 교수)
노 승 원 (LG이노텍 CTO)
노 태 문 (한국전자통신연구원 연구전문위원)
류 수 정 (서울대학교 교수)
문 영 식 (한림성심대학교 총장)
민 동 보 (이화여자대학교 교수)
박 성 욱 (차세대지능형반도체사업단 이사장)
박 성 한 (한양대학교 명예교수)
박 인 규 (인하대학교 교수)
박 준 석 (인하대학교 교수)
박 항 구 (소암시스텔 회장)
배 준 성 (강원대학교 교수)
백 종 덕 (연세대학교 교수)
변 대 석 (삼성전자 마스터)
서 병 석 (상지대학교 교수)
서 정 목 (연세대학교 교수)
서 철 헌 (숭실대학교 교수)
손 교 민 (삼성전자 마스터)
송 병 철 (인하대학교 교수)
송 준 영 (인천대학교 교수)
신 오 순 (숭실대학교 교수)
신 현 철 (광운대학교 교수)
안 광 호 (한국전자기술연구원 본부장)
안 성 수 (명지전문대학 교수)
양 준 성 (연세대학교 교수)
오 윤 제 (정보통신기획평가원 PM)
우 성 민 (한국기술교육대학교 교수)

강 석 주 (서강대학교 교수)
강 윤 희 (백석대학교 교수)
고 석 준 (제주대학교 교수)
고 정 환 (인하공업전문대학 교수)
공 규 열 (한성대학교 교수)
곽 진 태 (고려대학교 교수)
구 용 서 (단국대학교 석좌교수)
권 경 하 (한국과학기술원 교수)
권 종 원 (한국산업기술시험원 센터장)
권 호 열 (강원대학교 교수)
김 기 연 (한국산업기술시험원 선임연구원)
김 도 현 (국민대학교 명예교수)
김 동 식 (인하공업전문대학 교수)
김 명 선 (한성대학교 교수)
김 상 범 (한국폴리텍대학 교수)
김 선 욱 (고려대학교 교수)
김 소 영 (성균관대학교 교수)
김 승 환 (한국전자통신연구원 책임연구원)
김 영 로 (명지전문대학 교수)
김 영 선 (대림대학교 교수)
김 용 신 (고려대학교 교수)
김　윤 (서울시립대학교 교수)
김 재 준 (서울대학교 교수)
김 종 선 (홍익대학교 교수)
김 지 훈 (이화여자대학교 교수)
김 철 영 (충남대학교 교수)
김 학 구 (중앙대학교 교수)
김　현 (서울과학기술대학교 교수)
김 형 준 (한국과학기술연구원 소장)
김　훈 (인천대학교 교수)
남 기 창 (동국대학교 교수)
노 원 우 (연세대학교 교수)
동 성 수 (용인예술과학대학교 교수)
류 은 석 (성균관대학교 교수)
문　용 (숭실대학교 교수)
박 관 서 (연세대학교 교수)
박 성 욱 (강릉원주대학교 교수)
박 수 현 (국민대학교 교수)
박 종 선 (고려대학교 교수)
박 진 옥 (육군사관학교 명예교수)
방 승 찬 (한국전자통신연구원 원장)
배 현 철 (한국전자통신연구원 책임연구원)
백 준 기 (중앙대학교 교수)
변 영 재 (울산과학기술원 교수)
서 봉 상 (올포랜드 상무)
서 종 모 (서울대학교 교수)
선 우 경 (서울대학교 교수)
송 문 섭 (심텍 회장)
송 상 헌 (중앙대학교 교수)
송 진 호 (연세대학교 교수)
신 요 안 (숭실대학교 교수)
심 동 규 (광운대학교 교수)
안 병 구 (홍익대학교 교수)
안 현 식 (동명대학교 교수)
엄 낙 웅 (한국전자통신연구원 책임연구원)
오 윤 호 (고려대학교 교수)
우 운 택 (한국과학기술원 교수)

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

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

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

사무국 직원 명단

송기원 국장 – 산학연관 협력, 신규 사업, 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(ICCE–Asia), 하계종합학술대회, 신기술(총괄), 차세대리더육성, 시스템 및 제어소사이어티
배기동 부장 – 사업행사(기술워크숍 등), 국문논문, 표준화, 용역업무, 인공지능신호처리소사이어티
이소진 서기 – 국제학술대회(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), 차세대리더육성, 기타 지원업무, 반도체소사이어티
곽새롬 서기 – 본회/소사이어티/연구회 재무, 회원관리(개인회원 및 특별회원), 기획, 통신소사이어티

제1차 상임이사회 회의록

제1차 상임이사회가 1월 3일(금) 오후 5시 SC컨벤션센터 아나이스홀(한국과학기술회관 1관 12층)에서 개최되었다. 이번 회의 결과는 다음과 같다.

– 다 음 –

1. 성원 보고

– 54명의 상임이사 중 47명이 참석하여 성원되었음.

2. 위원회 보고

– 각 위원회 별로 위원장들이 서면 및 발표 보고사항을 진행하고 위원회별 주요사항을 논의함.

3. 논의사항

– 대한민국학술원 회원 추천 관련하여 회장단 진행으로 논의함.

4. 심의사항

– 주요 승인위원회 (재정위원회 / 선거관리위원회) 위원 선정에 대해 원안대로 승인함.

5. 기타

특별회원 및 유관기관 방문

학회 회장단은 2월 4일부터 13일까지 도쿄일렉트론코리아, 에스에스엔씨, 국가보안기술연구소, 텔레칩스, 한국전자통신연구원, 삼성전자, 호리바에스텍코리아, 히타치하이테크코리아, 케이케이테크코리아, 자람테크놀로지, DL아이엔씨 등을 방문하여 경영 및 기술 분야의 현안 과제를 논의하고 상호 협력방안을 모색하였다.

한국전자통신연구원 – 강성원 부원장

LG이노텍 – 문혁수 대표이사

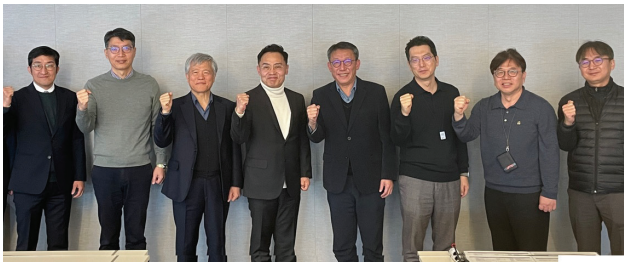
삼성리서치 – 김강태 부사장

자람테크놀로지 – 백준현 대표이사

텔레칩스 – 이장규 대표이사

에스에스앤씨 – 한은혜 대표이사

도쿄일렉트론코리아 – 원제형 대표이사



DL아이앤씨 – 배원복 부회장

호리바에스텍코리아 – 김성환 회장

케이케이테크코리아 – 김경하 대표이사

2025년 영상이해연구회 겨울학교

인공지능신호처리소사이어티 영상이해연구회[위원장: 김원준 교수(건국대)] 개최로 "2025년 영상이해연구회 겨울학교"는 1월 14일(화)~16일(목)까지 비발디파크 소노캄 다이아몬드홀(2층)에서 개최되었다.

이번 영상이해연구회 겨울학교에서는 컴퓨터 비전 분야에서 최근 널리 연구되고 있는 주제의 기초 이론부터 최신 연구 동향까지 폭넓은 주제로 진행되었다. 특히, 다양한 신경망 구조 및 학습 이론, 멀티모달 데이터 기반 응용 등 많은 관심을 받고 있는 연구 주제와 더불어, 좋은 연구자가 되기 위한 멘토링 강연이 진행되어 대학원 신입생뿐만 아니라 석/박사과정 학생, 산업계 연구원들이 많이 참여하였다. 참석은 약 175여 명이었다.

2025년 영상이해연구회 겨울학교 강연 모습(01.14–16)

2025년 바이오영상신호처리연구회 겨울학교

인공지능신호처리소사이어티 바이오영상신호처리연구회[위원장: 이종호 교수(서울대)] 개최로 "2025년 바이오영상신호처리 겨울학교"는 1월 16일(목)~17일(금)까지 호암교수회관 무궁화홀에서 개최되었다.

이번 바이오영상신호처리 겨울학교에서는 의료영상 분야에서의 최신 연구 결과를 공유하고 함께 성장하기 위하여 국내외에서 주목받고 있는 다양한 분야의 전문가들이 참여하여 의료영상 기술과 응용 분야 등 새로운 기술과 연구 동향에 대한 흥미로운 강연과 논의의 장이 되었다. 참석은 약 160여 명이었다.

2025년 바이오영상신호처리 겨울학교 강연 모습(01.16–17)

ICEIC 2025

제24회 International Conference on Electronics, Information, and Communication (ICEIC) 국제학술대회가 2025년 1월 19일부터 22일까지 4일에 걸쳐 일본 오사카에서 개최되었다. ICEIC는 대한전자공학회(IEIE)에서 주최하는 대표적인 국제학술대회로 우리나라 정보통신 기술 · 학문 · 산업의 글로벌 확산에 목적을 두고 매년 개최되고 있다. 이번 ICEIC 2025 또한 전 세계에서 학계, 산업계 연구자들이 모두 한자리에 모여 서로의 통찰을 공유하고 학술 교류의 장으로 기능하며 성황리에 진행되었다. 특히나 이번 학술대회에는 미래 기술의 원동력이 될 학부생 신진 인재들의 참가가 기존보다 대폭 늘어 더욱 의미 있는 학술대회를 이루었다.

ICEIC 2025는 대한전자공학회 학회장 백광현 교수(중앙대학교)가 General Chair, 권혁인 교수(중앙대학교)가 Organizing Committee Chair, 구민석 교수(서울시립대학교)가 Technical Program Committee Chair로 학술대회를 이끌었다. 본 학술대회에서는 총 628편의 논문이 17개국에서 제출되어 발표되었으며, 총 1,085명이 전 세계 14개국에서 참가 등록을 하여 학회장을 가득 채웠다.

학술대회 첫째 날인 1월 19일에는 한양대학교 추민성 교수와 동국대학교 조성인 교수의 Tutorial 강연과 학부생 특별 포스터발표 세션이 진행되며 학술대회 일정이 시작되었다. 둘째 날인 1월 20일에는 개회식에서 백광현 학회장의 환영사와 함께 학술대회의 공식적인 시작을 알렸으며, 이어서 TPC Chair를 맡고 있는 구민석 교수의 TPC Report 발표를 통해 이번 학술대회의 성과를 참가자들에게 공유하였다. 그다음으로는 대한민국 과학기술정보통신부 장관을 역임한 서울대학교 이종호 교수와 일본 Ritsumeikan University의 이주호 교수의 Plenary Talk가 진행되며 개회식장을 가득 채운 참가자들에게 의미 있는 강연을 전달하였다. 같은 날 Tutorial 강연으로 일본 The University of Aizu의 TOMIOKA Yoichi 교수와 KAIST 제민규 교수의 강연이 진행되었고, 삼성전자에서 준비한 Invited Talk 강연에서 송택상 박사, 백승재 박사, 이천안 박사의 강연 또한 많은 참가자의 관심 속에 진행되었다. 이와 동시에 각 세션 룸에서는 구두 발표와 포스터발표가 이루어지며 둘째 날 일정을 풍성하게 채웠다. 학술대회 셋째 날인 1월 21일에는 중국 Shanghai Jiao Tong University의 Jian Zhao 교수와 고려대학교 김진규 교수의 Tutorial 강연이 진행되었으며, 동시에 일반 세션 룸에서는 구두 발표와 포스터발표가 진행되었다. 이어서 마지막 날인 1월 22일에 모든 세션의 마무리와 함께 학술대회의 마침표를 찍었다.

이번 행사를 통틀어 총 23개의 스페셜 구두 발표 세션, 20개의 일반 구두 발표 세션, 1개의 학부생 구두 발표 세션, 7개의 포스터발표 세션, 3개의 학부생 포스터발표 세션이 진행되었다. 여러 초청 강연과 다양한 주제의 발표, 학부생 세션 등 풍성한 프로그램으로 채운 ICEIC 2025는 행사 일정 내내 뜨거운 열기와 인파 속에서 성공적으로 진행되었으며, 세계적인 연구자들이 모여 향후 정보통신기술 분야의 연구와 새로운 기술 발전을 도모하는 뜻깊은 기회를 제공하는 역할을 하였다.

개회식 – 백광현 회장 환영사


구두 발표 세션

Plenary Talk – 이종호 교수(서울대학교)


포스터 발표 세션

Plenary Talk – 이주호 교수(Ritsumeikan University)


Best Paper 시상식

Tutorial – TOMIOKA Yoichi 교수(The University of Aizu)

주요참가자 기념촬영 – 초청연사, 조직위원 등

학회일지

www.theieie.org

THE INSTITUTE OF
ELECTRONICS AND INFORMATION
ENGINEERS

(2025년 1월 18일 ~ 2월 16일)

1. 회의 개최

회의 명칭	일시	장소	주요 안건
제1차 하계 조직위원회 회의	2.7 (17:00)	학회 회의실	– 학술대회 운영 및 프로그램 구성 등

특 · 집 · 편 · 집 · 기

인공지능 · 신호처리 효율화 및 응용 기술

정 승 원 편집위원
(고려대학교)

최근 인공지능 기술은 신호처리 분야에서 효율적인 데이터 처리 및 분석 기술을 개발하는 데 있어 중요한 역할을 하고 있다. 특히, 신호처리의 효율화 및 최적화는 고성능과 저전력을 동시에 요구하는 IT 융합 시스템에서 경쟁력을 좌우하는 핵심 요소로 자리 잡고 있다. 본 특집호는 인공지능 신호처리 기술의 효율적 구현 및 응용을 중심으로 최신 연구 동향과 발전 가능성을 다룬다. 구체적으로, 딥러닝 기반 신호처리 알고리즘, 효율적인 인공지능 모델 설계 기술, 실제 응용 사례 (CCTV, LLM, 국방) 등 다양한 주제를 포함하며, 학계와 산업계 전문가들의 총 6편의 기고문으로 구성되었다.

첫째, "지능형 CCTV AI 신호처리 응용기술(민병석)"에서는 지능형 CCTV 시스템의 특징과 구성 요소 기술을 분석하며 객체 크기와 해상도, 시인성 저하 등의 문제를 해결하는 기술이 실제 산업 현장에 적용된 사례를 소개하였다. 둘째, "YOLO 아키텍처 기반의 자기증류를 활용한 객체 탐지 성능 향상(이동석 외)"은 지식 증류 방법을 개발하여 YOLO v6 모델에 적용한 사례를 통해 고효율 객체 탐지 전략을 제시하였다. 셋째, "AI의 메타인지 능력 개발(황중원 외)"에서는 스스로 학습역량을 인지하고 활용하여 결과를 제공하는 인공지능 기술의 개념과 강화학습에서의 메타인지 AI 구현 사례를 소개하였다. 넷째, "Scientific LLMOps: 과학기술 분야 AI 개발의 중추(이용 외)"에서는 기존의 인공지능 개발을 지원하는 다양한 Ops 기술들을 비교 분석하고 과학기술 LLM 개발과 운용에 요구되는 Scientfic LLMOps 특성을 도출하였다. 다섯째, "영상 초해상화 인공지능 모델의 양자화 기술(홍채은)"에서는 영상 초해상화 모델이 갖는 신호 특성을 활용한 다양한 양자화 기법을 소개하였다. 끝으로, "구조적 관점에서의 영상 초해상화 인공지능 모델의 경량화 기법 분석(손명준 외)"에서는 인공지능 신경망 구조적인 관점에서 연산 효율성 달성을 위하여 개발된 다양한 설계 기법들을 소개하였다.

연말연시의 바쁜 일정 중에 본 특집호를 위하여 옥고를 보내주신 집필진 여러분께 감사드리며, 본 특집호가 인공지능 신호처리 시스템의 효율화 기술 및 응용 개발을 위한 지침이 되어 우리나라 IT융합산업의 발전과 경쟁력 강화에 기여할 수 있기를 기원한다.

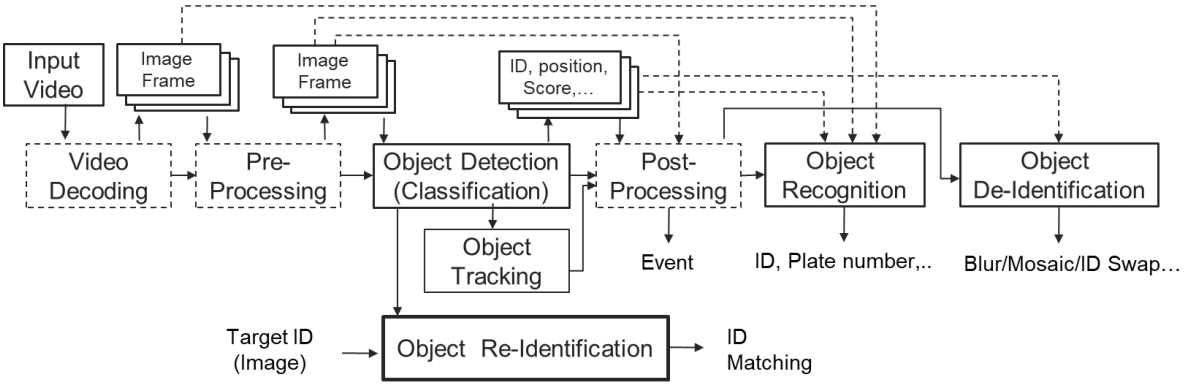
지능형 CCTV AI 신호처리 응용기술

Ⅰ. 서 론

현대 사회에서 CCTV(폐쇄회로 텔레비전)는 단순히 모니터링을 넘어 범죄 예방과 수사, 시설물 관리, 교통 통제, 마케팅에서의 고객 분석 등 다양한 분야에서 활용되고 있다. CCTV는 초기에는 사람, 자동차 외 다양한 사물이나 객체의 움직임이나 행동 등을 관찰하기 위해 설치되었으나, 범죄 예방과 재난 · 재해 방지 등 안전에 대한 요구가 증가하면서 빠르게 확산되어 전국의 대부분의 지역에 설치되어 관리되고 있고, 많은 기술 발전으로 CCTV 시스템의 통합이 가능해지며, 다수의 CCTV 영상을 한 곳에서 관리할 수 있는 통합관제센터가 운영되고 있다. 하지만 CCTV 카메라의 수가 많아지면서 소수의 인력이 여러 카메라 영상을 관제하는 데 한계가 생기자, 지능형 CCTV가 효과적인 대안으로 주목받고 많은 관련된 기술이 개발되어 적용되어 왔다. 지능형 CCTV는 관제구역에서 발생하는 상황을 자동으로 분석해 관제요원에게 알리는 시스템으로, 적은 인력으로도 여러 대의 카메라를 효율적으로 관리할 수 있고 이 기술은 정보보호 분야 중 물리보안에 속하며, 점차 매출 비중과 중요성이 커지고 있고, 물리보안을 넘어 정보보호 산업 전체의 성장을 견인하고 있다.

이러한 지능형 CCTV의 핵심기술은 사람, 사물의 행위를 얼마나 정확하게 식별하여 해당 시스템이 자동화되어 확장 될 수 있는가 하는 점이 가장 중요한 요소인데, 여러 기술적인 단계를 거쳐 그 성능이 개선 발전되어 왔다. 가장 초기의 1세대 검출방식은 모션 디텍션 기반 기술로, 해당하는 객체를 검출하여 판단하는 방식이 아니라, CCTV 내의 영상내 객체 움직임 등을 변화하는 지를 보는 방식이다. 그러나 이렇게 Frame과 Frame 사이의 픽셀 변화를 통해 움직임을 탐지하는 기법은 객체의 움직임이 아닌 밝기의 변화, 나뭇가지나 눈, 비와 같은 작은 물

민병석
MarkAny 연구소



〈그림 1〉 CCTV 전체 시스템 알고리즘 구성도

체의 움직임 만으로도 오검출이 발생하는 기술적인 한계를 가지고 있다. 이후 2세대의 방식은 영상 분석을 통한 탐지 방식으로, 객체 탐지 기술과, 배경과 객체의 분리기술, 객체 추적기술 등을 사용하여, 배경영역에 대한 영향도를 최소화 하면서 실제 객체의 움직임에 특화된 방식으로 1세대의 모션검출 방식대비 오검출이나 미검출에 대한 성능이 대폭 향상되었다. 특히 AI 딥러닝 네트워크를 통한 객체 탐지 및 분류, 객체 추적기술들을 통한 대폭적인 성능 향상과 함께, 고사양의 GPU가 CCTV 관제 시스템에 탑재되면서 AI를 통한 지능관제 시스템이 광범위하게 확대되고 있다. 영상분석을 통한 탐지 방식은 객체 분석을 통한 다양한 영상 내 이벤트 검출이 가능하여, 영상보안과 관련된 배회, 침입, 폭행, 유기 등의 이벤트 탐지뿐 아니라, 안전과 관련된 화재, 연기, 군중 밀집에 대한 이벤트 탐지, 비즈니스 인텔리전스와 관련된 매장 고객수, 마트 계산대 대기열 분석 등으로 확대되어 사용되고 있다. 특히 AI 기반 객체 검출에 대한 기술은 딥러닝 네트워크 및 모델 가중치가 공개되어 실제 전문가뿐 아니라 영상처리 및 컴퓨터 비전에 전문성이 낮은 사람들도 쉽게 다양한 분야에 활용할 수 있는 환경을 제공하고 있다. 특히 저가의 GPU에도 적용될 수 있는 경량화 구조를 가지면서도 Detection과 Classification을 동시에 수행하는 높은 성능을 가지는 YOLO와 같은 AI 모델이 매년 높은 성능을 가진 지속적으로 오픈되는 환경에서, 실질적인 산업제품군과 지능형 CCTV 시스템에 단일 객체 검출모델만으로도 상용화 수준의 서비스가 제공되고 있다. AI 개발을 위해 필요한 인공지능 학습용 데이터 측면에서는 기존의 논문 연구개발을 위한 학술적 목표의 학습 DB뿐 아니라, 한국정보진흥원의 AI Hub를 통한 실환경 학습 DB도 모든 사람이 사용할 수 있도록 공유되어 사용되고 있다[1]. 또한 지능형 CCTV 솔루션의 사용 확산 및 만족도를 높이고 실질적인 지능형 CCTV의 기술 경쟁력을 높이기 위해 우수한 솔루션에게는 기술 인증을 주는 지능형 CCTV 성능 시험 인증 체계가 한국인터넷진흥원(KISA)에 의해 시행되고 있고 실제로 많은 CCTV 솔루션 업체들이 해당 인증을 기반으로 많은 CCTV 사업에 참여하고 있다[2]. 해당 기관의 성능 시험 인증은 실제 CCTV 적용환경과 유사한 다양한 환경에서 촬영된 영상 DB를 통해 솔루션이 영상 DB내 특정 이벤트를 판별할 수 있는지 시험하는 것으로, 해당 DB의 다양한 CCTV 영상 조건들이 실제 단순한 학습 모델의 적용만으로는 해결하기 쉽지 않은 기술의 난이도를 보여주고 있다. 이 기고문에서는 현재 적용되고 있는 CCTV 지능화 시스템의 전체 구조 설명과 함께, CCTV 시스템이 가지고 있는 본질적인 성능 한계 요소와 실제 많은 미검출과 오검출이 발생하는 CCTV 환경에서 검출 판단 성능 향상을 위해 사용되는 다양한 영상 신호처리 기술을 응용한 전처리 기법들에 대해 소개를 할 수 있도록 한다.

Ⅱ. CCTV 지능화 시스템 구성

CCTV 영상이 영상처리 및 객체 분류 및 탐지시스템을 거쳐 최종 이벤트 탐지 및 자동차 번호판 및 얼굴 인식, 특정 객체 추적 및 탐색을 위한 시스템 구조도는 다양한

형태를 가질 수 있지만 일반적으로 〈그림 1〉과 같이 표현될 수 있다.

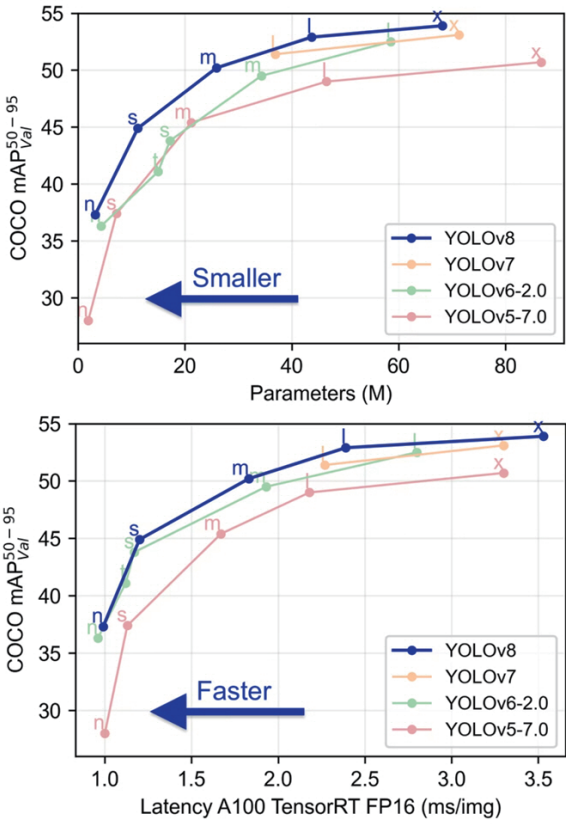
첫 번째 단계는 동영상으로 압축된 비디오 파일을 비디오 프레임으로 변환하는 단계로서, 일반적으로 많이 사용되는 H.264 및 H.265 동영상 압축방식에 대한 Decoding 과정에 해당한다. 디코딩 방식은 HW 장비를 통한 고속 디코딩 방식과, CPU를 통한 디코딩, GPU를 통한 병렬 고속 디코딩 방식들이 사용된다. 일반적으로 CCTV 채널수가 많지 않다면 CPU를 통한 디코딩 방식이 비용 측면에서 효율적일 수 있지만, CCTV 채널수가 늘어나거나 CPU를 통해 처리하는 영상 전처리/후처리 알고리즘의 개수나 사이즈가 커질수록 시스템 부하량이 늘어나는 한계가 존재한다. GPU를 통한 디코딩 방식은 여러 채널을 GPU에서 병렬처리하고 GPU처리된 결과를 Video 메모리에서 추가 작업없이 바로 AI 모델로 바로 처리가 가능하여 속도 및 리소스 측면에서 큰 장점을 가진다. 하지만, 이후의 영상 전처리 과정이 CPU에서 이루어지는 알고리즘인 경우, 초기에 CPU 메모리에서 GPU 디코딩을 위해 cpu2gpu 메모리 복사과정과, 디코딩 이후 CPU에서 전처리를 위해 다시 gpu2cpu 메모리 복사 과정, 전처리 후 AI 객체 검출을 위해 다시 cpu2gpu 메모리 복사하는 과정들이 수반되는 비효율성이 발생하므로, GPU 디코딩 방식은 모든 알고리즘이 GPU에서 동작할 때 적용하기 적합한 방식으로 알려져있다. 동영상 디코딩을 통해 나오는 결과물은 YCbCr형태의 비디오 프레임으로, 일반적으로 YCbCr420 방식이 많이 사용되며 이 경우 밝기 정보에 해당하는 Y 채널은 Full 해상도를 가지는 반면 컬러 정보에 해당하는 CbCr 채널은 Y채널의 1/4 해상도를 가지는 3채널 영상으로 구성되어 일반적으로 사용되는 RGB 3채널 형식과는 다른 포맷을 가진다[3].

디코딩 이후의 전처리과정은 포맷 변환, Resizing, Cropping, Geometric 변환, 영상 전처리, 영상 분석등 여러 과정을 포함한다. 대부분의 인공지능 모델이 RGB 3채널 방식의 입력을 기반으로 하고 있으므로 디코딩된 YCbCr 포맷을 RGB 포맷으로 변화하는데, 일반적으로 YCbCr420 포맷을 YCbCr444 포맷으로 해상도 변화하는 과정을 거친 뒤, YCbCr 컬러스페이스를 3x3 매트릭스 연산을 통해 RGB 컬러 공간으로 변환을 한다[3]. 이때 일반적으로 사용되는 객체 탐지 모델이 Floating 연산을 하는 모델인지, 8bit 경량화 연산을 하는 모델인지에 따라 RGB 영상을 Floating FP32 형태라 정수형 Int8 형태로 선택 저장하여 처리하는데, 일반적으로 정수형 연산이 실수형 연산에 대해 정확도가 일부 떨어질 수는 있지만 빠른 시스템 처리 시간과 적은 메모리 공간을 차지하는 장점이 있다. 또한 Object detection 모델이 입력 영상의 크기에 따라 검출 성능과 처리 속도 그리고 소요 메모리에 대한 상관관계가 존재하여 이에 따른 최적의 영상 크기로 Resize 하여 모델로 입력되는 과정을 거친다. 이때 일반적으로 객체의 Shape 특성에 해당하는 영상의 Aspect 비율을 유지하여야 객체 검출 성능이 유지되므로, Resize 수행 시 영상 Padding이나 Cropping 과정이 같이 실행된다.

전처리 과정 중 컬러 변환이나 Resizing 등과 같이 필수적으로 거치는 픽셀 연산 외에, 객체 분류나 탐지의 성능을 올리기 위해 추가로 전처리 없이 수행된다. 이는 CCTV 탐지시스템의 한계상황을 극복하기 위한 것으로, 해당 한계사항들은 악천후로 인한 CCTV 카메라의 흔들림이나 PTZ 카메라와 같이 고정되지 않은 카메라의 Panning 중 촬영되는 객체에 대한 탐지 성능 저하, 카메라로부터 가까운 객체는 충분히 큰 픽셀로 표현되는 반면 멀리 떨어진 객체는 작은 픽셀로 표현되어 객체 탐지율이 저하되는 카메라 해상도 문제를 포함한다. 또한 카메라의 외적인 조건에 해당하는 주간/야간의 조도 차이와 일반 RGB 카메라와 IR 카메라의 특성 차이로 인한 객체 탐지 및 오검출 성능 저하 문제, 안개/비/눈과 같이 실제 카메라의 시인성을 크게 저하로 인한 미검출/오검출 문제 등이 존재한다. 객체 특성 측면에서는, 동일한 객체에 대해 객체가 움직일 때와 움직이지 않을 때 발생하는 검출력 차이 문제 및 카메라 촬영 화각의 차이로 인해 발생하는 객체 탐지 성능 저하에 대한 문제가 존재한다. 이러한 성능 저하를 야기하는 문제들은 그 모든 경우를 포함하는 학습데이타의 확보를 통해서도 일부 개선 가능하지만, 실제 각 경우별로 발생하는 환경적인 요인들을 역으로 해결하는 다양한 신호처리 기반 전처리 알고리즘의 조합에 의

해 성능을 개선하는데, 이에 대한 각 경우별 신호처리 기술과의 연동을 통한 다양한 성능개선 법들을 본론에서 다루기로 한다.

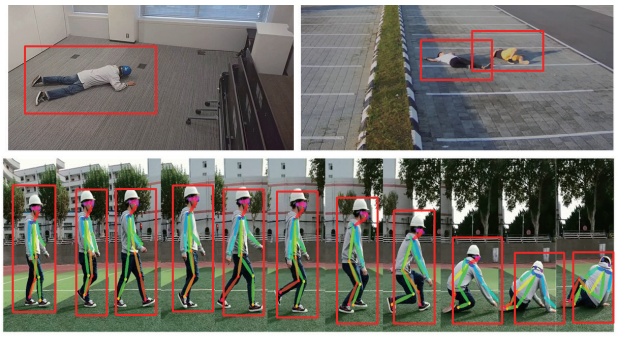
객체를 탐지하고 그 객체를 분류하는 방법은 크게 단일 stage detector 방법과 이중 Stage detector 방법으로 분류되는데, 이중 Stage detector 방식은 일반적으로 이미지 안에서 객체가 존재하는 영역(ROI)을 Bounding Box로 찾아 제안하는 Region proposal 방법을 거친 뒤, 제안된 후보 영역을 classificationn하는 방식으로 R-CNN 계열(R-CNN, Fast R-CNN, R-FCN, Mask R-CNN)과 같은 방식이 이에 속한다. 이때 후보 영역을 찾는 Region proposal 방식은 이미지 영역의 질감, 컬러, 휘도 차이 등을 갖는 인접 픽셀를 찾아 후보 영역을 제안하는 Selective search 방식과, 검출하고자 하는 입력 영상의 정해진 크기의 Bounding Box를 만들어 방향을 이동하면서 객체가 존재할 수 있는 영역을 추출하는 슬라이딩 윈도우 방식들이 사용되는데, 이 슬라이딩 윈도우 방식은 성능은 높은 장점을 가지나 전체 영역을 탐색해야 하므로 시간이 많이 소요되어 비효율적인 특성을 가진다. 이에 반해 단일 Stage 방식은 Feature를 추출하는 Convolution layer에서 Region proposal과 Classification을 동시에 수행되는 방식으로 SSD계열(SSD, DSSD, DSOD, RetinaNet)과 YOLO 계열들이 이에 해당하고, 이중 stage 방식에 비해 일부 검출 성능이 떨어질 수 있지만 빠른 처리 속도로 인해 현재 대부분의 객체 검출 및 분류 기술에 사용되고 있다.



〈그림 2〉 YOLO의 버전별 성능차이 비교

특히 가장 널리 사용되는 YOLO 방식은 2016년 V1이 나온 뒤로 Backbone 네트워크의 변화나 Loss 함수의 변화, 새로운 네트워크 구조의 추가를 통한 작은 객체들의 탐지 성능 개선, 파라미터 및 계산량 감소에 따른 처리 속도 개선 등을 통해 매년 새로운 버전들이 소개되고 있다. 객체 성능 탐지 성능을 보여주는 mAP(Mean Average Precion) 수치 또한 Yolo V2의 30.2%에서, Yolo V3의 57.9%, Yolo V5의 83.5%, Yolo V8의 86.4%까지 지속적으로 개선된 성능을 보여주고 있다. 또한 GPU에서 소요되는 딥러닝 네트워크의 파라미터 수와, 처리속도, 검출 성능에 대한 버전별 차이에 대한 상세한 성능 비교 분석이 〈그림 2〉와 같이 제공되어, 이러한 버전별 성능 및 속도 특성들을 참조하여 CCTV 지능화 시스템의 HW 사양 및 GPU 사양 등과 적용되는 모델 등을 결정하여 시스템을 구성하게 된다[4].

CCTV를 통한 주요 서비스가 사람, 고양이, 자동차, 건물과 같이 다양한 객체를 분류하는 작업도 있지만, 안전과 관련된 쓰러짐, 폭행, 군중/인파 밀집, 배회, 침입, 화재, 수위탐지와 같이 다양한 상황을 분류하는 작업들도 존재한다. 이미지 한 프레임을 보고 객체를 분류하는 방식은 영상을 학습하는 과정과 이미지를 추론하는 과정이 동일하며, 객체 탐지 모델 뒤에 후처리 알고리즘을 통해 추가 성능 개선이 일부 가능하여도, 분류성능의 대부분이 객체 학습 성능에 따라 전체 성능이 결정되므로 충분한 학습 데이터의 확보와 학습 방법이 가장 중요한 요소로 작용한다. 하지만 복잡한 상황이나 행동을 분류 및 인식하는 작업은, 일반적으로 매 프레임 단위로 수행되는 객체 분류 탐지 결과들을 통합 분석하여 현재 상황이나 행



〈그림 3〉 Box기반(위) 및 Skeleton기반(아래) 행동인식

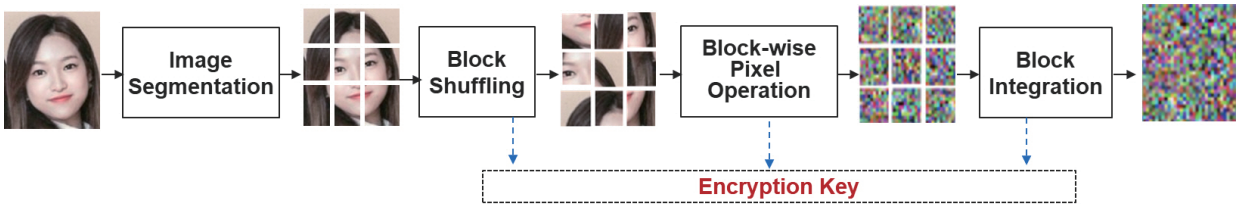
동을 분석하는 비디오 분석 방식을 사용하고, 객체 탐지를 위한 학습 성능 외에 매 프레임별 특징 들을 분석하여 판단하는 후처리 알고리즘이 수반된다.

예를 들어 사람의 쓰러짐의 행동을 탐지하는 시스템의 경우 여러 형태의 구현 방법들이 존재 하는데, 〈그림 3〉의 상단과 같이, 사람이 쓰러지는 다양한 형태의 영상을 Box 형태로 학습하여 쓰러짐 자체를 Box 형태로 추정하는 단일 프레임 기반 방법론을 사용하거나, 〈그림 3〉의 하단처럼 사람을 객체 탐지를 통해 Box 형태로 검출 후 검출된 객체의 Skeleton 정보 분석을 통해 Pose estimation을 하는 2단계 방식이 많이 사용되는데, 이 두 방법 역시 처리 속도와 판별 정확도 간의 상관관계가 존재하여 목표로 하는 시스템의 요구조건에 따라 상호 선택하여 사용된다[5]. 하지만 단일 정지영상 상태에서 쓰러짐 여부를 판단하는 방식만으로는 CCTV 카메라의 다양한 화각으로 인한 영상 모호성이 존재하여 오검출 확률이 높게 존재하고, 이를 개선하기 위해 복수의 프레임을 분석하여 사람이 쓰러지기 전과 후의 자세의 변화 부분을 탐지하거나, 쓰러진 후 사람이 움직이지 않는 정보들을 활용하는 후처리 알고리즘들을 연동하여 오검출 확률을 감소시키는 방식으로 구현된다.

객체를 탐지한 뒤, 해당 객체의 Pose를 추정하는 2단계 방식을 통해 쓰러짐과 폭행, 쓰레기 유기 등과 같은 다양한 행동을 인식하는 기술이 구현되는 것과 유사하게, 객체 인식을 통한 보안 시스템의 경우도 객체 탐지 알고리즘을 통해 자동차 번호판이나 사람 얼굴을 Box 형태로 검출 한뒤, 검출된 객체의 Box 영상을 별도의 자동차 번호판 인식 알고리즘이나 얼굴 인식 알고리즘을 통해 Identity 정보를 추출하는 기술들이 주차장 차량통제 시스템이나 보안시설의 사람 인식 및 통제 시스템 등에 널리 사용되고 있다. CCTV 객체 인식의 성능 또한 객체 탐지와 마찬가지로 먼 거리에 있는 작은 객체의 성능이 저하되는 문제가 존재하므로, 번호판이나 사람 얼굴을 인식하는 알고리즘 성능의 개선과 함께, 더 고해상도의 자동차 번호판이나 얼굴 영상이 인식 알고리즘에 입력될 수 있는 시스템적인 구성이나 전처리 기술이 같이 요구된다.

CCTV를 통한 객체의 탐지, 상황에 대한 탐지, 객체 Identity에 대한 인식과 함께, 얼굴이나 자동차 번호판과 같은 개인 Privacy 정보 보호를 위해 객체 영역을 알아볼 수 없도록 흐림이나 모자이크 처리와 같은 객체 비식별화(De-identification) 기술이 많이 사용되고 있고, 비식별화된 영상이 추가적으로 다양한 AI 학습에 사용되도록 객체 영역의 정보를 손실하지 않고 이미 정의된 임의의 얼굴이나 자동차 번호판으로 Swap하거나 AI기법을 통해 임의의 얼굴을 생성하여 Sawp하는 방식도 널리 사용되고 있다. 특히 비식별화된 원본을 향후 원본 비교 등의 목적으로 복원하는 경우를 위해 아래와 같이 객체의 분할된 영역을 Scrambling 하여 암호화한 뒤 다시 해당 정보를 기반으로 복호화 하는 방식의 비식별화 기법도 비식별화의 복적에 따라 선별적으로 사용되고 있다.[6]

CCTV 지능화 시스템의 핵심 기능 중의 하나인 객체 추적 기술은 범죄 용의자와 같이 특정 객체를 추적하거나, Privacy가 요구되는 영상에서 선별적으로 비식별화에서

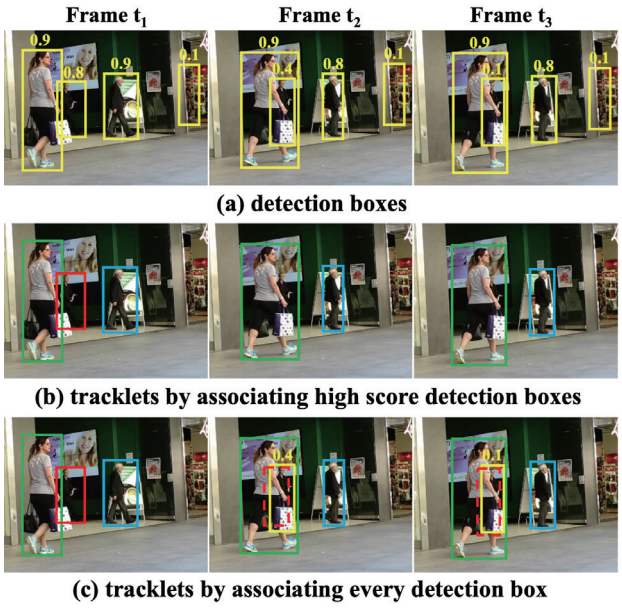


〈그림 4〉 원본 복원이 가능한 비식별화 방식

〈표 1〉 CCTV 지능화 시스템 구성 요소 기술

기술 분류	세 부 기 술	기 술 개 요
영상 전처리 기술	동적 배경 추출기술	영상내 배경을 분리하여, 움직이는 객체만을 동적 검출하는 기술
	카메라렌즈 보정기술	어안렌즈와 같은 특수렌즈 촬영 영상을 일반 형태로 변환하는 기술
	카메라 흔들림 보정기술	카메라의 흔들림을 보정하는 Video stabilization 기술
	영상 화질 개선기술	카메라 영상의 노이즈 제거, 해상도 향상, Contrast 개선등을 통해 고화질/고시인성의 영상으로 변환하는 기술
객체탐지 및 분류기술	객체 탐지 기술	사람, 동물, 차량등 관심있는 객체를 탐지하는 기술
	객체 분류기술	탐지된 객체를 상세한 Class로 분류 및 구분하는 기술
	객체 인식 기술	자동차 번호판 및 사람 얼굴등을 인식하는 기술
객체 추적기술	단일 카메라 추적기술	단일 카메라에서 특정 개체를 연속적으로 추적하는 기술
	다중 카메라 추적기술	특정 객체를 다중 카메라에서 연속적으로 추적하는 기술
	Re-identification 기술	자세/의상/주변 조명환경 변화에 강인한 특정 객체 탐지 기술
영상 보안기술	프라이버시 보호 기술 (De-identification 기술)	영상 내 얼굴/자동차 번호판과 같이 개인정보에 민감한 영역을 비식별화 하고 복원하는 프라이버시 마스킹 기술
	보안 침해 방지 기술	영상 센서의 해킹 방지 및 암복호화, 위변조 방지 기술

제외하는데 사용되는 것으로 트래킹 성능과 처리 시간 개선 등에 많은 연구가 이루어져 오고 있다. 일반적인 트래킹 방법은 〈그림 5〉 b와 같이 낮은 Score를 가진 객체를 제거하여 True positive / False positive trade-off 문제를 다루는 방식으로 이때 가려진 객체들을 탐지하지 못하는 단점이 존재한다.



〈그림 5〉 검출 Box score와 객체 추적 성능 비교

이에 반해, 최근 소개된 ByteTrack 알고리즘은 〈그림 5〉 c와 같이 Score가 낮더라도 detection box를 모든 Tracklet들과 연결하여 IOU 정보와 Kalman filter를 사용하여 모든 객체를 추적할 수 있도록 하여 높은 다중 객체 추적 성능을 보여주는 것으로 알려져 있다[7].

또한 객체 추적 및 검색 시스템의 일환으로 인적자의 정보가 담긴 사진을 기반으로 실종자를 검색하는 시스템이나, 이를 배경과 주변 환경이 다른 다중 CCTV에서 동일한 사람을 찾기 위해 객체의 옷이나 자세, 카메라 각도, 조명의 변화에도 동일한 객체를 인식하는 객체 재인식(Re-identificaion) 기술이 검출된 객체의 동일한 Box 내 정보를 따라 추적하는 Tracking 기술과 함께 시스템에 혼용되어 사용된다.

III. CCTV 지능화 시스템 성능 개선 기술

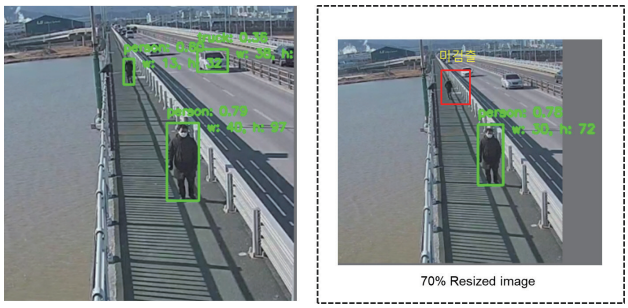
CCTV 지능화 시스템의 기본적인 알고리즘들의 구성과 함께, 고정형/PTZ 이동형/어안렌즈 카메라와 같이 다양한 CCTV 카메라의 종류와, 야외/실내/도심/자연등과 같

은 다양한 카메라 설치 환경 및 운용 환경, 지능화 시스템의 조건 등에 따라 많은 성능 제한 요소가 발생하고, 〈표 1〉에서와 같이 각 성능 제한 요소들을 해결하기 위한 여러 요소 기술이 기본 알고리즘과 함께 사용되고 있다. 이 장에서는 CCTV가 가지는 기본적인 성능 제한 요소들을 영상 요소 관점(해상도, 시인성, 기하학적)에 따라 분류하고, 각 요소별 한계를 해결하기 위한 기존의 많은 신호처리 기반 처리 방법들과 영상 분석기법들을 응용한 전처리 기술들, 시스템 플랫폼 관점에서 여러 알고리즘을 구성하는 방법들을 소개하고, 또 이를 하나의 고속 병렬 시스템으로 구성하기 위해 실제로 고려되는 부분들을 소개하기로 한다.

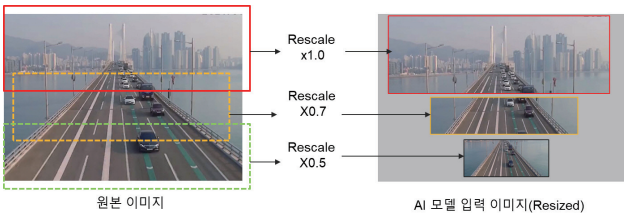
1. 객체 크기와 해상도에 대한 성능 개선

일반적으로 CCTV 지능화 시스템의 성능은 객체 탐지 및 분류 정확도에 의해 표현되지만, 동시에 CCTV의 환경적인 요소 및 객체 특성 조건에 대해 정확도가 같이 표현된다. 고정 CCTV가 가지는 근본적인 요소 중 하나로, CCTV로부터 어느 정도의 거리가 떨어져 있는 곳까지 객체 탐지 성능이 유지되는지가 성능 지표로 정의되기도 하는데, 이는 CCTV 카메라에서 멀리 떨어져 있는 객체일수록 객체 크기가 작게 표현되어 객체 검출 정확도가 저하되는 특성과 연관되고, 이에 따라 일정한 성능을 유지하기 위한 CCTV 카메라의 필요 개수와 배치 위치도 같이 결정된다.

거리에 따른 객체 크기의 차이는 실제 객체가 차지하는 Box의 픽셀 크기에 따라 정의되므로, 이 부분은 AI 네트워크 측면에서 입력되는 이미지 크기와 관련되고, 이미지 Resizing 크기와 이에 따른 GPU 소요 메모리의 크기 및 검출 정확도, 검출 소요 시간이 모두 연동되어 시스템 파라미터 구성시 중요한 요소 중 하나로 결정된다. Resizing에 의한 성능 관점의 해상도 최적화 외에 고해상도 입력을 객체 검출 및 인식 AI 모델에 넣기 위한 다양한 방법들이 제안되고 있다. 〈그림 7〉에서 보는 것과 같이 CCTV에서 가까운 영상의 하단 부분의 객체는 크게 촬영되는 반면, CCTV에서 먼 영상의 상단 부분의 객체는 작게 촬영된다는 점을 이용해서, 이미지를 Resizing 할 때 영역에 따라 Resizing scale을 다르게 하여 영상 구성을 한 뒤 AI모델에 입력하는 방법이 사용된다. 이때 관심 영역별로 다른 Scale을 적용하여 다양한 Patch 형태로 AI 모델 입력 영상을 구성하는 기법이 사용되고, 중첩 영역에 대한 후처리가 같이 수행되지만, 높은 정확도에 비해 작은 계산량이 요구되어 효율성 면에서 많이 사용된다.

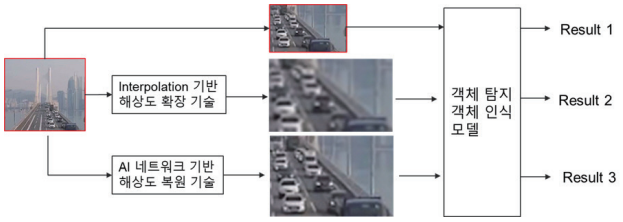


〈그림 6〉 Resizing/거리/객체크기에 따른 검출성능 차이



〈그림 7〉 Multi-rescaling AI 모델 입력

하지만 원본영상을 그대로 사용하는 경우에도 CCTV에서 멀리 위치하여 작은 객체로 표현되는 경우에 대해서는 추가 해상도 Upscaling이 요구되는데, 일반적으로 사용되는 Bicubic이나 Bilinear, Neeast neighbor와 같은 Interpolation 기반 알고리즘이 가지는 성능 한계를 넘어서기 위해 〈그림 8〉과 같이 딥러닝 모델에 기반한 Super-resolution 전처리 기술을 통한 고해상도 결과 영상을 객체 탐지 및 인식 모델에 입력하는 방법도 사용된다. Super-resolution 기술은 일반적으로 해상도를 4배 향상시켰을 때 원본 대비 유사성(SSIM)이나 PSNR 기준으로 성능 평가가 이루어지는데, CNN 네트워크 구조 기반의 SRCNN 방식으로부터 시작하여, GAN 방식의 SRGAN 네트워크 등 다양한 네트워크 구조와 네트워크 레이어, 전처리 방식의 결합 등을 통해 원본 복원 성능과 처리속도 부분에서 성능이 향상되어 객체 탐지된 기술들이 사용되고 있다[8, 9].

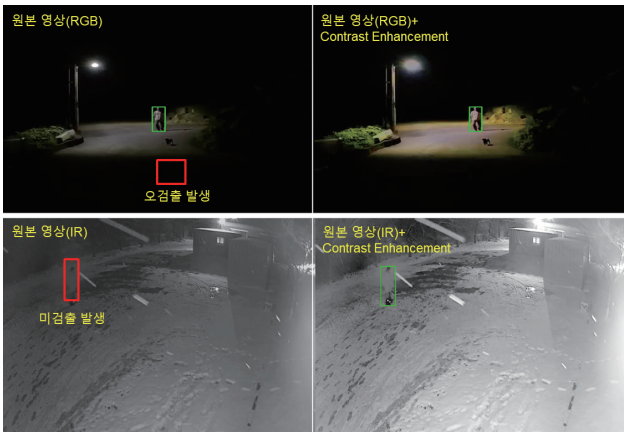


〈그림 8〉 Upscaling 전처리에 따른 객체 탐지 성능개선

〈그림 8〉에서의 다른 해상도 처리 방법에 대해, Bicubic과 같은 Interpolation 기반의 Upscaling 기법들은 실제 정보를 복원하는 성능이 제한적으로, Result 1과 Result 2가 유사한 성능으로 객체 탐지에 큰 효과를 주지 못하는 반면, Super-resolution기반의 Result 3의 결과는 더 높은 객체 탐지 성능을 보여준다. 하지만 영상 전체를 Super-resolution 전처리를 하는 것은 실시간 처리 시간 측면에서 비효율적이어서, 입력 영상은 〈그림 7〉과 같은 Multi-resolution 기법들을 통해 효율적으로 Resizing 하여 객체 탐지 성능을 확보하고, 크기가 작은 객체의 Box 형태 영상에 대해 선별적으로 Super-resolution 전처리 후 객체 인식을 하는 방식으로 고성능 확보와 동시에 실시간 처리 성능을 유지하는 방식들이 사용된다

2. 시인성 저하에 따른 객체탐지 저하 성능개선

일반적으로 객체 시인성이 높은 맑은 날씨와 조명이 충분한 낮 시간에는 CCTV지능화 시스템의 객체 탐지 및 분류 성능이 높은 반면, 사람의 육안으로도 분간하기 힘든 객체 시인성이 낮은 야간이나 악천후 및 안개가 있는 경우는 CCTV의 객체 탐지 성능이 현저하게 저하가 된다. 이를 위하여 악천후에 해당하는 많은 학습 데이타 추가를 통해 성능 개선이 이루어지기도 하지만, 학습 데이터가 충분히 주어져도 야간 또는 안개가 있는 상태의 영상이 시인성 높은 맑은 날씨의 낮에 촬영된 영상보다 낮은 객체 탐지율을 가지는 한계를 가진다. 따라서 영상의 시인성을 향상시키는 전처리 기법이 많이 사용되는데, 영상처리 측면에서는 영상 신호처리를 통한 Global contrast 및 Local contrast를 향상시키는 기법들이 사용되고, 유사한 효과를 내는 딥러닝 네트워크 기반의 High Dynamic Range (HDR)기법도 많이 사용된다[10].



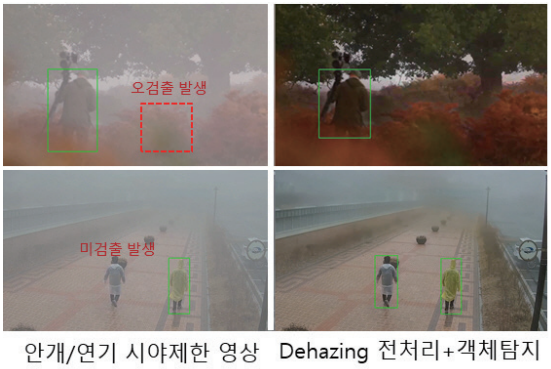
〈그림 9〉 시인성 개선 알고리즘 적용 전/후에 따른 객체 탐지 성능 영향 비교

야간 CCTV 시스템의 경우, 가로등이 있거나 조명이 충분히 밝은 경우는 일반적인 RGB카메라 모드의 CCTV로도 충분히 객체 시인성이 확보되어 낮과 객체 탐지 성능이 크게 나지 않는다. 하지만, 〈그림 9〉에서와 같이 조도가 낮은 환경에서는 RGB 카메라 모드로는 육안으로도 실제 사물을 분간하기 힘든 경우 객체 탐지율도 낮아 오검출이나 미검출이 많이 발생하고, 야간 모드에 동작하는 적외선 방식의 IR 카메라는 RGB 카메라의 Gray 모드 영상과는 그 특성이 달라서 AI 학습을 위해 IR 카메라에 의해 촬영된 충분한 학습 데이터가 요구된다. 따라서 낮은 저조도의 RGB 카메라 영상과 IR 카메라의 시인성 개선을 위해, 화질 개선 알고리즘의 적용을 통한 객체 탐지 성능 개선 방법론이 사용된다. 많이 사용되는 알고리즘 중 하나는 이미지 전체의 히스토그램 분포를 균등하게 변환하는 Histogram Equalization(HE) 방식인데, 이 방법은 프레임 전체의 픽셀 분포에 기반하여 Global contrast를 개선하는 방식으로, 프레임 내에 어두운 영역이 많이 있거나 밝은 하이라이트 영역이 많은 경우 그에 따른 영상의 변화가 프레임간에 급격하게 변하는 문제를 가진다. 따라서 어두운 영상에 HE 적용시 암부 영역의 노이즈가 많이 발생하여 프레임간의 결과 영상이 많이 다르게 나올 수 있고, 이에 따라 객체 탐지 결과도 크게 영향을 받으므로 실제 적용에는 많은 한계가 따른다.

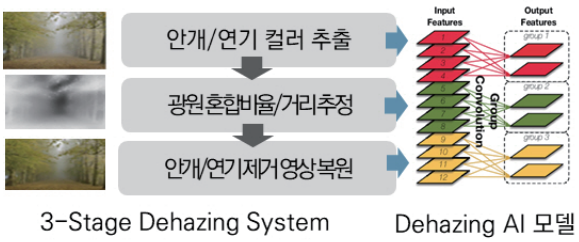
이를 보완하려는 방법으로 프레임을 블록 단위로 나

〈그림 10〉 Contrast enhancement 적용 전/후 결과비교

누어 HE하는 CLAHE 방법이 많이 사용되는데 이는 블록 단위로 Local contrast를 향상하는 효과를 가지는 알고리즘으로, 〈그림 10〉에서 보는 것처럼 영상 전체가 크게 변하지 않으면서 암부와 하이라이트 영역의 객체 시인성이 향상되는 효과를 볼 수 있다[11]. CLAHE 방법은 블록의 크기가 작아져서 프레임내의 블록 수가 많아질수록, 각 블럭별 Contrast의 강도가 크게 올라가는 효과를 가지지만 동시에 여러 블록을 처리해야 하므로 영상 처리 시간이 늘어나게 된다. 하지만 최근 Nvidia에서 제공하는 CUDA기반 병렬처리 방식의 CLAHE가 구현되어 GPU에서의 실시간 영상 전처리가 가능해져서, 프레임별 CLAHE 전처리 및 객체 탐지 분류 모델 실시간 처리를 통해 야간 영상에서 객체 탐지 및 분류 성능을 향상하는 구조로 많이 사용되고 있다.



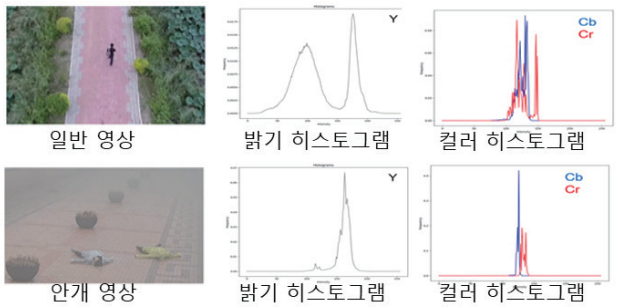
〈그림 11〉 Dehazing 전처리 적용 전/후의 결과 비교



〈그림 12〉 Dehazing 전처리 알고리즘 구조

눈과 비가 내리는 날씨의 경우 객체 탐지의 성능이 일부 저하 되기는 하지만 일반적으로 학습을 통해 개선이 가능하다. 하지만 안개가 존재하는 경우는 〈그림 11〉에서 보는 것처럼 실제 사람의 육안으로도 객체를 인지하기 어려운 상태가 발생하여 이 경우 CCTV에 의한 객체 탐지 성능도 크게 저하되어, 이 부분은 학습 영상의 추가만으로는 성능 향상에 한계를 가진다. 이를 위해 오랜 기간 연구되어 온 Dehazing 기법은 일부 영상에만 효과가 있고 영상 내 Artifact가 많이 존재하는 한계가 있어 왔는데, 최근 AI 모델을 통한 Dehazing 기법이 큰 성능 향상을 가져오면서 실제 영상 전처리 기법으로 많이 사용되고 있다.

Dehazing 알고리즘은 일반적으로 〈그림 12〉와 같이 안개나 연기의 컬러를 추출하고, 그에 따른 광원 복원 정보를 추정하여 해당 정보에 기반하여 Dehazing AI 모델을 통해 안개와 연기를 제거하는 과정을 거친다. 하지만 AI Dehazing 알고리즘이 GPU에서 처리되어도, 상대적으로 긴 처리시간을 가지므로 모든 프레임에 대해 처리하지 못하고 Scene 분류 기술을 통해 안개가 존재하는 프레임에 대해서만 Dehazing 알고리즘을 적용하여 실시간 처리를 하는 방식이 사용된다. 이때 안개 영상을 분류하



〈그림 13〉 Scene 분류를 위한 밝기/컬러 정보 비교

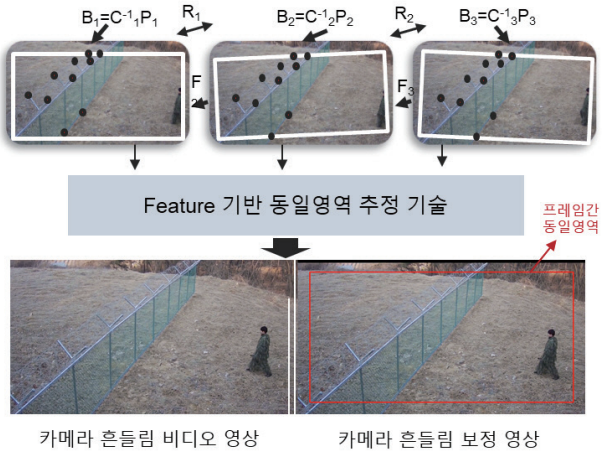
는 지표로 〈그림 13〉과 같이 전체 프레임의 밝기와 컬러 히스토그램을 분석하여, 안개가 낀 영상은 특정 밝기의 픽셀들이 분포한다는 점과, 안개로 인해 전체 영상의 컬러 채도가 낮아지면서 특정 컬러 영역에 컬러가 분포한다는 특성을 사용하여 낮은 계산량으로 Scene을 구분하는 방식이 많이 사용된다.

3. 카메라 Geometry 변화 보상을 통한 성능 개선

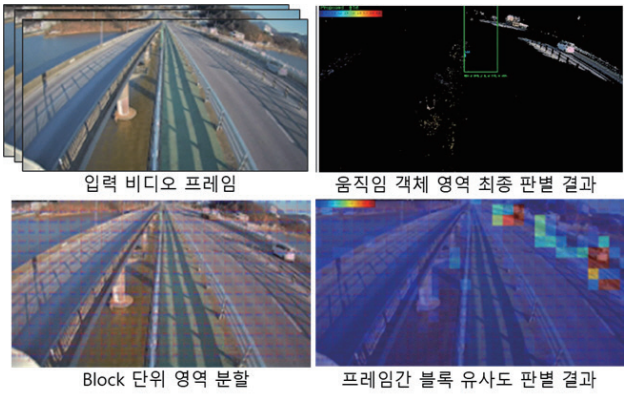
CCTV 지능화 시스템에서 오검출을 유발하는 가장 큰 요소 중 하나로, 동일한 객체에 대해서 카메라가 고정되어 있을 때와 악천후 및 바람으로 인해 카메라의 흔들림이 크게 발생할 때 객체 탐지 및 분류 성능이 크게 차이가 발생하는 문제가 발생한다. 일반적으로 악천후의 경우에 발생하므로, 실제 문제 발생 원인과 해결책을 악천후 상황의 학습 데이터를 더 확보하여 객체 탐지 성능을 향상하는 방향으로 진행하는 경우가 많이 발생한다. 하지만 기본적으로 프레임간의 비교를 통해 객체의 움직임 여부를 판단하여 객체 탐지 및 분류를 하는 Background subratraction과 같은 객체 움직임 판단 알고리즘은 거의 동작하지 못하는 상황이 발생한다. 이에 대한 해결 방법으로 카메라의 흔들림 문제를 근본적으로 해결하는 방식의 알고리즘을 사용하거나, 카메라의 흔들림이나 밝기 변화가 존재하는 상황에서도 강인한 객체 움직임 판단 알고리즘이 사용되고 있다.

카메라의 흔들림을 보상하는 방식으로, 핸드폰이나 캠코더 등에서 많이 사용되는 Video Stabilization 기법들이 활용된다. 이는 〈그림 14〉에서와 같이 영상 프레임의 특정 특징점을 추축하여 프레임간 특징점 매칭을 통해 Homography transformation이 수행되는데, 일반적으로 카메라가 고정된 상태에서 좌우로 흔들리는 경우가 많으므로 3x3 matrix 형태의 Affine transformation 보정만으로도 카메라의 흔들림을 보상할 수 있다. 이때 실시간 Stabilization 구현을 위해, 영상의 특징점 추출은 딥러닝 네트워크를 통한 방법을 통해 GPU상에서 실시간 처리를 하고, Affine transformation의 3x3 matrix 연산은 CPU를 통해 수행하거나 CUDA 병렬처리를 통해 수행하여 고속 Stabilization 전처리를 구현하는 방식이 가능하다.

객체의 움직임 여부와 배경을 판별하는 Background subtraction 알고리즘은 프레임간의 움직임과 조명조건의 변화로 인한 작은 픽셀값의 변동이 성능에 큰 영향을 받는다. 또한 Background subtraction 알고리즘은 매 프레임 작동해야 하는 실시간 처리 능력과 높은 성능이 동시에 요구되므로, GPU에서 CUDA 병렬처리 적용이 가능한 알고리즘 구조를 통해 실제 CCTV 시스템에 적용된다. 이를 위해 프레임과 프레임간의 픽셀 차이를 보는 방식보다는 〈그림 15〉와 같이 프레임을 복수개의 Block으로 구분하고, 해당 Block 간의 유사도를 보는 방식을 통해 픽셀 위치 변화와 픽셀값 변화에 덜 민감하게 하고, Block 간의 유사도를 판별하기 위한 방법으로 Block variance를 사용하여 프레임간의 전체 밝기 변화에도 영



〈그림 14〉 카메라 흔들림 영상 보정 기술



〈그림 15〉 Block 분석 기반 동적 객체 판별 알고리즘

Fisheye 렌즈 영상　Fisheye 영상 Dewarping 처리 결과

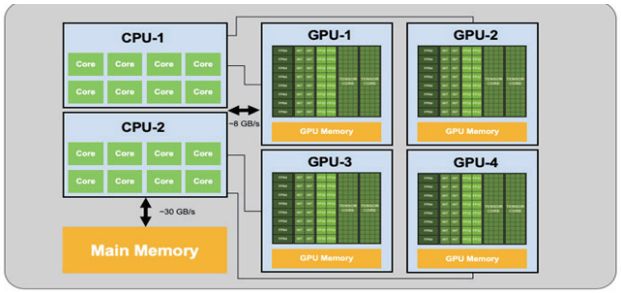
〈그림 16〉 어안렌즈 영상 보정 기술

향을 적게 받게 하면서 동시에 CUDA 구현을 통해 고속 처리가 가능하도록 하는 방법이 사용될 수 있다.

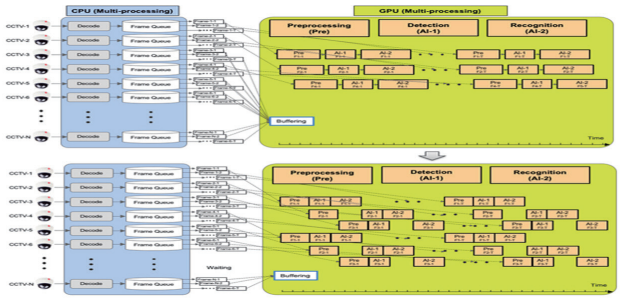
일반 지향성 CCTV외에 어안 카메라(Fisheye Lens)가 사용되는 경우, 지향성 CCTV에서는 일반적으로 사람과 같은 객체가 머리가 위쪽방향으로 탐지되는 일관적인 방향성을 가지는 것에 비해 Fisheye 카메라는 객체의 머리 위쪽에서 촬영되는 특성으로 인해 객체의 방향이 여러 방향으로 형성된다. 이런 특성으로 인한 학습의 한계를 보완하기 위해, 촬영된 Fisheye 영상을 많은 각도로 Rotation 하여 Augmentation을 한 뒤 학습을 하는 방식이 많이 사용되지만, 지향성 CCTV 카메라 대비 현저하게 낮은 객체 탐지율을 가지는 한계를 가지게 된다. 이러한 기하학적인 한계를 극복하기 위하여 〈그림 16〉과 같이 Fisheye 영상을 Dewarping 하여 객체의 방향성을 고정형 CCTV와 유사하게 한 뒤 객체 탐지 및 분류를 하는 기법이 많이 적용되고, 이때의 Dewarping 알고리즘도 AI 모델기법을 통해 GPU에서 고속 처리될 수 있도록 하여 실제 실시간 시스템에서 객체 탐지의 전처리 모듈로 같이 사용된다[12].

4. Multi-AI 모듈의 실시간 처리 성능 개선

실제 CCTV에서 발생하는 많은 성능 한계를 극복하기 위해 다양한 전처리 기법들이 소개되었지만, 실제로 이 모든 모듈을 실시간 처리하는 방식은 시스템 플랫폼 구조로 인한 한계가 존재하고, 이를 위한 고속 병렬 처리기법이 요구된다. 가장 큰 요소로는, 〈그림 17〉에서 보는 것처럼 CPU와 GPU간의 데이터 전송으로 인한 병목이 최소화 되도록 하는 인터페이스 최적화 기법과, 영상이 CPU 메모리로부터 GPU메모리로 전송된 후 모든 전처리와 객체 탐지 분류의 모듈이 모두 GPU에서의 단일 파



〈그림 17〉 메인 메모리와 GPU 메모리간 인터페이스



〈그림 18〉 다채널/멀티 AI 모듈에 대한 고속 최적화

이프라인을 따라 처리되도록 하는 GPU 모듈 파이프라인 기술이 요구된다.

또한 〈그림 18〉과 같이, 동시에 여러 채널의 CCTV 입력에 대해, 여러 개의 전처리 모듈과 객체 탐지 모듈들이 고속 병렬처리 되도록 영상 프레임과 각 모듈을 Queuing 관리하고 최종 Delay가 최소화 되도록 하는 스케쥴링 기반 플랫폼을 구성하는 부분이 중요하다.

IV. 전망과 결론

주요 CCTV 관제시스템의 일반적인 운용 현황은, 넓은 벽면에 설치된 여러 개의 디스플레이를 통해 한 사람당 5~10개의 모니터의 내용을 육안으로 관찰하면서 범죄와 안전 등에 관련된 사항들을 관제하는 단계로부터, 한 사람이 10~30개의 모니터에서 AI 시스템에서 감지하는 여러 이벤트의 결과를 참조하여 모니터링 하는 방식으로 진화하고 있다. 이와 함께 사람이 하는 일을 AI가 대체하는 효과와 함께 사람이 놓칠 수 있는 작은 이벤트에 대해서도 AI를 통해 놓치지 않는 비약적인 기술 발전을 경험하고 있다. 하지만 다양한 CCTV가 설치된 다양한 물리적

인 환경 차이와, 날씨나 조도와 같은 카메라 촬영 조건, 객체들의 움직임이나 행동 차이에 따른 탐지조건의 차이 등으로 인해 발생하는 미검출 오검출 등과 같은 AI 엔진의 한계가 해결되지 못하는 현황에서는 추가적인 AI 시스템의 확장이 제한적일 수 있는 한계를 지니고 있다. 기본적으로 AI 시스템은 학습이라는 가장 중요한 요소에 의해 성능이 결정되지만, 이외에도 다양한 영상 분석 기술이나 전처리 신호처리 기술들을 통해 성능 저하 발생 요소들을 해결하고 있다. 특히 가까운 미래의 CCTV 지능화 시스템의 발전 요소들로 최근 생성형 AI 기술이 CCTV 지능화 시스템과 접목되고 있다. 이 중 하나로 AI 시스템이 요구하는 학습데이터의 부족 문제와 국방과 같은 제한된 환경에서의 학습데이터를 해결하기 위해 생성형 AI 기법을 통한 학습데이터 영상 생성 기술들이 국방과 다양한 산업현장에서 적용되고 있다, 또한 최근 비약적인 LLM 및 VLM(Vision Language Model)의 발전으로 인해, 영상의 Context를 파악하고, QnA가 가능하면서, 또 이미지와 텍스트의 Multi-modal 검색이 가능한 기술들이 시스템에 적용되고 있다. 이를 통해 1세대의 CCTV 육안 모니터링에서, 이벤트 검출 위주의 2세대 CCTV지능화 시스템으로 진화된 현재를 지나, 실제 영상 내의 객체간 관계 및 상황을 설명 및 요약해주고 QnA가 가능한 3세대 CCTV 지능화 시스템으로 진화가 계속되고 있어서, 이에 대한 관련 기술들의 개발과 함께 생성형 AI를 시스템화하는 기술들의 발전도 지속될 것으로 예상된다.

참고문헌

[1] www.aihub.or.kr

[2] "지능형 CCTV 성능시험인증", www.kisa.or.kr

[3] "BT.709: Parameter values for the HDTV standards for production and international programme exchange", www.itu.net, 1993.

[4] Degang Yang et al. "PGDS-YOLOv8s: An Improved YOLOv8s Model for Object Detection in Fisheye Images", Applied Sciences 2024, 14(44)

[5] W. Chen et al. "Fall Detection Based on Key Points of Human-Skeleton Using OpenPose", Symmetry", 2020, 12(5)

[6] K. Madono et al. "Block-wise Scrambled Image Recognition Using Adaptation Network", AAAI 2020

[7] Y. Zhang et al. " ByteTrack: Multi-Object Tracking by Associating Every Detection Box", ECCV 2022

[8] C. Dong et al. "Image Super-Resolution Using Deep Convolutional Networks", ECCV 2014

[9] C. Ledig et al. "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network", CVPR 2017

[10] L.Wang et al. "Deep Learning for HDR Imaging: State-of-the-Art and Future Trends', TPAMI 2022

[11] J.P.Eriksen et al. "MAHEM: a Multiprocessor Engine for Fast Contrast-Limited Adaptive Histogram Equalization", SPIE 1990

[12] A. Konrad et al. "FisheyeSuperPoint: Keypoint Detection and Description Network for Fisheye Images", VISIGRAPP 2021

민병석

- 2001년 2월 연세대학교 전기공학과 학사
- 2003년 2월 연세대학교 전기전자공학과 석사
- 2009년 8월 미국 Purdue대학교 전기컴퓨터공학과 박사
- 2009년 6월 ~ 2015년 11월 삼성전자 DMC연구소 책임연구원
- 2015년 11월 ~ 2020년 3월 삼성전자 VD사업부 수석연구원
- 2020년 3월 ~ 2024년 1월 ㈜자비스 연구소장
- 2024년 1월 ~ 현재 ㈜마크애니 연구소장

〈관심 분야〉
Computer vision, Image processing, Industrial AI Application, Image forensic application.

YOLO 아키텍처 기반의 자기증류를 활용한 객체 탐지 성능 향상

이동석
한화시스템

김해문
한화시스템

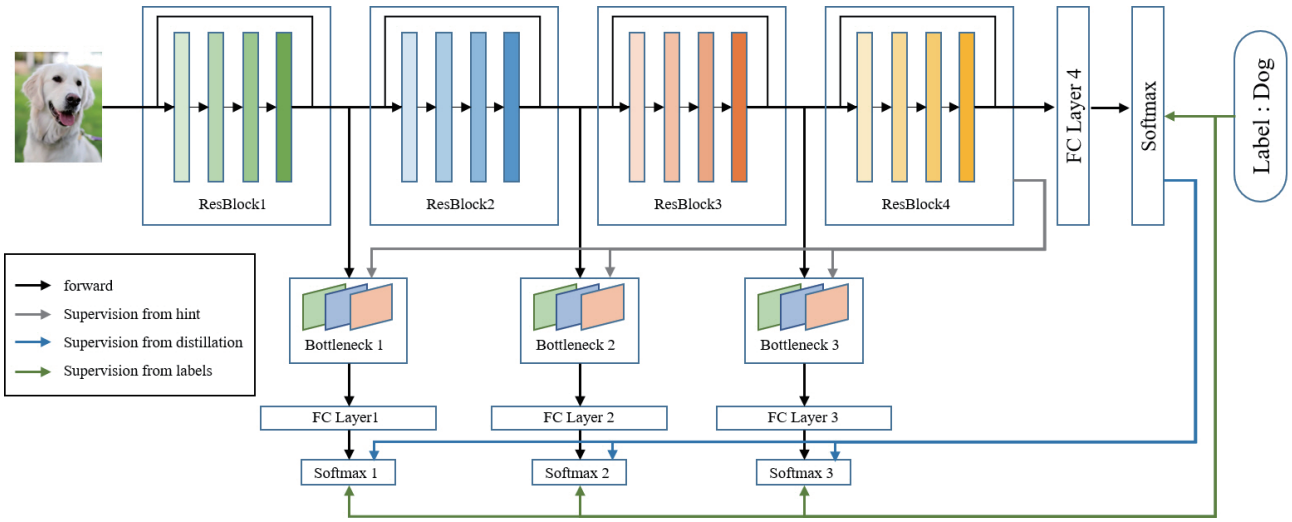
Ⅰ. 서 론

딥러닝 분야에서 지식 증류(Knowledge Distillation)는 대규모의 복잡한 모델(이하 '교사')에서 소규모 모델(이하 '학생')로 지식을 전달하여 모델 성능을 향상시키는 기법이다. 이 기술은 처음에는 Hinton et al.[1]에 의해 제안되었으며, 대규모 모델의 비효율성과 높은 계산 비용 문제를 해결하기 위해 고안되었다. 큰 모델을 직접 배포하는 대신 지식을 소규모 모델로 증류함으로써, 성능 저하를 약간 감수하더라도 보다 효율적인 대안을 제공한다. 교사 모델이 생성한 소프트 라벨은 클래스 분포에 대한 더 풍부한 정보를 포함하고 있어, 단순히 정답만 학습하는 것에 비해 학생 모델이 더 잘 일반화할 수 있도록 해준다.

이 기술은 이미지 분류(image classification)[2]뿐만 아니라 자연어 처리(natural language processing)[3]와 객체 탐지(object detection)[4] 같은 다양한 응용 분야에서도 계산 비용을 줄이고 모델의 크기를 줄이기 위해 널리 사용되고 있다.

그러나 지식 증류(KD)가 출력 로짓(output logits)에만 의존하는 경우, 더 복잡한 작업에서 효과가 제한될 수 있다. 이는 교사 모델의 출력 계층에서 생성된 로짓 만을 통해 지식을 전이하는 방식이 중간 계층에 존재하는 중요한 특징 표현(feature representation)을 충분히 활용하지 못하기 때문이다. 중간 계층은 저수준부터 고수준까지 캡슐화 된 계층적 정보를 포함하며, 이는 객체 탐지(object detection)나 의미론적 분할(semantic segmentation)과 같은 복잡한 작업에서 성능을 향상시키는 데 필수적이다.

이러한 한계를 극복하기 위해 피처 증류(Feature Distillation, FD)가 도입 되었다. FD는 교사 모델의 중간 계층에서 추출된 피처 표현을 학생 모델로 전달하여 KD의 한계를 극복한다. 중간 계층은 다양한 스



〈그림 1〉 이 그림은 Zhang [9]이 제안한 자가 증류 프레임워크를 보여준다.
이는 각 병목 계층(bottleneck layer)에서 특징을 추출하여 교사와 학생 간의 지식 전이 과정을 시각화한 것이다.

케일 및 입력 데이터의 미세 구조에서 중요한 정보를 포함하고 있어, FD는 더욱 복잡한 작업에서 우수한 성능을 제공한다. Romero et al.[5]는 FitNets 접근법을 제안하여, 교사 모델의 계층에서 추출된 피처를 학생 모델이 학습하도록 함으로써 학생이 네트워크의 더 깊고 복잡한 지식을 효과적으로 활용할 수 있도록 했다. 마찬가지로, Zagoruyko et al.[6]는 교사와 학생 모델 간의 주의(attention) 맵을 정렬하여, 학생이 교사로부터 중요한 공간 정보를 학습할 수 있도록 하여 성능을 향상시켰다.

FD는 객체 탐지 및 의미론적 분할과 같은 복잡한 작업에서 특히 강점을 보인다. 이러한 작업에서는 공간적 및 맥락적 이해가 중요하며, 중간 계층에서 추출된 특징 표현이 이 정보를 풍부하게 만들어 출력 로짓 전송만으로는 달성할 수 없는 성능 향상을 이끈다. 예를 들어, Wang et al.[7]은 객체 탐지 모델에서 특징 피라미드와 객체 특화 표현을 증류하여 탐지 정확도를 향상시켰고, Li et al.[8]은 다양한 데이터셋에서 다중 스케일 특징을 정렬하여 의미론적 분할에서 성능을 개선했다.

그러나 교사–학생 모델 접근 방식의 주요 단점은 긴 훈련 시간과 그로 인한 자원 소비이다. 이는 교사 모델의 크기와 복잡성으로 인해 훈련 시간이 길어지며, 모델 훈련이 교사의 성능에 제약을 받을 수 있기 때문이다. 만약 교사 모델이 충분히 훈련되지 않으면, 학생 모델도 성능 저하를 겪을 수 있다. 이러한 문제를 해결하기 위해 자가 증류 (Self–Distillation)가 도입되었으며, 이는 더 깊은 계층을 교사로, 더 얕은 계층을 학생으로 지정하여 단일 모델 내에서 지식을 전달한다. 이를 통해 외부 교사 모델 없이도 지식 전달이 가능하며, 일반화와 훈련 효율성을 개선할 수 있다.

Zhang et al.[9]은 CNN 모델에 자가 증류(Self–Distillation)를 적용하여 이미지 분류 성능에서 큰 향상을 달성했다. 이 접근 방식은 네트워크를 깊이에 따라 나누어 하나의 교사와 여러 학생을 정의하고, 정확도를 향상시키기 위해 교차 엔트로피 손실, KL 발산 손실, L2 손실과 같은 다양한 손실 함수를 적용한다.

객체 탐지에서는 자가 증류가 YOLOv6 모델[10]에 적용되어 성능 향상을 이끌었다. 이 접근 방식에서는 동일한 교사 모델과 학생 모델이 사용되며, 학생 모델은 사전 훈련된 교사 모델로부터 지식을 전달받는다. 두 모델은 동일한 아키텍처를 공유하며, 지식 전달은 사전 훈련된 교사 모델에서 학생 모델로 이루어지기 때문에 자가 증류라고 불린다. 그러나 교사 모델을 사전 훈련해야 하므로 훈련 시간이 두 배로 늘어나는 것은 기존의 교사–학생 모델과 다르지 않은 단점이다.

본 원고에서는 YOLOv6[10]의 자가 증류 방법의 한계를 다루고 단일 모델에서의 간단한 자가 증류 방법을 통한

두 가지 기여를 제안한다.

첫째, Zhang et al.[9]의 이미지 분류에 처음 적용된 자가 증류 접근 방식을 YOLOv6 모델에 확장하여, Backbone과 Neck을 교사-학생 구조로 구성한다. Backbone은 학생 역할을 하고, Neck은 교사 역할을 하여 Neck의 풍부한 표현력을 Backbone이 직접 학습을 가능하게 한다.

둘째, 교사와 학생 모델을 동일 네트워크 내에서 동시에 훈련하는 방법을 제안하여 교사 모델의 사전 훈련이 필요 없도록 하며, 훈련 시간을 단축시킨다. 이 접근 방식은 YOLOv6 모델의 훈련 효율성을 극대화하며, VOC 데이터셋에서 YOLOv6-s 모델에 비해 작은 객체에서 2.9% 성능 향상, YOLOv6-m 모델에 비해 작은 객체에서 4.6% 성능 향상을 달성했다.

Ⅱ. 관련 연구

A. 이미지 분류에서의 자가 증류

Zhang et al.[9]은 ResNet 기반의 자가 증류 프레임워크를 제안하였으며, 〈그림 1〉은 해당 프레임워크를 나타낸다. 기본적인 ResNet 구조를 기반으로 다양한 깊이의 계층에서 병목(bottleneck) 모듈을 활용하여 특징을 추출하며, 네트워크는 소프트 맥스 손실 함수로 이미지 분류를 수행하도록 설계되었다. 추가적으로, 이러한 병목에서 추출된 각 특징에 대해 완전 연결(FC) 계층과 소프트 맥스 손실을 적용하여, 하나의 교사(layer 4)와 세 개의 학생(layer 3 이하)으로 구성된 증류 구조를 형성한다. 이 과정에서 사용되는 증류 손실은 세 가지 주요 구성 요소로 이루어진다. 첫째, 교사 계층에서 추출된 교사 특징과 각 병목에서 추출된 학생 특징 간의 L2 손실이다. 둘째, 교사의 소프트맥스 출력이 교사 레이블로, 병목의 소프트맥스 출력이 학생 레이블로 작용하는 KL 발산 손실이다. 마지막으로, 최종 계층 출력과 실제 레이블(GT 레이블) 간의 교차 엔트로피 손실이 포함된다.

Deep supervision[11]은 이미 여러 계층에 분류기를 추가하여 중간 계층이 직접 학습하도록 함으로써 기울기 소실 문제를 해결하는 방식을 제안한 바 있다. 하지만 중간 계층이 지나치게 높은 목표를 설정할 경우, 최종 출력 계층의 성능을 저해할 수 있다. 이러한 문제를 해결하기 위해 Zhang et al.[9]은 병목 계층을 사용하여 얕은 계층의 영향을 줄이면서, 중간 계층이 최종 계층에서 핵심 특징을 학습하도록 하여 Deep supervision의 장점은 활용하면서 단점을 보완하였다.

B. 실시간 객체 탐지에서의 자가 증류

실시간 객체 탐지에서는 YOLOv6 3.0[10]에서 자가 증류를 적용하였다. 기존의 지식 증류는 일반적으로 큰 교사 모델과 작은 학생 모델을 포함하지만, YOLOv6 3.0[10]은 동일한 모델을 교사와 학생으로 사용한다. 먼저 교사 모델을 훈련한 후, 학생 모델은 동일한 데이터셋으로 훈련되며, 이때 교사 모델의 클래스 예측과 회귀 예측을 기반으로 KL 발산을 측정하여 증류 손실을 계산한다. 이 증류 손실은 학생 모델의 탐지 손실과 결합되어 최종 손실을 계산한다. 훈련 초반에는 교사의 소프트 라벨에 더 많은 비중을 두어 학생 모델의 일반화 성능을 향상시킨다. 훈련이 진행되면서 학생 모델의 성능이 교사 모델에 가까워지면, 증류 손실의 가중치를 줄이고 학생의 하드 라벨에 더 집중한다.

그러나 이 방법은 기존의 교사-학생 구조의 한계를 여전히 가지고 있어 훈련 시간이 늘어난다는 단점이 있다. 예를 들어, 300 에포크 훈련의 경우, 교사 모델에 300 에포크, 학생 모델에 또 다른 300 에포크가 필요하므로 훈련 시간이 두 배가 된다. 이에 본 원고에서는 Zhang et al.[9]의 자가 증류 방식을 YOLOv6 3.0[10] 프레임워크에 적용하여 훈련 시간을 크게 늘리지 않으면서도 일반화 성능을 향상시키는 방법을 제안한다.

Ⅲ. 제안하는 기법

최근 객체 탐지 방법에서는 다양한 크기의 객체를 인식 성능을 향상시키기 위해 특징 피라미드 네트워크(Feature Pyramid Network, FPN)[12] 구조가 널리 사용된다. FPN은 다양한 해상도의 특징 맵을 결합하여 작은 객체부터 큰 객체까지 강력한 다중 스케일 성능을 발휘한

다. 또한, Backbone에서 고해상도 저수준 특징이 추출되고, 저해상도 고수준 특징이 생성되어 Neck에서 위에서 아래로 결합되며 다중 스케일에 걸친 풍부한 정보를 제공한다.

이에 본 원고에서는 Neck 출력(교사 특징)과 Backbone 출력(학생 특징) 간의 특징 증류를 통해 모델 훈련을 강화하는 방법을 제안한다. 다중 스케일 정보를 통합한 Neck의 출력이 교사 역할을 하며, Backbone은 학생으로서 교사로부터 학습한다. 이를 통해 학생 Backbone은 Neck 출력에 포함된 풍부한 정보를 효율적으로 학습할 수 있으며, 모델의 일반화 성능을 향상시킬 수 있다.

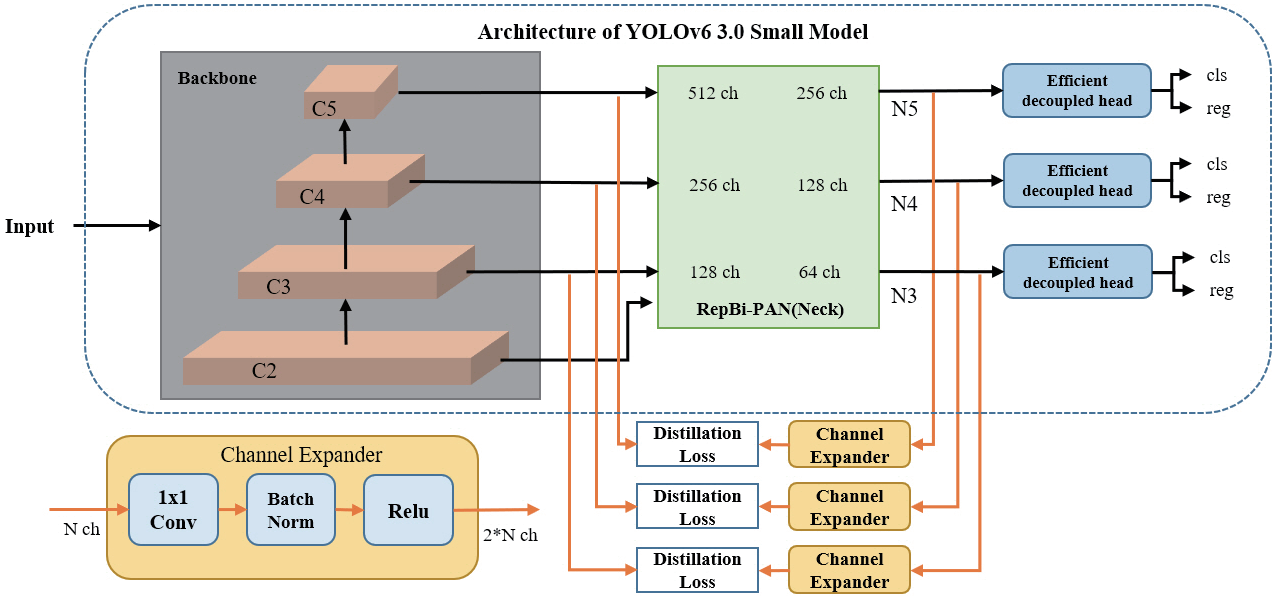
A. YOLOv6의 neck 아키텍처

YOLOv6[13]의 Neck은 경로 통합 네트워크(Path Aggregation Network, PANet)[14]를 기반으로 하며, 상위 및 하위 계층의 특징을 결합한다. PANet은 다중 스케일 특징을 통합하여 작은 객체부터 큰 객체까지 탐지할 수 있도록 하며, 상향식 경로를 추가하여 저수준에서 고수준 특징으로의 정보 경로를 단축시켜 정확한 탐지를 가능하게 한다. 이 구조는 효율성과 빠른 추론 속도를 유지하면서도 탐지 정확도를 향상시킨다.

또한, YOLOv6 3.0[10]의 Neck은 PANet의 개선 버전인 RepBi-PAN을 활용하여 다중 스케일 특징을 통합함으로써 객체 탐지 성능을 향상시키도록 설계되었다. BiC(Bi-directional Concatenation) 모듈은 세 개의 인접 계층에서 특징 맵을 연결하여 Backbone의 저수준 특징과 고수준 특징을 융합한다. 이는 작은 객체를 정확히 탐지하는 데 중요한 위치 정보를 보존하는 데 도움을 준다. 추가적으로, SPPF 블록[15]을 간소화한 SimCSPSPPF 블록[10]이 도입되어 특징 표현력을 강화한다. 구조는 또한 작은 모델에는 RepBlock[16], 큰 모델에는 CSPStackRepBlock[13]을 적용하여 너비와 깊이를 조정함으로써 더욱 최적화되었다. 이와 같은 설계를 통해 YOLOv6 3.0[10]의 Neck은 추론 속도와 탐지 정확도 간의 균형을 유지한다. Neck의 이러한 강력한 특성은 모델의 전반적인 성능에 중요한 역할을 하며, Backbone의 출력과 비교했을 때 Neck의 출력은 증류 과정에서 교사 역할을 수행할 수 있는 탁월한 능력을 보여준다.

B. 네트워크 디자인

제안한 네트워크 구조는 〈그림 2〉와 같다. 기존



〈그림 2〉 이 그림은 YOLOv6 3.0[10]을 기반으로 제안된 셀프 증류 프레임워크를 보여준다. Neck의 출력(N3, N4, N5)은 교사 특징으로 사용되며, Backbone의 출력(C3, C4, C5)은 학생 특징으로 사용된다. 채널 일치를 위해 채널 확장기가 사용되며, 이후 교사와 학생 간의 증류 손실이 계산된다.

YOLOv6 3.0[10] 아키텍처에서 Backbone의 다중 스케일 특징 출력 {C3, C4, C5}와 Neck의 출력 {N3, N4, N5}를 추출한다. C3는 고해상도의 저수준 특징 맵으로, 작은 객체를 탐지하는 데 효과적이며, C4는 중간 크기 객체, C5는 큰 객체에 적합하다. 다중 스케일 특징 표현을 강화하기 위해 N3, N4, N5를 교사로, C3, C4, C5를 학생으로 설정하여 각각 N3–C3, N4–C4, N5–C5 쌍을 형성한다.

그러나 YOLOv6–s에서 C3, C4, C5는 각각 128, 256, 512 채널을 가지는 반면, N3, N4, N5는 각각 64, 128, 256 채널을 가진다. 따라서 증류 손실을 계산하려면 채널 매칭이 필요하다. 이를 해결하기 위해 교사의 채널 수를 두 배로 늘리거나 학생의 채널 수를 절반으로 줄일 수 있다. 본 원고에서는 교사의 채널 수를 증가시키는 방법을 선택하였다. 학생의 채널 수를 줄이면 특징 표현이 감소하여 복잡한 객체 탐지에서 중요한 세부 사항을 놓칠 가능성이 있어 학생이 교사로부터 충분히 학습하지 못할 위험이 있다. 반면, 교사의 채널 수를 증가시키면 풍부한 표현력을 유지할 수 있어 효과적인 증류가 가능하며, 학생에게 더 세부적인 특징 정보를 제공할 수 있다

이렇듯 채널 불일치를 해결하기 위해 N3, N4, N5 출력은 채널 확장기를 통해 채널 수를 두 배로 늘려 Backbone 출력과 매칭된다. 채널 확장기는 1x1 컨볼루션을 적용하여 채널 수를 두 배로 늘린 후, 배치 정규화를 사용해 교사 특징의 평균과 분산을 정규화하여 훈련 안정성과 수렴 속도를 향상시킨다. 또한 ReLU 활성화 함수를 추가하여 비선형성을 도입한다. 추론 시에는 채널 확장기가 제거되고 기존 모델의 구조를 그대로 사용하므로 계산 속도와 성능은 동일하게 유지된다.

C. 손실 함수

가장 기본적인 특징 증류 손실 함수는 힌트 손실(Hint Loss)로, 교사와 학생의 특징 간의 제곱 차이를 최소화하며, 이는 평균 제곱 오차(Mean Squared Error, MSE) 손실과 유사하다. 만약 교사의 특징이 이미 잘 학습된 상태라면, 힌트 손실만으로도 좋은 결과를 얻을 수 있다. 그러나 제안된 자기 증류 접근법에서는 동일한 모델 내에서 교사와 학생이 처음부터 학습을 시작하기 때문에, 훈련 초기 단계에서는 두 모델의 특징이 불안정하고 성능이 낮다. 이런 경우에는 KL 발산(KL Divergence, KL Div)이 더 적합하다. KL Div는 교사와 학생의 확률 분포 간의 차이를 최소화하여, 교사의 특징이 완전히 학습되지 않은 상태에서도 학생이 교사로부터 세밀한 확률 정보를 학습할 수 있도록 한다. 이를 통해 초기 훈련의 불안정성을 완화하고, 교사와 학생 모델이 함께 훈련됨에 따라 성능 향상을 가능하게 한다.

제안된 방법에서는 Backbone의 출력은 학생 특징 $P_S = \{C3, C4, C5\}$로 정의하고 Neck의 출력은 교사 특징인 $P_T = \{N3, N4, N5\}$로 정의하고, 각 특징 쌍 간의 KL Div 손실을 계산한다. KL Div를 위해 학생 모델과 교사 모델의 특징 맵에 로그 소프트맥스(log–softmax) 함수를 적용한다. 로그 소프트맥스를 사용하는 이유는 원시 특징 맵을 확률 분포로 변환하기 위함이다. 이는 KL Div 손실 계산에 필수적이며, 모델 출력이 확률의 맥락에서 수치적으로 안정적이고 해석 가능하게 만들어준다. 또한, 로그 소프트맥스를 사용하면 연산이 로그 공간에서 이루어져 훈련 중 수치적 안정성을 높일 수 있다.

추가적으로, 로그 소프트맥스 함수 내에서 온도 스케일링(temperature scaling)을 적용하여 출력 확률 분포의 매끄러움을 조절한다. 이는 출력 분포의 선명도를 조정하여 모델의 예측이 과도하게 확신하는 것을 방지한다. 로그 소프트맥스 변환 이후, 교사와 학생 모델 간 KL Div 손실을 계산하며, 배치 크기와 채널 차원을 기준으로 정규화하여 그래디언트 업데이트의 균형을 맞춘다.

교사 모델과 학생 모델의 확률 분포는 다음과 같이 정의된다.

$$P_F(x) = \log\left(\frac{\exp(F(x)/t)}{\sum_i \exp(F(x_i)/t)}\right) \quad (1)$$

여기서 $F(x)$는 특징 맵 $F(x) \in \mathbb{R}^{C\times H\times W}$를 나타내며, t 는 온도 스케일링 값이다. 따라서, 교사 모델의 특징에 대한 확률 분포는 $P_T = \{P_N^3, P_N^4, P_N^5\}$로, 학생 모델의 특징에 대한 확률 분포는 $P_S = \{P_C^3, P_C^4, P_C^5\}$로 정의된다. 이러한 확률 분포 함수를 사용하여 i 번째 다중 스케일 특징에 대한 KL Div 손실은 다음과 같이 정의된다.

$$L_{KL}(P_T^i \| P_S^i) = \sum P_T^i(x)\log(P_T^i(x)/P_S^i(x)) \quad (2)$$

여기서 $P_T^i(x)$는 i 번째 교사 특징의 확률 분포를, $P_S^i(x)$는 i 번째 학생 특징의 확률 분포를 나타낸다. 따라서, 각 교사–학생 특징 쌍에 대한 KL Div 손실은 다음과 같이 정의된다.

$$L_{distillation} = \sum_{i=3}^{5} L_{KL}(P_T^i \| P_S^i) \quad (3)$$

최종 손실 함수는 YOLOv6 3.0의 탐지 손실과 증류 손실을 결합하여 다음과 같이 정의된다.

$$L_{total} = \mathrm{a} \cdot L_{detection} + (1 - \mathrm{a}) \cdot \mathrm{b} \cdot L_{distillation} \quad (4)$$

여기서 $L_{detection}$은 YOLOv6의 탐지 손실을 나타내며, α는 탐지 손실과 증류 손실 간의 균형을 조정하는 하이퍼파라미터이다. 본 연구에서는 교사와 학생이 동시에 학습되는 점을 고려하여 안정적인 학습을 보장하기 위해 α 값을 0.9로 고정하였다.

β 값은 YOLOv6 3.0에서 제안된 하이퍼 파라미터로, 자기 증류 손실 계산 시 학습 과정에서 동적으로 조정된다. YOLOv6 3.0에서는 학습 초기에는 교사 모델의 소프트 라벨에 더 많은 가중치를 부여하고, 학습이 진행됨에 따라 학생 모델의 하드 라벨이 더 중요한 비중을 차지하도록 설계되었다. 이러한 전략은 학습 초기에는 교사 모델의 소프트 라벨의 효과를 최대한 활용하고, 학습 후반부에는 학생 모델의 성능에 집중할 수 있도록 돕는다. YOLOv6 3.0에서 사용된 가중치 감소(weight decay) 공식은 다음과 같다:

$$\mathrm{b} = -0.99 \times \left(\frac{1 - \cos(\mathrm{p} \times E_i E_{max})}{2}\right) + 1 \quad (5)$$

여기서 E_i는 현재 에포크를, $E_{\max}$는 총 에포크 수를 나타낸다.

본 연구에서도 β를 증류 손실에 적용하였다. 학습 초기 단계에서는 Backbone이 아직 완전히 학습되지 않은 상태이기 때문에 Neck에서 추출된 고수준 특징이 Backbone으로 효과적으로 전달되는 것이 중요하다. 따라서 학습 초기에는 증류 손실에 높은 가중치를 부여하여 Backbone이 Neck으로부터 활발히 학습할 수 있도록 하였다. 그러나 학습이 진행되고 Backbone이 안정화됨에 따라, 증류 손실이 탐지 손실과 충돌하여 학습 과정에 방해가 될 가능성이 있다. 이를 방지하기 위해 학습 후반부에서는 증류 손실의 비중을 줄이고 탐지 손실에 더 집중하도록 하여 학습 효율성을 높였다.

Ⅳ. 실험 결과

A. 실험 환경

제안한 자기 증류 방법을 VOC0712와 VisDrone-DET2019 두 개의 데이터셋을 사용하여 평가하였다. VOC 데이터셋은 VOC2007과 VOC2012에서 추출된 16,551개의 학습 및 검증 이미지와 VOC2007의 4,952개의 테스트 이미지로 구성되며, 비행기, 자전거, 새, 보트 등 20개의 클래스를 포함한다. VisDrone-DET2019 데이터셋은 드론으로 촬영된 8,629개의 이미지로 이루어져 있으며, 이 중 6,471개는 학습용, 548개는 검증용, 1,610개는 테스트용으로 사용된다. 이 데이터셋은 보행자, 자동차, 밴, 버스 등 10개의 카테고리로 분류된다.

VOC 데이터셋의 경우, 모든 모델이 300 에포크 이내에 최적 성능에 도달했으며, 이에 따라 추가 실험은 진행되지 않았다. 그러나 VisDrone-DET2019 데이터셋에서는 300 에포크 이후에도 성능 개선이 지속되어 실험은 400 에포크까지 확장되었다.

B. 구현 세부사항

본 연구에서 사용된 방법은 GitHub에 공개된 YOLOv6 3.0 소스 코드를 기반으로 구현되었다. 실험은 소스 코드에서 제공하는 기본 하이퍼파라미터를 사용하여 진행되었다. mAP 성능 비교를 위해 YOLOv6s와 YOLOv6m 모델이 사용되었다. 학습과 평가는 Nvidia RTX A6000 GPU에서 수행되었으며, 입력 이미지는 640×640 픽셀로 리 사이즈 되었다.

학습 과정에서는 확률적 경사 하강법(Stochastic Gradient Descent, SGD) 옵티마이저와 람다 학습률 스케줄러가 사용되었다. 모든 모델은 초기부터 사전 학습된

가중치를 사용하지 않고 학습되었으며, 이는 학습 초반부터 제안된 방법의 효과를 객관적으로 평가하고 예상치 못한 성능 향상을 방지하기 위함이다.

C. YOLOv6 3.0과의 비교

VOC 데이터셋에서는, YOLOv6s 모델에 비해 mAP가 0.1% 증가했으며, YOLOv6m 모델에 비해서는 1.8% 증가하였다. 작은 모델에서의 전반적인 성능 향상은 크지 않았지만, 작은 객체에 대해서는 성능이 2.9% 증가하였다. 중간 모델에서는 작은 객체 성능이 4.6% 향상되었다.

VisDrone 데이터셋에서는, YOLOv6s 모델에 비해 mAP가 0.3% 증가했으며, YOLOv6m 모델에 비해서는 0.1% 증가하였다. 작은 모델에서는 작은 객체에 대한 성능이 0.2%, 중간 객체에 대한 성능이 0.4% 증가하였다. 큰 객체의 경우, YOLOv6m 모델에 비해 제안된 방법이 0.1%의 성능 향상을 보였으나, 다른 경우에서는 성능이 베이스라인 모델에 비해 감소하였다(〈표 1〉 참조).

YOLOv6에서 제안된 자기 지식 증류 방식의 경우, YOLOv6m 모델은 VisDrone 데이터셋에서 본 원고에서 제안한 방법과 동일한 mAP를 달성하였다. 그러나 YOLOv6s 모델의 경우, 성능이 베이스라인 모델에 비해 상당히 감소하였다.

D. 증류 손실 함수에 대한 분석

VOC 데이터셋에서 특징 증류를 수행하기 위해 Hint Loss와 KL Divergence Loss를 비교하였다(〈표 2〉 참조). YOLOv6s 기본 모델을 300 에포크 동안 학습한 후 mAP 성능은 63.9%였다. 그러나 Hint Loss만을 사용하여 증류 방법을 적용한 경우, mAP가 63.7%로 감소하며 기본 모델에 비해 성능 저하가 발생하였다.

E. 가중치 감소에 대한 분석

학습이 진행됨에 따라 증류 손실 가중치를 점진적으로 감소시키는 가중치 감소(weight decay)의 효과를 검토하기 위해 분석 연구를 수행하였다. 〈표 3〉에 나타난 바와 같이, 제안된 방법에 가중치 감소를 적용했을 때 적용하지 않은 경우에 비해 mAP 성능이 0.7% 높아지는 결과를 얻었다.

〈표 1〉 YOLOV6 3.0과의 성능 비교

Method	epoch	mAP_{50}	mAP_{75}	mAP	mAP_s	mAP_m	mAP_l
VOC Dataset							
YOLOv6s	300	84.3	70.1	63.9	23.8	45.9	70.2
YOLOv6s–Self distillation	300+300	84.7	70.0	63.6	27.2	46.1	69.8
YOLOv6s–Self distillation (Proposed)	300	84.2	70.2	64.0	26.7	46.0	70.2
YOLOv6m	300	87.6	75.2	68.8	30.6	53.7	74.6
YOLOv6m–Self distillation	300+300	86.8	74.3	67.2	27.2	51.7	73.1
YOLOv6m–Self distillation (Proposed)	300	87.7	75.3	69.0	35.2	53.6	74.7
VisDrone Dataset							
YOLOv6s	400	29.8	16.6	16.6	6.9	26.6	40.6
YOLOv6s–Self distillation	300+300	16.2	8.3	8.8	2.2	14.0	29.8
YOLOv6s–Self distillation (Proposed)	400	30.5	16.8	16.9	7.1	27.0	36.6
YOLOv6m	400	32.4	18.3	18.3	8.0	28.6	42.0
YOLOv6m–Self distillation	300+300	33.2	18.3	18.4	8.1	28.9	43.8
YOLOv6m–Self distillation (Proposed)	400	32.5	18.5	18.4	8.0	29.2	41.2

〈표 2〉 Hint Loss vs. KL Divergence

Distillation Method	mAP
Base model(YOLOv6s)	63.9
Hint loss	63.7
KL divergence loss (Proposed)	64.0

〈표 3〉 가중치 감소에 대한 분석

Weight Decay	mAP
Base model(YOLOv6s)	63.9
Proposed(w Weight decay)	64.0
Proposed(w/o Weight decay)	63.3

Ⅴ. 결론

본 원고에서는 YOLOv6 모델에 적용된 자가 증류(self-distillation) 기법을 제안하여 객체 탐지 성능을 향상시켰다. 특히, Neck의 출력을 교사 특징으로, Backbone의 출력을 학생 특징으로 설정함으로써 다중 스케일 특징 증류를 통해 모델의 일반화 능력을 개선했다. 실험 결과, YOLOv6-s와 비교하여 작은 객체 탐지 성능이 2.9% 향상되는 등, 작은 객체와 중간 크기 객체에서 유의미한 성능 향상을 확인했다. 또한, 학습이 진행됨에 따라 증류 손실의 영향을 점진적으로 줄이는 가중치 감소(weight decay)를 적용함으로써 학습 후반 단계에서 모델 성능을 추가로 개선할 수 있었다.

Ⅵ. 한계점 및 향후 연구

본 원고에서는 VOC와 VisDrone 데이터셋에만 검증되었으며, 다른 데이터셋에 대한 일반화 가능성은 검토되지 않았다. 따라서, 향후 연구에서는 더 다양하고 복잡한 데이터셋에서 제안된 기법을 평가하여 일반화 가능성을 확인할 필요가 있다. 또한, 제안된 셀프 증류 기법을 FCOS[17], RetinaNet[18] 등 FPN을 사용하는 모델에 적용하는 것도 유의미한 연구 방향이 될 것이다. 이를 통해 다양한 객체 탐지 아키텍처와 데이터셋에서 셀프 증류의 효과와 효율성을 보다 포괄적으로 평가할 수 있을 것이다.

참 고 문 헌

[1] G. Hinton, O. Vinyals, and J. Dean, “Distilling the knowledge in a neural network,” arXiv preprint arXiv:1503.02531, 2015.

[2] J. Yim, D. Joo, J. Bae, and J. Kim, “A gift from knowledge distillation: Fast optimization, network minimization and transfer learning,” in CVPR, 2017.

[3] V. Sanh, L. Debut, J. Chaumond, and T. Wolf, “DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter,” arXiv preprint arXiv:1910.01108, 2019.

[4] G. Chen, W. Choi, X. Yu, T. Han, and M. Chandraker, “Learning efficient object detection models with knowledge distillation,” in NeurIPS, 2017.

[5] A. Romero, N. Ballas, S. E. Kahou, A. Chassang, C. Gatta, and Y. Bengio, “FitNets: Hints for thin deep nets,” in International Conference on Learning Representations (ICLR), 2015.

[6] S. Zagoruyko and N. Komodakis, “Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer,” in International Conference on Learning Representations (ICLR), 2017.

[7] T. Wang, L. Yuan, X. Zhang, and J. Feng, “Distilling object detectors with fine-grained feature imitation,” in CVPR, 2019.

[8] T. Furlanello, Z. Lipton, M. Tschannen, L. Itti, and A. Anandkumar, “Born again neural networks,” in ICML, 2018.

[9] L. Zhang, J. Song, A. Gao, J. Chen, C. Bao, and K. Ma, “Be your own teacher: Improve the performance of convolutional neural networks via self distillation,” in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3713-3722, 2019.

[10] C. Li, L. Li, Y. Geng, H. Jiang, M. Cheng, B. Zhang, Z. Ke, X. Xu, and X. Chu, “Yolov6 v3.0: A full-scale reloading,” arXiv preprint arXiv:2301.05586, 2023.

[11] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” in CVPR, 2015.

[12] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S.

Belongie, "Feature pyramid networks for object detection," in CVPR, 2017.

[13] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, L. Li, Z. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei, and X. Wei, "Yolov6: A single-stage object detection framework for industrial applications," CoRR, abs/2209.02976, 2022.

[14] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, "Path aggregation network for instance segmentation," in CVPR, 2018.

[15] G. Jocher, "YOLOv5 release v6.1," 2022. Available: https://github.com/ultralytics/yolov5/releases/tag/v6.1.

[16] X. Ding, X. Zhang, N. Ma, J. Han, G. Ding, and J. Sun, "Repvgg: Making vgg-style convnets great again," in CVPR, 2021.

[17] Z. Tian, C. Shen, H. Chen, and T. He, "FCOS: Fully convolutional one-stage object detection," in ICCV, 2019.

[18] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in Proceedings of the IEEE International Conference on Computer Vision, pp. 2980-2988, 2017.

이동석

• 2008년 2월 인하대학교 컴퓨터정보공학 학사
• 2007년 12월 ~ 현재 한화시스템 SW팀(지능형)

〈관심 분야〉
Electro-Optics, Image Processing, Deep Learning, Real-time Object Detection, Knowledge Distillation, Continual learning, Domain adaptation

김해문

• 2022년 8월 한양대학교 인공지능융합학과 석사
• 2022년 7월 ~ 현재 한화시스템 SW팀(지능형)

〈관심 분야〉
Object Detection, Knowledge Distillation, Image Processing

AI의 메타인지 능력 개발


황중원
한국전자통신연구원

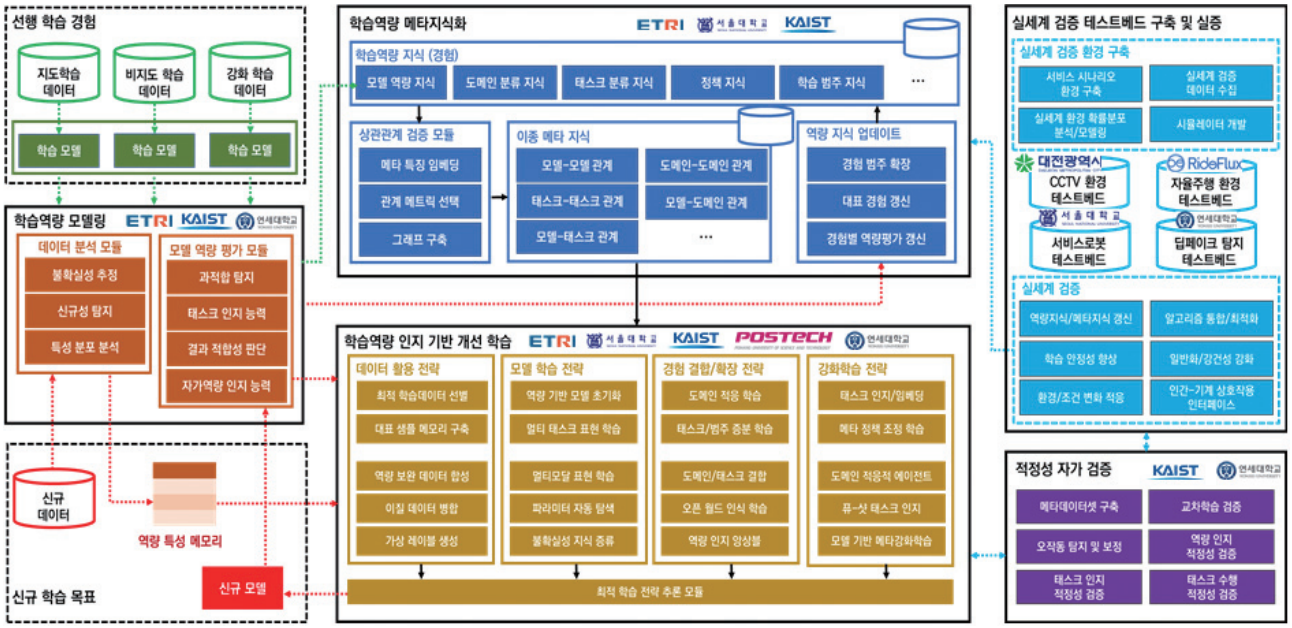

윤기민
한국전자통신연구원


배유석
한국전자통신연구원

Ⅰ. 서 론

AI 기술이 발전함에 따라서 우리의 생활은 AI와 점점 밀접해지고 있다. 거리에서는 심심치 않게 자율주행 자동차가 보이고, 생성형 AI는 창작 활동의 패러다임을 흔들고 있으며, 식당에서는 로봇을 통해 서빙을 받는 경우도 늘어나고 있다. 상상을 초월한 AI의 발전속도에 힘입어, 우리는 AI에게 중요한 결정을 맡기기 시작했지만, 때로는 AI가 인간의 기대를 벗어난 결과를 도출하여 신뢰를 떨어뜨리기도 한다. 일례로 2021년 온라인 부동산 시장인 Zillow는 부동산 평가를 하는 AI의 예측치 못한 실패로 약 8억 8천만 달러의 손실을 입고 전체 직원의 25%를 감축해야하는 어려움을 겪었다. 또한 생성형 AI가 실제로는 존재하지 않는 사실을 사실인 것처럼 만들어내는 환각현상(Hallucination)이 존재한다는 사실은 널리 알려져 있다. 이와 관련하여 미국 뉴욕에서 한 변호사가 AI가 환각현상으로 생성한 가짜 판례를 법적 서류로 제출하였다가 곤혹을 치르기도 하였다.

이와 같은 문제는 AI의 메타인지 능력(Meta cognition)[1] 부족에서 발생한다. 인간의 지능 역시 문제풀이에 있어서 실패의 순간이 등장 한다. 그러나 인간의 경우에는 대부분 자신의 사고과정을 성찰하여 실패의 가능성을 예측할 수 있는 능력이 있다. 이렇게 자신의 사고를 성찰할 수 있는 능력을 메타인지 능력이라고 부르며, 메타인지 능력은 실패가 일어날 수 있는 순간에 대응할 수 있는 중요한 전제조건이 된다. 예를 들어, 어두운 밤길을 운전할 때, 앞이 잘 보이지 않는다는 점을 깨닫고 천천히 속도를 줄이는 것은 자신의 판단과 상황을 성찰하여 행동을 조정하는 메타인지 능력을 활용한 예라 할 수 있다. 마찬가지로, 혼자서 해결하기 어려운 문제가 생겼을 때, 다른 사람의 조언을 구하거나 문제를 해결하기 위한 정보를 좀 더 구하는 일 역시 메타인지 능력이



〈그림 1〉 스스로 학습역량을 인지하고 활용하여 적정한 결과를 제공하는 인공지능 기술 개발 과제 구조도

작용한 결과라고 볼 수 있다.

반면 AI는 자신의 판단을 평가하거나 수정할 수 없는 구조적 한계를 지닌다. 현재 AI 학계에서 중심이 되는 방법인 다층신경망 방식은 우리가 데이터라고 부르는 입력-결과쌍의 관계를 학습하고 해당 관계를 관계성으로 일반화(Generalization)하는 엔드-투-엔드(End-to-End) 방식으로 지능이 구축된다. 이러한 학습방식은 기계가 어떠한 사고의 과정을 걸쳐서 결과를 도출하였는지에 대해서, 제한하거나 평가하지 않는다. 즉 AI가 입력에 따라 올바른 결과를 내더라도, 그 결과를 내는 과정이 합리적인지 여부에 대해서는 알 수 없다. 이러한 이유로 일부 학자들은 AI를 확률적으로 때려 맞추는 원숭이(Stochastic Monkey)라고 비판하기도 한다. 이는 세상에 대한 참된 이해 없이 확률적인 패턴만을 맞추고 있는, 저급한 지능을 가진 원숭이라는 뜻이다. 이렇게 AI의 사고과정이 알 수 없는 일종의 블랙박스로 존재하기 때문에, 그 사고과정을 평가하는 메타인지 능력을 구현하는 것은 쉽지 않다. 결과적으로 AI는 자신의 실패 가능성을 예측하지 못하고, 실패에 대응할 수 있는 기회를 놓치게 된다.

메타인지 능력은 인간의 학습 과정에서도 중요한 역할을 한다[2]. 교육학 연구에 따르면, 메타인지 능력이 뛰어난 학습자는 자신의 학습 목표와 전략을 효과적으로 조정하며, 학습 과정에서의 오류를 스스로 인식하고 수정할 수 있다. 이러한 이유로 메타인지 능력은 학업 성취도와 밀접한 상관관계를 보이며, 자기 주도 학습의 핵심 요소로 간주된다. AI에서도 이러한 메타인지적 접근을 구현할 수 있다면, 학습의 효율성을 향상시킬 가능성이 있다.

최근 AI 분야에서는, 비록 메타인지라는 단어를 직접적으로 사용하지 않더라도, 메타인지 능력을 구현하기 위한 다양한 접근법들이 발표되고 있다. 염두에 두어야 할 것은 메타인지 능력은, AI를 통해 구현하고자 하는 인지능력이 무엇이냐에 따라서 요구사항과, 구현 방식, 활용 형태가 달라진다는 것이다. 현재까지 AI 연구의 대표적인 트렌드를 꼽자면 이미지와 영상을 처리하는 컴퓨터 비전 기술, 환경에 맞는 행동 전략을 찾는 강화 학습 기술, 그리고 언어 데이터를 처리하고 생성하는 언어 모델 기술이 있다. 이러한 AI 기술들은 각각 다른 목적을 충족하기 위한 과정을 갖고 있기 때문에, 이를 고찰하는 메타인지 또한 다른 요구사항과 방법론을 통하여 구현되어야 한다.

Ⅱ. 메타인지

메타인지란 단어는 1976년 미국의 발달 심리학자 John H. Flavell이 처음으로 사용하였으며[1], 현재는 교육학과 심리학에서 사고에 대한 사고 능력을 광범위하게 지칭하는 용어로 사용되고 있다. Flavell의 주장에 따르면 인지는 사람, 태스크, 전략에 대한 메타인지적 지식을 바탕으로 목표와 전략을 수립하며, 메타인지적 경험이 메타인지적 지식과 목표, 전략을 수정하거나 강화한다고 한다[3].

이후 교육학에서의 메타인지 연구는 인지에 대한 이해를 넘어, 이를 어떻게 관리하고 조정할 수 있는지에 중점을 두며 발전해왔다. 특히 메타인지를 메타인지적 지식(Metacognitive Knowledge)과 메타인지적 조절(Metacognitive Regulation)이라는 두 가지 주요 측면으로 구분하면서, 이 두 가지가 학습자의 학습 및 자기조절 능력과 밀접한 관련이 있음을 밝히고자 했다[4][5]. 인지에 대한 지식은 학습자가 자신의 사고 과정을 이해하는 능력을 의미하며, 조절은 학습 활동을 구성하고 평가하는 데 필요한 구체적인 행동을 포함한다. 이러한 조절 과정은 학습자가 자신의 목표를 달성하기 위해 학습 전략을 설정하고, 그 과정을 지속적으로 점검하며, 필요에 따라 수정할 수 있도록 돕는다.

메타인지적 조절은 학습자가 자신의 학습 과정을 능동적으로 관리할 수 있도록 돕는 계획(Planning), 모니터링(Monitoring), 평가(Evaluation)의 세 가지 주요 활동으로 이루어진다[4][6]. 계획은 학습 목표를 설정하고, 과제 수행에 필요한 전략과 자원을 배분하는 과정으로, 학습의 방향성을 제시한다. 모니터링은 학습자가 자신의 진행 상황을 실시간으로 점검하며, 목표 달성 여부를 확인하고 필요한 경우 조정을 통해 효율성을 높이는 활동이다. 평가는 학습 후 결과를 검토하고, 사용된 전략의 효과성을 분석하며, 이후의 학습 과정에서 개선점을 도출하는 단계이다. 요컨데 주어진 작업에 대한 사전지식을 바탕으로, 학습되는 도중에 실시간으로, 그리고 학습이 완료된 이후에 최종적으로 자신의 전략과 조율 방식을 검토하는 활동이 메타인지적 조절이다. 이러한 조절 활동들은 독립적으로 작동하기보다 서로 유기적으로 연계되어 학습자의 자기조절 능력을 강화하며, 학습 성과를 향상시키는 핵심적인 역할을 한다.

이러한 메타인지 능력은 학생의 학습 및 문제풀이 능력과 매우 밀접한 관계가 있다는 사실이 여러한 연구 결과로 나타나고 있다. 특히, 주목해야할 점은 메타인지 능력이 때로는 학생이 직접적으로 가진 지능이나 지식보다 중요하게 작용할 때도 있다는 것이다. 다양한 전략과 문제풀이에 필요한 지식을 많이 알고 있고 개개별 전략을 잘 수행하는 능력이 뛰어난 학생보다 제한된 전략을 적재적소에 잘 선택할 수 있는 능력이 있는 학생이 문제 해결 능력이 더 높다는 사실도 종종 관찰된다[7].

교육학 분야에서 메타인지 연구는 인간 학습자의 자기조절 능력을 탐구하는 과정에서 AI 기술 개발에도 중요한 시사점을 제공한다. 특히, 대규모 언어 모델(Large Language Model, LLM)이 활용되기 이전의 AI는 상대적으로 제한된 데이터와 명시적으로 정의된 작업을 수행하는 기계적인 성질이 강했다. 이는 앞서 서술한 문제 풀이에 필요한 지식과 지능이 부족한 상황에 해당하며, 이를 보완할 수 있는 메타인지 능력을 외부적으로 설계하여 AI가 제한된 조건에서도 더 효율적이고 유연하게 작동할 수 있도록 하는 접근 방식이 더욱 요구되었다. 또한 AI 개발의 패러다임이 LLM을 적극적으로 활용하는 방식으로 발전한 이후에도, LLM의 방대한 지식을 학습한 지식을 다양한 작업에 활용하기 위해서 스스로 학습 전략을 설정하고 조정하는 메타인지 능력은 여전히 요구될 것으로 예측된다.

한국전자통신연구원에서는 이러한 메타인지 능력을 갖춘 AI를 구현하기 위하여 '스스로 학습역량을 인지하고 활용하여 적정한 결과를 제공하는 인공지능 기술 개발' 과제(RS-2022-II220124, 이하 학습역량 과제)를 수행하고 있다. 해당 과제는 정보통신기획평가원의 사람 중심 인공지능 핵심 원천 기술 개발 사업의 일환으로, 2022년부터 2026년까지 5년간, AI가 자체적으로 메타인지 능력을 사용하여 효율적인 학습을 수행하고, 추론 결과의 적절성을 사람의 도움 없이 스스로 평가할 수 있도록 여러 가지 방법들을 연구 중에 있다.

해당 과제에서는 메타인지 능력을 활용하여 학습하고 추론하는 프레임워크의 구조를 만들고, 구조 내의 각 부분에 필요한 구성 요소들을 개발하는 방향으로 과제를 수행한다. 이 프레임워크는 크게 메타인지적 지식, 특히 모델과 작업에 대한 지식을 활용하는 학습역량 모델링 모듈, 메타인지적 경험을 누적하여 메타인지적 지식을 강화하는 메타지식화 모듈, 그리고 이를 구체적인 학습 전략에 반영하는 개선 학습 모듈로 구성된다.

학습역량 과제와 관련하여 한국전자통신연구원과 더불어 한국과학기술원, 연세대학교, 서울대학교, 포항공과대학교, 중앙대학교 5개의 학교와 자율주행기업 라이드플럭스가 참여하고 있다. 각 학교는 컴퓨터 비전 혹은 강화학습에 특화하여 연구개발을 진행하고 있으며, 라이드플럭스는 이렇게 개발된 기술을 자율주행에 적용시키기 위한 노력을 진행하고 있다.

III. AI와 메타인지

들어가기에 앞서, AI의 메타인지 개발은 본질적으로 인간의 메타인지 개발과 차이가 존재한다는 점을 짚고 가겠다. 인간 메타인지의 핵심은 스스로 자신의 판단과 사고 과정을 점검하는 능력을 발전시키는 데 있다. 이는 인간 개인 사고의 자율성에 대한 간섭에 한계가 있어 개인의 내적 성장에 의존할 수 밖에 없기 때문이다. 이에 따라 교육학에서는 학생들의 메타인지 능력을 키우기 위하여 학생들에게 메타인지 능력과 활동의 중요성을 가르치고, 메타인지적 요소에 대하여 가르쳐주어 다양한 학습과 태스크 컨텐츠에 직접 적용할 수 있도록 간접적으로 유도한다[4][6].

반면, AI는 설계자가 정의한 알고리즘과 시스템에 의해 구현되며, 이는 AI의 메타인지 역시 마찬가지이다. 따라서 개인의 성장을 통하여 간접적으로 유도되는 인간의 메타인지능력 배양 방식과 다르게, AI에서는 직접적으로 개발자가 어떻게 AI의 사고과정을 판단할지, 그 판단 결과로 어떻게 학습할지를 결정하여 구현한다. 개발자가 메타인지의 원리와 필요성을 염두에 두어, 메타인지적 기능을 AI가 갖도록 개발하는 점에서 차이가 있다. 이러한 이유로 인간의 메타인지를 자신의 사고과정에 대한 사고라고 표현한다면, 기계의 메타인지를 논할 때는 자신이란 수식어가 희석된다. 그에 따라 AI를 통하여 컨텐츠에 대한 사고를 하는 것과 대비하여, 사고과정에 대한 사고를 할 수 있는 시스템과 알고리즘을 추가하는 과정을 AI의 메타인지 개발이라고 말할 수 있다.

1. 컴퓨터 비전에서의 메타인지

컴퓨터 비전 기술은 이미지에 나타나는 시각적 패턴을 분석하고 이해하는 기술로, 사람의 시각적 인식을 컴퓨터에 구현하는 것을 목표로 한다. 과거에는 수작업으로 특징을 추출하고 이에 맞는 알고리즘을 개발하여 인식을 하였으나, 2012년 Alexnet이 영상의 내용물을 판별하는 대회인 Imagenet에서 획기적인 성능 향상을 보여준 이후 딥러닝 기반 접근법이 표준이 되었다. 이후로 다양한 딥러닝 기술들이 컴퓨터 비전 문제를 해결하기 위하여 개발되었다. 이러한 이유로, 이 세션에서 소개하는 메타인지 관련 기술 주제들은 비단 컴퓨터 비전 분야에서만 국한되지 않고, 2012년부터 약 10년동안 기계학습 분야 전반에서 중요시 다뤄진 내용들을 포함한다.

컴퓨터 비전에서 사용되는 딥러닝 기술은 CNN (Convolution Neural Network)의 형태로서, 공간적 계층 구조를 효과적으로 학습하는 일에 특화된다. CNN의 기본적인 구현 방법과 학습 알고리즘은 매우 간단하다. 그러나, 내부적으로는 선택해야만 하는 튜닝요소가 굉장히 많고, 이러한 요소들을 학습하고자 하는 태스크가 달라질 때마다 매번 새롭게 튜닝해야 한다. 이 부분을 설계가 아닌 튜닝이라고 부르는 이유는, 각 요소의 최적의 상태가 논리적인 귀결을 통한 일반론에 의거해서 정해지기보다는 여러 번의 실패를 동반한 실험의 반복으로 찾아지는 경우가 대다수이기 때문이다. 이렇게 튜닝에 많은 노력이 들기 때문에 이를 자동화 하는 AutoML이 제안되었다.

AutoML은 풀고자 하는 태스크에 따라 개별적으로 적용해야 하는 기계학습의 구성요소들을 자동으로 선택하고 최적화하기 위해 알고리즘을 사용하는 방법론을 말한다. AutoML의 과정에서 최적의 학습 전략과 구조에 대

한 계획과, 학습 성과를 실시간으로 모니터링하는 메타인지적 조절 과정이 수반되게 된다. AutoML에 포함되는 대표적인 기술 갈래로는 NAS(Network Architecture Search), HPO(Hyper Parameter Optimization), 메타학습 (Meta Learning)등을 예시로 들 수 있다[8].

NAS는 딥러닝 모델의 실체가 되는 다층신경망의 구조를 학습이 진행됨과 동시에 동적으로 구성하는 방법을 탐구한다. NAS는 크게 ENAS[9]를 포함하는 강화학습 기반의 방식과 DARTS[10]를 포함하는 미분기반의 방식으로 나뉘어서 개발되었다. 강화학습 기반의 방식은 구성하고자 하는 다층신경망의 구조를 더 작은 단위인 블럭의 연속으로 간주하여 다음의 블럭이 무엇이 좋을지를 외부의 controller를 통하여 선택하는 것이다. 반면 미분기반 방식은, 거대한 슈퍼네트워크를 다층신경망의 시작으로 하여, 학습이 진행됨에 따라 필요 없는 부분을 점진적으로 제외해 나가며 최종적인 네트워크를 조각해 내는 방식이다.

HPO는 기계학습 모델의 성능을 극대화하기 위해 모델 학습 과정에서 설정되는 하이퍼파라미터를 체계적으로 탐색하고 최적화하는 방법론이다. HPO를 통하여 최적화되는 하이퍼파라미터들은 학습률, 배치 크기, 모델 깊이, 정규화 강도 등이 있으며, 검색 공간의 효율적인 탐색을 위해 그리드 검색, 랜덤 검색, 베이지안 최적화, 진화 알고리즘 등 다양한 최적화 기법이 활용된다.

메타학습은 모델이 새로운 태스크를 학습하는 능력을 학습하도록 설계된 방법론으로, 일반적으로 "학습하는 법을 학습한다"는 개념으로 설명된다. 메타러닝은 기계가 처리하고자 하는 태스크의 데이터가 적고 빠른 적응이 요구되는 상황에서, 다른 태스크의 데이터를 통해 터득한 학습 방법을 활용하여 빠르게 학습할 수 있도록 한다. 컴퓨터 비전 분야에서 활용된 대표적인 메타학습 기법으로는 MAML(Model Agnostic Meta Learning)[11]과 REPTILE[12]이 있다. 이들은 모델의 초기 파라미터를 어떻게 세팅해야할지를 다른 태스크들로부터 학습하며, 새로운 태스크에 대해 적은 학습 과정을 통해도 높은 성능을 발휘할 수 있도록 설계되었다.

AutoML을 통한 튜닝 이외에, 학습하는 데이터의 중요도나 순서를 메타인지를 통하여 선별하는 방법에 대한 연구들도 진행되었다. 딥러닝은 대규모 데이터를 학습데이터로서 필요로 하는데 데이터 확보 단계에서 비용이 가장 많이 드는 부분은 사람이 데이터에 대해 필요한 정보를 다는 단계, 즉 레이블링을 하는 단계이다. 액티브 학습(Active Learning)은 모델이 학습 과정에서 가장 유용한 데이터의 레이블을 선택적으로 요청하여 학습 효율을 높이는 방법론이다. 어떤 데이터가 중요한지를 판단하는 과정에서 메타인지적 판단 능력이 중요한 역할을 한다.

한편, 커리큘럼 학습(Curriculum Learning)은 학습의 안정성을 높여 최종 성능을 향상시킬 수 있는 접근법이다. 이는 모델이 간단한 태스크나 데이터 샘플부터 학습을 시작하여 더 점진적으로 더 복잡한 문제를 학습하도록 설계된 학습 방법으로, 인간 학습의 점진적인 학습 패턴을 모방한다. 이는 학습 과정에서 난이도를 조정하고 최적의 학습 순서를 정하는 메타인지적 프로세스를 반영한다.

앞서 설명한 기술들이 메타인지 능력을 학습과정에 적용하여 학습 효율을 높이는 방법들에 대한 설명이었다면, 추론 과정에 메타인지적 능력을 활용하여 문제 상황에 대응하는 방법에 대한 연구들 역시 진행되었다. 이에 해당하는 대표적인 연구 분야는 불확실성 추정(Uncertainty Estimation)과 분포 외 탐지(Out-of-Distribution Detection) 분야, AutoEval 분야 등이 있는데, 이 분야들은 원리와 방법론을 상당히 많이 공유한다.

불확실성 추정 기술은 AI 모델의 개개별의 입력에 따른 추론 결과의 신뢰도를 수량화하는 것에 목표를 둔다. 이는 모델이 출력값에 대한 신뢰도를 평가할 수 있도록 하여, 잘못된 결과를 사전에 경고하거나 후속 행동을 유보할 수 있는 기술적 발판이 된다. 불확실성을 추정하기 위해 불확실성 조정(Uncertainty Calibration)[13], MC Dropout[14] 을 포함하는 베이지안 신경망 방식(Bayesian Neural Networks), 에너지 기반 모델(Energy-Based Models)[15] 등의 기법 등이 연구되었다.

분포 외 탐지 기술은 AI가 학습한 데이터 분포 밖에 있는 입력을 마주했을 때 발생하는 문제를 논한다. AI가 오류를 일으킬 때 일어나는 현상에 초점을 두고 연구하는

불확실성 추정과 다르게, 분포 외 탐지 기술은 AI가 오류를 일으키는 원인과 상황에 보다 초점을 두고 기술이 개발된다. 앞서 설명한 대로, 엔드-투-엔드로 학습된 AI는 패턴과 결과간의 관계를 학습하여 맞추는 방식으로 작동하기 때문에, 추론 상황에서 패턴과 결과간 관계가 학습 상황과 달라지는 분포 변화(Distribution shift)가 발생하는 경우 제대로 동작할 수 없게 된다. 이러한 분포 변화를 AI의 성능이 떨어지는 핵심 원인으로 상정한 후, 이러한 현상이 일어나는지 파악하기 위한 기술이 분포 외 탐지 기술이다.

분포 외 탐지 기술은 어떤 방식으로 추론 상황이 학습 데이터셋과 달라지는지에 따라 세부기술로 나뉘어지게 된다. 대표적인 예시로는 이상점 탐지(Anomaly detection), 오픈셋 인식(Openset recognition), 신규성 탐지(Novelty detection) 등이 있다[16]. 이상점 탐지는 비정상을 규정하고 데이터를 구하기 어려운 상황에서 정상 데이터를 학습하여 비정상적인 샘플을 탐지하는 기술이다. 오픈셋 인식은 추론 시 학습하지 않은 범위의 클래스가 입력되었음을 식별하고 이를 기존 클래스와 구별하는 기술이다. 이를 어떠한 수준으로 구분하냐에 따라 오픈월드 인식(Open world recognition)으로 확장되기도 한다. 신규성 탐지는 이상점 탐지와 오픈셋 인식의 중간 정도에 존재하는 기술으로서, 학습 시 보지 못한 데이터인지 아닌지를 판단하는 작업을 뜻한다. 이 때 보지 못한 데이터란 벤치마크에서 보통 보지 못한 클래스의 데이터로 치환되는데, 오픈셋 인식의 경우 학습한 클래스를 올바르게 구분하는 과정도 포함되지만 신규성 발견의 경우 새로운 클래스인지 아닌지 여부만 판단하는 차이가 있다.

특정한 분포 외 데이터를 가정하는 분포 외 탐지 기술과 별개로, 특정 데이터셋에 대한 AI 모델의 성능을 예측하고자 하는 기술에 대한 연구 역시 진행되고 있다. AutoEval[17] 이라고 불리는 갈래의 기술들은 인간이 제공하는 데이터에 대한 정보, 즉 레이블이 없는 상태에서 데이터셋의 메타특징을 사용하여 모델이 자신의 성능을 예측 및 평가하는 방법을 연구한다.

이렇게 자신의 추론 결과를 스스로 평가할 수 있는 기술은 현재까지도 활발하게 연구가 진행되고 있다. 이는 기존 기계학습의 알려진 한계점이 메타인지 능력의 부재와 밀접하게 관계되기 때문이다. 학습역량 과제에서 역시 객체검출에서 위치 불확실성을 측정하는 방법과[18], 학습 과정에서 손실함수를 모니터링하여 과신뢰를 방지하는 방법[19], 무비용-프록시를 결합한 효율적인 NAS 기술[20]들을 개발하여 발표하는 등 다양한 방법에 대한 연구를 진행하고 있다.

2. 강화학습에서의 메타인지 구현 예

강화학습은 AI가 상황에 맞는 최적의 선택을 할 것이 요구될 때, AI를 학습하는 방법이다. 아마도 강화학습이 사용된 예시 중 현시대 한국인들에게 가장 익숙한 예시는 알파고일텐데, 바둑이란 게임에서 얼마나 무궁무진한 국면이 나오며, 이에 맞춰 놓는 수가 미래의 국면에 어떠한 영향을 끼치는지 생각해보면 강화학습이 도전하는 문제가 어떠한 성질을 가졌는지 문제의 본질을 상상할 수 있다. 문제를 단계별로 쪼갰을 때, 각 단계의 상황 상태의 가짓수가 방대하고, 이에 맞는 선택이 미래 결과에 영향을 미치는 순차적 의사 결정 방법(Sequential decision making)을 학습해야할 때, 강화학습이 주로 사용된다.

강화학습은 AI를 탑재한 주체인 AI Agent와 AI Agent가 활동하는 무대인 Environment로 구성된다. AI Agent는 주어진 임무를 수행하기 위해 Environment에서 행동하며, 그 결과로 Reward를 받는다. 이 Reward는 AI Agent가 Environment의 상태를 관찰한 후, 그 상태에 적합한 행동을 선택하여 최종 목표 달성을 위해 조건적으로 행동할 때 극대화된다. 이렇게 Environment의 상태를 AI Agent의 행동으로 매핑하는 AI Agent의 내재적 행동 원리를 Policy라고 부른다.

강화학습의 AI Agent는 Policy에 따라 행동하여 데이터를 얻고 최종적으로 Reward가 어땠는지 모니터링하여 행동을 수정하고 또 다시 데이터를 얻기 위한 행동을 한다. 이 과정에서는 필연적으로 메타인지적 조정인 모니터링 능력을 바탕으로 Policy의 효율성에 대한 지식이 누적되게 된다. 이러한 원리상 강화학습은 어느 정도 기본적으로 메타인지적 특성을 모방한다고 볼 수 있다.

최근의 강화학습 연구는 메타인지 능력을 더 다면적으

로 활용하는 방향으로 연구를 확장하고 있다. 메타강화학습(Meta reinforcement learning)은 이러한 연구 흐름의 대표적인 사례로 AI Agent가 과거에 태스크를 학습한 경험을 반추하여 새로운 태스크에 적응적으로 전략을 세울 수 있는 능력을 구현한다. 즉, 일반적인 강화학습이 하나의 정해진 태스크에 대한 Policy를 학습한다면, 메타강화학습은 여러 태스크의 학습을 통하여 새로운 태스크의 Policy를 구축하는 법을 학습한다. 이러한 확장은 기존 강화학습에서 요구하는 메타인지적 능력에 더하여, 새로운 태스크를 이해할 수 있도록 태스크의 성질을 체계화하는 태스크에 대한 메타인지적 지식과, 그렇게 모델링된 태스크에 어떠한 Policy가 적합한지 전략에 대한 메타인지적 지식을 요구한다.

메타강화학습은 크게 '어떻게 Policy를 모델링할 것이냐', '어떻게 태스크와 환경을 모델링 할 것이냐', '어떻게 효율적으로 데이터를 탐색할 것이냐'에 대한 세 가지 디자인 요소를 갖는다[21]. 이러한 요소들은 서로 어느 정도 연결되지만, 학습역량 과제에서는 태스크와 Environment를 어떻게 모델링할 것이냐의 문제에 초점을 두고 연구가 진행되었다. 이러한 연구의 일환으로 환경을 풀고자 하는 태스크와 상관 없이 범용적으로 모델링하여 다소 비효율적이던 기존의 방식을[22] 태스크에 맞춰 효율적으로 환경을 모델링하고 이에 맞춤형 Policy를 세우는 방법으로 발전시키거나[24], 태스크를 매개변수로 정의하여 단순한 태스크 변화만을 모델링 할 수 있던 기존의 방법을[23] 하위 태스크로 분해하여 정성적인 태스크 변화에도 대처할 수 있도록 하는 방법[25]으로 발전시켰다. 이들은 태스크와 전략의 관계에 대한 메타인지적 지식을 고도화한 결과에 해당한다.

태스크, 전략에 대한 메타인지적 지식의 활용이 요구되는 강화학습 분야는 메타 강화학습 이외에 멀티태스크 강화학습(Multi-task Reinforcement Learning)이 있다. 멀티태스크 강화학습은 여러가지 태스크를 해결할 수 있는 하나의 Policy를 생성하는 것으로, 태스크별로 알맞는 Policy를 생성할 수 있는 Controller를 학습하는 메타강화학습과 차이가 있다.

이러한 멀티태스크 강화학습은 공유된 표현(Shared Representation)을 갖기 때문에, 태스크간 정보를 활용할 수 있어 태스크가 유사하거나 관련된 경우 메타 강화학습에 비해 샘플활용과 연산의 측면에서 효율적이다. 그러나 반대로, 태스크가 이질적일 경우에는 서로 부정적인 간섭을 일으킬 수 있고, 여러가지 이질적인 태스크 중 특정한 태스크 하나에만 집중하여 학습하게 되는 문제가 생길 여지가 있다. 학습역량 과제에서는 이러한 문제를 풀기 위하여 태스크간 영향력과 AI Agent의 태스크별 능력을 고려하여 부정적 간섭과 태스크 편향성을 줄일 수 있는 방법을 개발하였다[26][27].

또한 멀티태스크 강화학습을 메타 강화학습처럼 학습시 관찰하지 못한 태스크에 적용하는 방법도 연구하였다. 어떠한 Reward 함수가 있을 때 어떠한 전략을 취하는 것이 가장 효율적인지에 대한 지식을 여러 태스크를 통하여 확립하여 새로운 태스크가 주어졌을 때 간단한 연산만으로 적응이 가능한 방법 역시 개발하였다[28].

3. 대규모 언어모델 시대의 메타인지

최근 몇 년간 인공지능 AI의 주류 패러다임은 급격한 변화를 겪어서 특정한 문제를 해결하기 위해 설계된 데이터와 전문화된 알고리즘을 개발하는 추세에서 범용적으로 활용 가능한 대규모 언어 모델(Large Language Model, LLM)을 적극적으로 활용하는 방향으로 발전하였다. 이러한 패러다임 전환은 트랜스포머(Transformer)와[29] 같이 순차적 패턴을 익히는데 강력한 성능을 보이는 다층신경망 구조가 획기적으로 발전에 힘입은 결과다. LLM은 대량의 텍스트 데이터를 학습하여 언어 패턴과 구조를 이해하고 예측하는 데 초점을 맞추는 방식으로 학습된다.

언어의 순서를 학습하는 방법론은 크게 세 가지 이유에서 큰 가능성을 제공한다. 첫 번째는 언어의 순서란 추론 과정과 같이 논리적 연속성과 같은 추론적 구조가 내재되어 있다는 것이다. '언어는 생각의 감옥'이라는 표현도 있듯이, 언어는 인간의 사고 과정과 밀접한 관계가 있고 이를 모방하는 것은 인간의 고차원적 지적 활동을 모방하는 첫 단추가 된다. 이 방법론의 강력한 두 번째 이유는 세상에 존재하는 학습 가능한 데이터 중 가장 방대한 비중

이 언어 형태로 존재한다는 점에 있다. 언어 데이터는 양적으로 클 뿐만 아니라, 과학 논문부터 소셜 미디어 게시물, 문학작품, 법률 문서, 기술 매뉴얼에 이르기까지 도메인을 가르지 않고 존재하기 때문에 다양한 논리와 지식을 습득하여 통섭할 수 있다. 세 번째 이유는 이러한 방대한 데이터를 인간의 노력 없이 학습할 수 있다. 단순히 다음으로 올 단어를 추론하는 방식을 사용하기 때문에 특별히 레이블링 작업이 필요하지 않은 것이다. 이러한 방법을 자가지도학습(Self-supervised Learning)이라 부르는데, 언어의 순서를 학습하는 방법은 이 갈래의 학습 방법에 해당하여 큰 데이터를 학습할 수 있다. 이러한 장점들은 데이터양의 증가에 따라 성능이 일정 수준에서 포화되지 않는 트랜스포머의 특성과 상승작용을 일으켜 AI 연구의 패러다임을 크게 바꾸게 되었다.

AI 연구에서 언어모델의 대두는 언어 처리 분야쪽 이외에도 큰 파장을 미치게 되었다. 처리해야 할 입력이 텍스트 형태가 아니라 영상, 음성, 혹은 테이블과 같은 그 데이터 자체여도 언어 모델이 갖고 있는 지식을 활용할 수 있도록 연결하는 멀티 모달 모델에 대한 연구가 활발하게 진행되고 있다. 멀티 모달 모델에 대한 연구를 포함하여 LLM의 대두는 AI의 메타인지와 관련하여 여러 가지 화두를 던지는데, 본 특집원고에서는 크게 세 가지를 언급하고자 한다.

첫 번째는 모달리티간 크로스체킹이 더 고차원적으로 이루어질 수 있게 된다는 점이다. 입력을 두 가지 방식으로, 예를 들면 사진과 그 사진에 대한 언어적 기술로 받아 그 결과를 크로스체킹하여 결과의 불확실성을 측정하는 것은 LLM이 대두되기 이전의 시기에도 가능했다. 이렇게 두 가지 방향으로 독립적인 과정을 통하여 판단한 결과를 취합하여 보다 신뢰성 있는 결과를 얻는 것은 앙상블(Ensemble)을 통한 성능 향상의 기본적인 원리이다. 그러나 언어모델은 이러한 앙상블을 통한 크로스체킹보다 한 단계 더 높은 수준의 크로스체킹을 가능하게 할 것으로 전망된다. 이에 관한 첫 번째 이유는 각 모달리티들이 분석 결과를 의미론적 영역으로 전환하여 보다 통합되고 고차원적인 크로스체킹이 가능해진다는 것이다. 최근 멀티모달 모델 중 일부는 모든 모달리티를 LLM이 이용가능한 시맨틱 임베딩 공간(Semantic Embedding Space)으로 옮겨서 통합한다[29][30]. LLM이 이용가능하다는 뜻은 모달리티의 결과를 단순한 패턴분석을 넘어서 개념적 이해와 매핑이 가능한 특징을 출력하게 된다는 뜻이고, 이를 통하여 의미론적 연관성을 기반으로 정보를 크로스체킹할 수 있게 된다는 뜻이 된다. 앙상블 보다 더 고차원적인 융합이 가능한 두 번째 이유는 입력과 지식간의 크로스체킹이 가능해지기 때문이다. CLIP[31]과 같은 멀티모달 모델은 학습하지 않은 클래스에 대해서도 높은 분류성능을 보여주는데, 이는 언어적 지식을 활용해서 처음 보는 클래스가 어떤 것인지 추측해 낼 수 있기 때문이다. 마찬가지 원리로 학습된 지식이나 언어로 표현된 외부의 지식을 바탕으로 어떠한 모달리티가 낸 결론이 얼마나 합당한지에 대한 크로스체킹이 가능할 것으로 전망된다.

두 번째는 상황과 태스크에 맞춰 모달리티간의 활용 정도를 결정하는 능력이 중요해질 것으로 예측된다. 멀티모달 모델을 설계한다고 해도, 실제 데이터 환경에서는 모든 모달리티의 데이터 수집이 균일하게 이루어질 것이라고 보장하기 어렵다. 가령 소음환경이 심한 경우나 어두울 때, 혹은 하나의 센서가 고장났을 때 모달리티가 제공하는 데이터가 불완전할 수 있다. 또한 태스크별로 근본적으로 요구되는 모달리티가 달라질 수 있다. 가령 와인을 시음하는 태스크를 AI가 수행한다면 이때는 미각 모달리티가 시각 모달리티 보다 중요할 수 있고, 문자를 읽는 태스크를 수행한다면 시각 모달리티가 다른 어떠한 모달리티보다 중요할 것이다. 이러한 점을 생각했을 때, 태스크 및 환경의 특성을 이해하고, 이를 어떻게 활용할지에 대한 전략을 세울 수 있는 메타인지 능력이 더더욱 요구될 것으로 전망된다[32].

세 번째는 설명 가능성(Explainability), 논증(Reasoning)에 대한 연구가 혁신적으로 진행되어 기계가 자신의 메타인지를 사람에게 표현할 수 있는 능력이 점점 향상되고 있다는 사실이다. LLM은 언어구조에 내재된 논리적 구조를 익혀 마치 인간과 소통하듯이 명시적 지시를 학습하고 수행한 후, 그 과정과 결과를 설명할 수 있다는 특성을 갖고 있다. 지금까지 LLM 분야에서는 Prompt[33]나 Instruction[34]형태를 학습하여 명령을

수행하게 하되, 단계적인 생각을 구성할 수 있을 뿐 아니라[35], 자신이 수행한 단계적인 생각을 언어로 다시 나타낼 수 있는 방향의 연구[36]가 진행되고 있다. 이러한 연구들은 인간이 AI에게 직접적으로 구조화된 사고를 유도하고 검증할 수 있는 방법론을 제시한다. 결과적으로 AI는 메타인지를 반영한 결과를 인간과 자연스럽게 소통할 수 있게 되며 경고 없는 실패의 가능성이 크게 줄어들 수 있을 것으로 예상된다.

특히, LLM의 중요한 한계로 꾸준히 지적된 사항 중 하나는, LLM이 일상생활 수준의 단순한 판단은 잘하지만, 복잡한 사고과정이 필요한 논리적인 추론 등의 태스크를 실패하는 경우가 많다는 것이다. 이러한 복잡한 태스크는 사고의 도구들을 적재적소에 효과적으로 채택하여 활용하는 메타인지적 능력이 필수적이다. 따라서 LLM의 중요한 약점을 극복하기 위하여 AI의 메타인지 능력을 구현하는 방법의 다면적인 탐구가 필연적으로 필요할 것으로 전망된다.

Ⅳ. 전망과 결론

AI 기술은 이제 우리의 일상에 더욱 밀접하게 다가오고 있다. AI의 메타인지 능력은 위험을 대처할 수 있도록 불확실성을 관리하며, 사회적, 윤리적 영향을 더 잘 이해할 수 있는 기반을 제시하여 인간과 AI 간의 신뢰성 있는 상호작용을 바탕으로 한 협력적 관계를 강화하는 데 필수적인 역할을 한다. 또한 다양한 태스크와 환경에 대한 일반화 능력을 높이고, 학습의 효율성을 높이기도 한다. 이러한 중요성에 따라 메타인지를 구현하기 위한 연구와 기술 개발이 AI의 모든 분야에서 활발히 진행되고 있다.

2024년 노벨물리학상을 수상한 Geoffrey Hinton교수는 Royal Swedish Academy of Engineering Sciences (IVA)에서 'LLM의 사고과정은 인간의 사고하고 학습하는 방식과 같다.'라는 강연을 하였다. 그는 환각현상을 근거로 AI가 때려맞추는 원숭이일 뿐이며, 인간은 내재된 논리구조로 사고를 하며 AI와 다르다는 언어학자 Noam Chomsky의 주장을 공개적으로 반박한다. Geoffrey Hinton의 주장은 인간 역시 일부 기억된 사실을 바탕으로 논리적 맥락을 통해 자신이 인식하는 사실을 재구성하는 방식의 사고과정을 갖고 있으며, 오히려 Noam Chomsky가 주장하는 지능이 구현된, 엄격하게 정의된 언어학적 논리를 바탕으로 작동하는 심볼릭 모델 기반 AI는 잘 작동하지 않는다고 주장하였다. 만약 Geoffrey Hinton교수의 주장대로 LLM이 인간 지능과 같은 방식으로 동작한다면, 앞으로 LLM에 메타인지 능력을 강화할 때, 기존 교육학, 사회심리학에서 제안된 메타인지 강화능력을 도입하여 더 효율적인 AI를 개발하는 것이 가능할 것으로 전망할 수 있다.

하지만 여기서 한 가지 짚고 넘어가야 할 점은, 언어모델의 메타인지 능력이 인간에 가까운 수준으로 발전한다고 하더라도, 이를 통해 컴퓨터 비전이나 강화학습과 같은 다른 AI 기술 분야에서 제기되는 모든 문제를 자동적으로 해결할 수 있을 것이라는 기대는 과도할 수 있다는 것이다. 논리적 능력이 뛰어나고 사고가 엄밀한 사람이 꼭 시력과 운동신경이 좋은 것은 아니듯이, 고도화된 언어모델이 컴퓨터 비전과 강화학습이 푸는 문제를 모두 해결할 것으로 예측하긴 어렵다. 따라서 컴퓨터 비전과 강화학습도 각각의 메타인지를 바탕으로 더 효율적으로 동작할 방법에 대한 개발이 요구된다. 이러한 연구는 각각의 모달리티의 한계를 상호보완하기 위하여 필요하며, 또한 다양한 방식으로 인공지능이 결합될 수 있는 여지를 열어둔다.

AI 기술에 대한 연구는 놀라울 정도로 빠른 속도로 진행되고 있으며, 세계 각국은 이러한 기술의 선두를 잡기 위해 치열한 경쟁을 벌이고 있다. 이러한 상황에서 특히 메타인지 기술은 AI 시스템의 신뢰성과 효율성을 높이는 핵심 요소로, 앞으로도 연구와 개발의 중요한 한 축이 될 것으로 예측한다. 한국은 이미 AI 연구와 기술 개발에서 세계적인 주목을 받고 있으며, 메타인지 기술을 포함한 차세대 AI의 선도적 역할을 수행할 잠재력을 갖추고 있다고 생각한다. 이를 실현하기 위해 정부, 학계, 산업계가 협력하여 혁신을 이어가 글로벌 AI 시대를 주도하는 중요한 역할을 맡을 수 있기를 기대한다.

Acknowledgements

이 글은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (RS-2022-II220124, 스스로 학습 역량을 인지하고 활용하여 적정한 결과를 제공하는 인공지능 기술 개발)

참고문헌

[1] Flavell, J. H. "Metacognitive aspects of problem solving." The nature of intelligence/Erlbaum (1976).

[2] Pintrich, Paul R. "The role of metacognitive knowledge in learning, teaching, and assessing." Theory into practice 41.4 (2002): 219-225

[3] Flavell, John H. "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry." American psychologist 34.10 (1979): 906.

[4] Schraw, Gregory. "Promoting general metacognitive awareness." Instructional science 26.1 (1998): 113-125.

[5] Jacobs, Janis E., and Scott G. Paris. "Children's metacognition about reading: Issues in definition, measurement, and instruction." Educational psychologist 22.3-4 (1987): 255-278

[6] Hartman, H.H. & Sternberg, R.J. (1993). A broad BACEIS for improving thinking. Instructional Science 21: 401-425.

[7] Swanson, H. Lee. "Influence of metacognitive knowledge and aptitude on problem solving." Journal of educational psychology 82.2 (1990): 306.

[8] Hutter, Frank, Lars Kotthoff, and Joaquin Vanschoren. Automated machine learning: methods, systems, challenges. Springer Nature, 2019.

[9] Pham, Hieu, et al. "Efficient neural architecture search via parameters sharing." International conference on machine learning. PMLR, 2018.

[10] Liu, Hanxiao, Karen Simonyan, and Yiming Yang. "Darts: Differentiable architecture search." arXiv preprint arXiv:1806.09055 (2018).

[11] Finn, Chelsea, Pieter Abbeel, and Sergey Levine. "Model-agnostic meta-learning for fast adaptation of deep networks." International conference on machine learning. PMLR, 2017.

[12] Nichol, Alex, and John Schulman. "Reptile: a scalable metalearning algorithm." arXiv preprint arXiv:1803.02999 2.3 (2018): 4.

[13] Guo, Chuan, et al. "On calibration of modern neural networks." International conference on machine learning. PMLR, 2017.

[14] Gal, Yarin, and Zoubin Ghahramani. "Dropout as a bayesian approximation: Representing model uncertainty in deep learning." international conference on machine learning. PMLR, 2016.

[15] Liu, Weitang, et al. "Energy-based out-of-distribution detection." Advances in neural information processing systems 33 (2020): 21464-21475.

[16] Yang, Jingkang, et al. "Generalized out-of-distribution detection: A survey." International Journal of Computer Vision 132.12 (2024): 5635-5662.

[17] Deng, Weijian, and Liang Zheng. "Are labels always necessary for classifier accuracy evaluation?." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021.

[18] Lee, Youngwan, et al. "Localization uncertainty estimation for anchor-free object detection." European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022.

[19] Park, Hyekang, et al. "Acls: Adaptive and conditional label smoothing for network calibration." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2023.

[20] AZ-NAS: Assembling Zero-Cost Proxies for Network Architecture Search

[21] Beck, Jacob, et al. "A survey of meta-reinforcement learning." arXiv preprint arXiv:2301.08028 (2023).

[22] Rakelly, Kate, et al. "Efficient off-policy meta-reinforcement learning via probabilistic context variables." International conference on machine learning. PMLR, 2019.

[23] Shin, Jaeuk, et al. "On task-relevant loss functions in meta-reinforcement learning." 6th Annual Learning for Dynamics & Control Conference. PMLR, 2024.

[24] Zintgraf, Luisa, et al. "Varibad: A very good method for bayes-adaptive deep rl via meta-learning." arXiv preprint

arXiv:1910.08348 (2019).

[25] Lee, Suyoung, Myungsik Cho, and Youngchul Sung. "Parameterizing non-parametric meta-reinforcement learning tasks via subtask decomposition." Advances in Neural Information Processing Systems 36 (2023): 43356-43383.

[26] Cho, Myungsik, Whiyoung Jung, and Youngchul Sung. "Multi-task reinforcement learning with task representation method." ICLR 2022 Workshop on Generalizable Policy Learning in Physical World. 2022.

[27] Cho, Myungsik, et al. "Hard Tasks First: Multi-Task Reinforcement Learning Through Task Scheduling." Forty-first International Conference on Machine Learning.

[28] Shin, Jaeuk, and Insoon Yang. "Multiparametric Analysis of Multi-Task Markov Decision Processes: Structure, Invariance, and Reducibility." IEEE Control Systems Letters (2024).₩

[29] Alayrac, Jean-Baptiste, et al. "Flamingo: a visual language model for few-shot learning." Advances in neural information processing systems 35 (2022): 23716-23736.

[30] Liu, Haotian, et al. "Visual instruction tuning." Advances in neural information processing systems 36 (2024).

[31] Radford, Alec, et al. "Learning transferable visual models from natural language supervision." International conference on machine learning. PMLR, 2021.

[32] Akbari, Hassan, et al. "Alternating gradient descent and mixture-of-experts for integrated multimodal perception." Advances in Neural Information Processing Systems 36 (2023): 79142-79154.

[33] Brown, Tom, et al. "Language models are few-shot learners." Advances in neural information processing systems 33 (2020): 1877-1901.

[34] Wei, Jason, et al. "Finetuned language models are zero-shot learners." arXiv preprint arXiv:2109.01652 (2021).

[35] Wei, Jason, et al. "Chain-of-thought prompting elicits reasoning in large language models." Advances in neural information processing systems 35 (2022): 24824-24837.

[36] Hu, Yushi, et al. "Visual program distillation: Distilling tools and programmatic reasoning into vision-language models." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

황중원

• 2009년 8월 서울대학교 전기컴퓨터공학부 학사
• 2014년 2월 KAIST 전기공학과 석사
• 2014년 7월 ~ 현재 한국전자통신연구원 선임연구원

〈관심 분야〉
머신러닝, 컴퓨터 비전, 멀티모달 모델, 분포 외 탐지, 신뢰성 있는 인공지능

윤기민

• 2010년 2월 서울대학교 전기공학부 학사
• 2017년 2월 서울대학교 전기컴퓨터공학부 박사
• 2017년 2월 ~ 2019년 2월 한국전자통신연구원 연구원
• 2019년 3월 ~ 현재 한국전자통신연구원 선임연구원

〈관심 분야〉
컴퓨터 비전, 머신러닝, 비디오 이해, 지능형 영상 처리 응용, 멀티모달 인공지능

배유석

- 1995년 2월 경북대학교 컴퓨터과학과 학사
- 1997년 2월 경북대학교 컴퓨터과학과 석사
- 2011년 2월 경북대학교 컴퓨터과학과 박사
- 1996년 12월 ~ 1998년 5월 시스템공학연구소 연구원
- 1998년 6월 ~ 현재 한국전자통신연구원 책임연구원

〈관심 분야〉
딥러닝, 머신러닝, 컴퓨터비전, 시각지능, 인공지능, 멀티모달, 파운데이션 모델

Scientific LLMOps: 과학기술 분야 AI 개발의 중추


이 용
한국과학기술정보연구원


구자현
한국과학기술정보연구원


장래영
한국과학기술정보연구원


이건우
한국과학기술정보연구원


이경하
한국과학기술정보연구원

Ⅰ. 서 론

GPT 또는 LLM(Large Language Models)으로 대표되는 생성형 인공지능이 일상화되어 가고 있다. 인공지능은 단순한 분류와 예측 중심의 모델에서 문장과 그림을 만들어 내는 능력을 가진 모델로 진화하였으며, 사람과 유사한 수준의 지능을 갖춘 AGI(Artificial General Intelligence) 시대가 곧 도래할 것으로 예측되고 있다. 다양한 분야에서 더 많은 인공지능을 원하고, 개인부터 국가적 차원까지 다양한 태스크와 의사결정 수행을 위한 자신만의 인공지능을 갖길 원하고 있다.

과학기술 분야에서도 LLM을 활용하여 수많은 논문과 특허와 같은 기술문서들을 학습하고, 과학연구의 보조자 역할이나, 인공지능 스스로 연구를 수행하기 위한 AI 과학자에 대한 아이디어들이 등장하고 있다. 과학기술 분야에 특화된 LLM 개발은 기술 혁신과 연구 역량 강화를 위한 핵심 전략으로 주목받고 있다. 이러한 특화된 LLM은 방대한 과학 데이터와 고도의 전문 지식을 효과적으로 학습해 복잡한 과학 문제 해결과 혁신적 발견을 가속화하는 데 크게 기여할 것으로 기대된다.

일반적으로, 생성형 모델의 대표 주자인 LLM 모델은 방대한 문서들을 비지도 방식으로 사전 학습하고, Q&A 데이터셋을 통한 지도 학습으로 지식을 급속히 흡수하며, 탁월한 질의응답 능력을 발휘한다. 이로

인해, 과학기술 분야에서 축적된 방대한 지식의 보고들을 빠르게 학습시켜 활용할 수 있는 잠재력을 지니고 있다. 이러한 능력은 단순한 데이터 검색이나 문장 요약을 넘어서, 논리적 추론과 문제 해결 능력을 통해 연구자들을 보조하거나, 일부 분야에서는 연구자의 역할을 대체할 수 있을 정도로 강력하다.

LLM을 과학기술 분야에 특화시키는 것은, 마치 어린 아이를 과학기술 연구자로 성장시키는 과정처럼 매우 복잡하고 정교한 과정을 요구한다. 일반적으로 웹의 방대한 자료와 수많은 서적을 통해 기초적인 지식과 사고, 대화 능력을 갖춘 파운데이션 모델을 기반으로, 전문 분야에 특화된 지식들을 빠르게 습득하여 과학기술 연구 현장에서 바로 적용할 수 있는 체계를 구축해야 한다. 그러나 현재까지 많은 과학기술 연구자들에게 인공지능과 LLM 기술 자체를 충분히 이해하고 활용하는 것조차 쉽지 않은 일이기 때문에, LLM 모델 개발에 직접 참여하는 것은 사실상 불가능에 가까운 일이다. 특히, LLM 모델에 대한 알고리즘적 이해와 더불어, S/W 개발 능력, 학습 데이터 준비, 모델 훈련을 위해서는 막대한 컴퓨팅 자원 확보 등 고도의 전문성과 시간이 요구된다. 더욱이 과학기술 분야에서는 단순히 문장을 생성하는 수준을 넘어서, 과학적 사고와 이론에 대한 깊은 이해를 바탕으로 즉시 관련된 정보를 제공하고, 필요에 따라 과학적 연구과정을 지원할 수 있는 능력이 요구되고 있다.

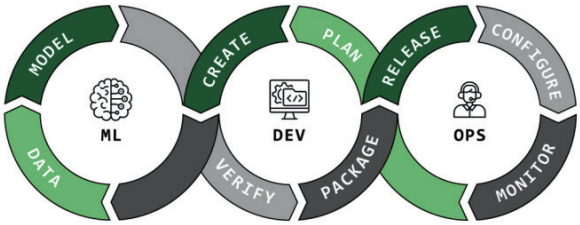
특히, 다양한 과학 분야에서 요구되는 전문화된 LLM을 효과적으로 개발하고 운영하기 위해서는 체계적이고 정교한 방법론이 필수적이다. 이러한 요구를 충족시키기 위해 도입되는 방법론 중 하나가 바로 LLMOps(LLM Operations)이다. LLMOps는 모델 운영을 단순히 관리하는 것을 넘어, 학습 데이터의 수집과 관리, 데이터 전처리, 모델 훈련 및 튜닝, 평가와 검증, 배포, 모니터링, 유지보수에 이르는 전체 라이프사이클을 통합적으로 지원한다. 이를 통해 맞춤형 모델 제작을 보다 효율적으로 수행할 수 있으며, 반복적이고 복잡한 작업들을 자동화하여 개발자의 부담을 경감시킨다. 그러나 과학기술 분야에서는 단순한 모델 운영을 넘어서, 전문적인 과학 지식과 논리적 사고, 문제 해결 능력을 LLM에 효과적으로 반영하기 위해서는 더욱 고도화되고 전문화된 Scientific LLMOps의 구축이 필요하다. Scientific LLMOps는 과학 연구의 특수성을 반영하여 데이터의 품질 관리, 도메인 특화된 전처리 기법, 과학적 검증 절차 등을 포함하며, 이를 통해 과학 연구 현장에서 요구하는 정확성과 신뢰성을 갖춘 LLM을 지속적으로 개발하고 개선할 수 있는 환경을 제공해야 한다.

본 기고문에서는 먼저 기존의 인공지능 개발을 지원하는 다양한 Ops 기술들을 비교 분석한다. 이를 통해 각 기술의 장단점과 과학기술 분야에의 적용 가능성을 평가하고, 과학기술 LLM 개발과 운용에 요구되는 Scientfic LLMOps 특성을 도출한다. 그리고, 한국과학기술정보연구원(KISTI)에서 개발 중인 LLMOps 플랫폼을 소개하며, 이를 통해 과학기술 특화 LLM의 개발과 운영에 어떻게 기여할 수 있는지를 설명한다. 마지막으로, Scientific LLMOps의 향후 과제와 전망에 대해 고찰한다. 이러한 논의를 통해 Scientific LLMOps가 과학기술 분야에서 LLM의 잠재력을 최대한 발휘하고, 연구자들과의 협업을 통해 지식 탐구와 혁신을 가속화하는 핵심 인프라로 자리매김할 수 있기를 기대한다.

Ⅱ. Scientific LLMOps

1. MLOps와 LLMOps

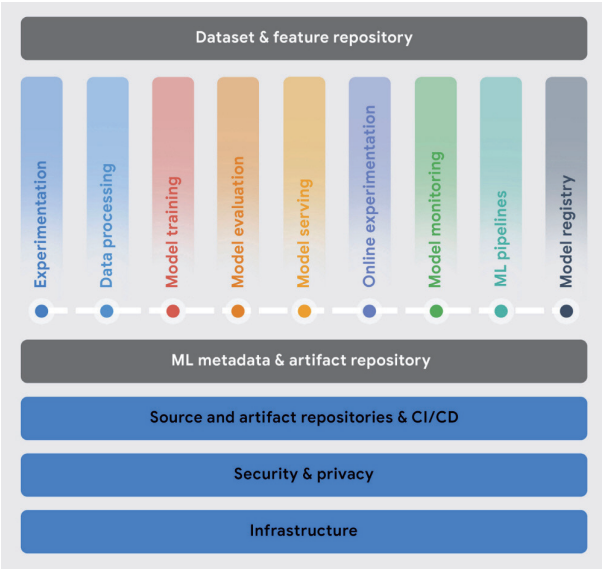
MLOps(Machine Learning Operations)는 머신러닝 모델을 운영하기 위해 필요한 작업을 하나로 통합한 시스템이다[1]. 〈그림 1〉과 같이 소프트웨어 개발과 배포 프로세스를 최적화하는 DevOps의 원칙에 데이터와 모델 관련 요소를 추가한 것으로 데이터 관리부터 모델 모니터링까지 일련의 과정을 효율적으로 수행할 수 있다[2]. 머신러닝 모델을 운영하기 위해 개별적으로 수행되던 단계를 하나의 시스템으로 통합하고 전 과정을 자동화하여 모델의 성능을 향상시키고 머신러닝 시스템의 생산성을 꾸준하게 개선하는 것이 목적이다. MLOps는 DevOps의 주요 원칙이라 할 수 있는 Continuous Integration(CI)과 Continuous Delivery(CD)에 Continuous Training(CT)이라는 개념을 추가한 것이다[3]. CI는 개발자들이 지



〈그림 1〉 MLOps의 워크 플로우 [1]

속적으로 코드를 통합 및 점검하여 코드 품질을 개선하는 과정을 의미하고, CD는 테스트 환경이나 운영 환경에 개선된 코드를 신속하게 배포하여 최종 사용자에게 제공하는 것을 목표로 한다. 여기에 머신러닝 모델은 학습 데이터에 크게 의존하기 때문에, CT 개념을 도입하여 정기적으로 학습 데이터를 개선하고 이를 모델에 재학습시킨다. 이를 통해 모델의 성능을 유지하고, 변화하는 환경에 모델이 대응할 수 있도록 최적화한다. MLOps는 단순히 소프트웨어 코드뿐만 아니라 데이터와 모델을 함께 관리하고, 모델의 학습과 모니터링 과정을 자동화하는 것을 포함한다.

MLOps의 구성 요소로는 데이터 처리, 모델 훈련 및 평가, 모델 서빙 및 모니터링, 자동화된 파이프라인 구축 4가지로 나눌 수 있다[4]. 데이터 처리 단계는 데이터 수집, 전처리, 변환, 피쳐 엔지니어링 등 데이터 관리와 전처리 기능을 지원하며, 모델 학습에 필요한 데이터를 안정적으로 제공하는 역할을 한다. 모델 훈련 및 평가 단계는 학습 환경을 구성하고, 하이퍼파라미터 최적화 같은 기능을 통해 모델을 효율적으로 훈련시킬 수 있으며, 사용자가 상황에 맞는 최적의 모델을 선택할 수 있도록 모델의 성능을 평가 및 비교할 수 있는 기능을 포함한다. 모델 서빙 및 모니터링 단계에서는 학습된 모델을 운영 환경에 배포하고, 성능이 급격하게 떨어지는 경우를 탐지하여 필요한 경우 새로운 데이터를 수집하거나 모델을 재학습시킬 수 있는 시스템을 제공한다. 이를 통해 배포된 모델의 안정성과 신뢰성을 유지하고, 모델이 배포 환경에서도 최적의 상태를 유지할 수 있도록 지원한다. 마지막으로, 자동화된 파이프라인 구축 단계는 모델 개발 및 운영에 필요한 과정을 자동화하여 단순 반복되는 작업을 줄이고 사용자가 시스템을 효율적으로 활용할 수 있도록 도와주는 역할을 한다.



〈그림 2〉 MLOps의 구성 요소 [4]

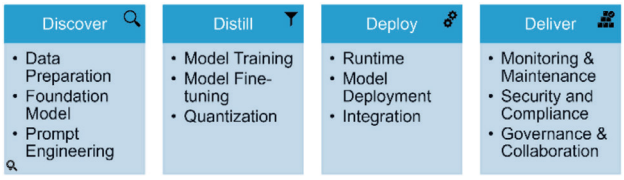
MLOps를 활용하면 소프트웨어를 효율적이고 안전하게 개발할 수 있다는 장점이 있다[5]. 통합된 하나의 시스템에서 데이터 전처리부터 모델 배포까지 자동화된 기능으로 모델을 신속하게 개발할 수 있고, 여러 가지 데이터와 모델을 버전 추적 기능과 일관성 있는 시스템으로 효과적으로 관리하면서 위험 요소를 줄일 수 있다.

LLMOps(Large Language Model Operation)는 MLOps의 원칙을 기반으로 LLM(Large Language Model)을 효율적으로 운영하기 위한 시스템이다[6]. MLOps가 데이터와 머신러닝 모델의 배포 및 관리에 중점을 두었다면, LLMOps는 LLM의 학습 데이터 전처리부터, 모델 학습, 배포, 모니터링, 재학습까지 전 과정을 효율적으로 관리하고, 단계별 복잡한 작업을 간단하게 수행할 수 있도록 설계되어, 모델 학습 및 운영에 필요한 비용을 줄이고, 관련 전문지식이 부족한 사람도 모델을 쉽게 운영할 수 있도록 한다.

LLMOps는 〈그림 3〉처럼 모델을 선택하고 학습 데이터를 구축하는 Discover, 효율적인 학습을 지원하는 Distill, 학습된 모델을 실제 운영 환경에 배포하는 Deploy, 모델을 모니터링하는 Deliver의 4단계로 구분할 수 있다[7]. 가장 먼저, Discover 단계는 모델의 크기, 속도, 정확성 등을 고려하여 배포 환경에 가장 적합한

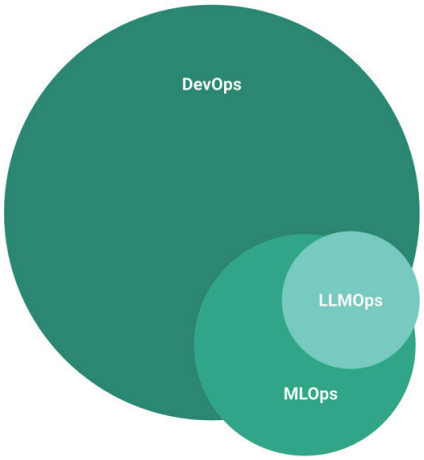
Foundation model을 선택하고, 도메인에 특화된 학습 데이터셋을 구축하는 단계이다. 모델의 성능을 위해선 학습 데이터의 품질이 중요하기 때문에, 데이터셋 구축 단계는 데이터 정제, 전처리, 익명화 등 다양한 기능을 지원하여 학습 데이터의 품질을 개선하고 데이터를 모델이 학습할 수 있는 형태로 변환한다. 또한, 프롬프트 엔지니어링을 통해 사용자가 원하는 출력 형식대로 LLM이 답변을 생성할 수 있도록 입력값의 구조를 설계함으로써 모델의 성능을 최대한 끌어낸다. Distill 단계는 효율적이고 확장 가능한 모델 학습 파이프라인을 구축하는 단계로, 분산 학습과 병렬 처리 기술을 활용하여 학습 시간과 자원을 효율적으로 사용할 수 있도록 지원한다. 양자화 기법을 통해 모델 사이즈를 대폭 줄이고 실행 속도를 향상시킬 수 있다. Deploy 단계는 학습된 모델을 실제 운영 환경에 배포하는 단계로, 모델이 배포 환경에서 안정적으로 작동할 수 있도록 컴퓨팅 자원과 데이터 저장 공간 등 기존 IT 인프라와 통합하는 단계다. 마지막으로, Deliver 단계는 배포 환경에서 동작하는 모델을 지속적으로 모니터링 하는 단계로, 모델의 성능이 떨어지는 상황을 파악하고 관련 데이터를 재학습한다. 그리고 모델의 편향과 환각을 방지하고, 사용자의 적대적 공격으로부터 모델을 보호할 수 있도록 관리하는 과정을 포함한다.

MLOps와 유사하게, LLMOps도 효율성, 위험 감소, 확장성 측면에서 3가지의 장점이 있다[8]. 효율성 측면에서 모델을 운영하기 위해 거쳐야 할 단계를 통합된 하나의 시스템으로 구현하고 이를 자동화함으로써 단순 반복되는 작업을 줄이고 시간과 비용을 절감할 수 있다. 다음으로, 위험 감소 측면에서 모니터링과 피드백 단계를 통해 모델을 향한 사용자의 악의적인 공격 및 환각과 편향 같은 문제를 신속히 감지하고, 대응함으로써 모델의 안정성과 신뢰성을 강화할 수 있다. 마지막으로, 확장성 측면에서 LLMOps는 다양한 환경에 모델을 손쉽게 배포하고 확장할 수 있는 기반을 제공한다. LLMOps는 다양한 도메인과 환경에서 LLM을 효율적으로 운영할 수 있도록 도와주는 통합 시스템으로, 사용자가 LLM에 대한 전문 지식이 없더라도 도메인 데이터를 갖고 있다면, 이를 활용해 모델을 쉽게 학습시키고, 자신만의 특화된 LLM을 구축할 수 있도록 지원한다.



〈그림 3〉 LLMOps의 구성 요소 [7]

〈그림 4〉처럼 DevOps, MLOps 그리고 LLMOps 모두 소프트웨어 개발부터 운영까지의 과정을 효율적으로 통합 및 관리하기 위한 기능들을 제공하지만, 몇 가지 차이점이 있다[9]. DevOps는 주로 코드의 버전 관리와 신속한 배포에 중점을 두며, 상대적으로 데이터나 모델과 같은 요소를 고려하지 않는다. MLOps와 LLMOps는 데이터와 모델 요소도 함께 고려하지만, 시스템을 운영하기 위해 필요한 컴퓨팅 자원의 규모에서 차이가 나는데, MLOps는 일반적인 소규모 GPU 환경에서도 동작하는 반면, LLMOps는 GPU 클러스터나 TPU와 같은 고성능 하드웨어가 필수적이다. 학습 방식에서도 차이가 있다. MLOps는 주로 모델을 처음부터 학습하는 방식이 일반적이지만, LLMOps는 대량의 데이터로 사전 학습된 모델을 기반으로 도메인 특화 데이터를 활용하여 미세 조정을 수행하는 것이 일반적이다. MLOps와 LLMOps 모두 모니터링을 통해 정기적으로 학습 데이터를 개선하고 재학습을 수행하지만, LLMOps는 사용자로부터 모델의 답변에 대한 피드백을 실시간으로 수집하여 RLHF(Reinforce Learning



〈그림 4〉 DevOps, MLOps, LLMOps의 관계 [12]

Human Feedback)[10]를 진행할 수 있다. RLHF는 사용자로부터 모델이 생성한 답변에 대한 평가 데이터를 수집하고, 이를 기반으로 사람이 선호하는 답변을 생성하도록 모델을 학습시키는 방법이다.

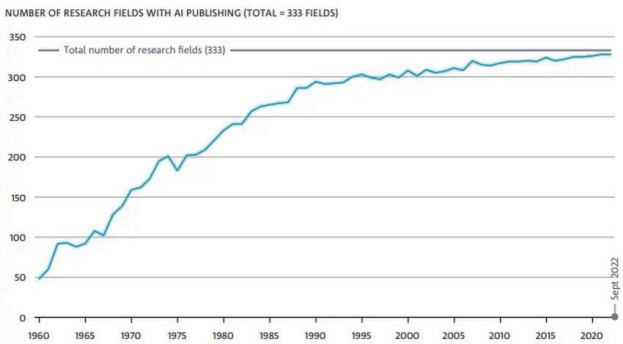
최근에는 멀티모달 LLM의 등장으로 LLMOps의 역할이 더욱 커지고 있다. 멀티모달 LLM은 거대 언어 모델을 기반으로, 텍스트뿐만 아니라 이미지, 비디오, 음성 등 다양한 유형의 데이터를 처리하고 추론할 수 있는 모델로. LLM보다 더 복잡한 작업을 수행할 수 있다[11]. 이 모델을 학습시키기 위해선 이미지, 텍스트, 음성 등 다양한 유형의 데이터가 요구되며, 각 유형에 따라 서로 다른 전처리 기법이 필요한데, LLMOps는 하나의 통합된 플랫폼에서 각 데이터 유형에 적합한 전처리 기능을 지원하여 효율성과 일관성을 제공해야 한다. 또한, LLMOps는 PDF, 이미지, CSV, JSON, 오디오 파일 등 데이터의 다양한 확장자와 형식을 처리할 수 있는 기능을 제공하여, 사용자들이 복잡한 형식 변환 과정 없이 데이터를 쉽게 다룰 수 있도록 도와야 한다.

2. Scientific LLMOps의 필요성

최근 의료 상담을 돕는 AI 챗봇처럼 특정 도메인에 특화된 LLM은 업무의 효율성과 생산성을 향상시키며, 뛰어난 자연어 처리 성능으로 여러 분야에서 적극적으로 활용되고 있다[13]. 이러한 트렌드는 과학기술 분야에서도 마찬가지로, 생산적인 연구를 위해 Scientific LLM에 대한 기대와 수요가 증가하고 있다. 과학기술 분야에 대한 정보량의 증가로 신규 기술을 분석하고 습득하는 데 소요되는 시간도 급격히 늘어나 LLM을 활용한 과학기술 정보 분석 활용 기술 개발에 대한 필요성은 갈수록 높아지고 있다[14]. LLM의 언어 이해 능력을 바탕으로 과학기술 분야의 방대한 문헌들을 빠르게 처리하고, 연구자에게 필요한 정보를 적시에 제공할 수 있다면, 연구 생산성 개선과 기술 혁신에 큰 도움이 될 것이다. 실제로 KISTI는 과학기술 지식인프라 서비스인 ScienceON을 통해 AI가 논문을 읽어주는 서비스를 운영하고 있으며[15], IBM은 NASA와 함께 과학 문헌에 특화된 LLM을 선보였다[16].

구체적으로 Scientific LLM을 활용하여 연구자가 얻을 수 있는 이점은 크게 3가지로 나누어 볼 수 있다[17]. 먼저, 새로운 연구 주제를 빠르게 탐색할 수 있다. 연구자가 아이디어를 모델에 입력하면 Scientific LLM은 학습한 과학 문헌 데이터를 바탕으로 이미 수행된 연구가 있는지 분석하고, 연구 아이디어의 한계점을 파악하여 연구자가 제시한 아이디어를 개선하는 데 도움을 줄 수 있다. 그리고 시간과 비용을 절약할 수 있다. 연구 주제와 방향이 정해지면, 모델은 관련 데이터와 함께 실험 단계를 설계하고, 실험을 반복적으로 수행할 수 있다. 실험을 반복하는 과정에서 발생하는 데이터를 실시간으로 분석하고, 실험 결과에 대한 해석과 함께 시각화 자료를 생성할 수 있다. 마지막으로, 논문을 작성할 때 도움을 받을 수 있다. 연구 결과가 정리되면 모델은 실험 결과와 데이터를 바탕으로 논문의 각 섹션을 작성하고, 관련 문헌을 검색하여 필요한 자료를 제공할 수 있다. 그리고 불필요한 문구를 제거하거나, 문장 구조를 개선하는 등 논문의 품질을 개선할 수 있다. 위와 같은 여러 장점들로 물리학, 생명 공학 등 여러 연구 분야에서 AI가 활용되고 있다. 호주 연방과학산업연구기구의 보고서에 따르면, 333개의 과학 분야 논문을 분석한 결과, AI를 활용하여 논문을 작성하는 분야가 점점 늘고 있으며, 22년에는 98%까지 증가한 것을 〈그림 5〉를 통해 확인할 수 있다[18].

Scientific LLM이 연구자에게 여러 가지 이점을 제공하지만, 이를 도입하기에는 여러 어려움이 따른다[19]. 우선, 해당 도메인과 관련된 고품질 데이터셋이 필요하다. 특히, 과학 분야의 데이터는 복잡한 수식이나 그래프로 존재하는 경우가 많아 일반적인 텍스트 데이터보다 더 복잡하고 까다로운 전처리 과정을 요구한다. 또한, 데이



〈그림 5〉 AI를 사용한 연구 분야 추이 그래프 [18]

터의 품질을 유지하고, 학습에 불필요한 자원 소모를 방지하기 위해 참고 문헌처럼 학습 데이터로 가치가 낮은 데이터를 필터링하는 과정도 필요하다. 현재 Scientific LLM 학습용 데이터셋을 구축하기 위해 많은 연구가 진행되고 있는데, 다음 두 가지의 전략을 채택하고 있다[20]. 첫 번째 전략은 과학 문헌에서 고품질의 과학 corpora를 확보하는 것이다. 대부분의 연구 논문과 실험 결과 보고서는 PDF 형식으로 제공되는데, 이를 텍스트 데이터로 변환하는 과정에서 데이터의 품질이 떨어지기도 한다. PDF 문서에서 텍스트를 정확하게 추출하고, 수식과 그래프 같은 비언어적 요소도 파싱할 수 있는 다양한 방법론이 연구되고 있다. 두 번째 전략은 다양한 과학 분야의 instruction data를 구축하는 것이다. 모델이 과학 문헌을 효과적으로 이해하고 instruction following 작업을 수행하려면 과학 전문 지식에 대한 대량의 고품질 instruction data가 필요하다. 그러나 현재까지 체계적으로 설계된 과학 instruction data가 부족하고, 사람이 수작업으로 구축하고 있지만 큰 비용이 요구되기 때문에, 사람 대신 LLM을 활용한 데이터 합성 기법도 함께 연구되고 있다[21]. LLM을 활용한 데이터 합성은 고품질과 다양성을 갖춘 대규모 데이터를 효율적으로 생성하여 데이터셋 구축 비용을 절감하고 모델의 성능과 확장성을 개선할 수 있다는 장점이 있다. 다음으로, 데이터 보안 문제가 있다. 민감한 연구 데이터나 개인 정보가 포함된 데이터를 모델이 학습할 경우, 모델의 생성물에 해당 정보가 포함될 위험이 있다. 컴퓨팅 자원과 비용 문제도 해결해야 할 중요한 부분이다. LLM은 수십억 개 이상의 파라미터를 포함하는 거대 모델로, 학습과 운영 단계에서 고성능 GPU나 TPU를 포함한 막대한 컴퓨팅 자원과 비용이 필요하다. 하지만 대부분의 연구 기관과 기업은 이러한 자원을 보유하고 있지 않거나, 제한된 예산으로 고성능 인프라를 도입하는 것은 쉽지 않다. 마지막으로, Scientific LLM은 부정확하거나 허위 정보를 생성하는 환각 문제를 겪을 수 있다. 이는 모델의 생성물에 대한 신뢰성을 저하시켜 연구에 부정적인 영향을 미칠 수 있다.

Scientific LLMOps는 이 장벽들을 해결하고 모델을 도입하는 데 도움을 줄 수 있다. 먼저 보안 문제를 해결하기 위해 전처리 단계에서 민감한 정보를 익명화하거나, 모델이 답변을 생성할 때 RAG(Retrieval-Augmented Generation) 기반 데이터 생성 방식을 지원할 수 있다. RAG는 외부 데이터베이스에서 정보를 검색하여 LLM의 응답을 생성하는 기술로[22], 이 방식은 민감한 데이터를 모델에 직접 학습시키지 않고, 모델이 답변을 생성할 때 필요한 데이터를 외부에서 참조하기 때문에 데이터가 모델 내부에 저장되지 않으므로 정보 유출 위험이 줄어든다. 또한, 모델이 답변을 생성하기 위해 참고한 외부 데이터베이스를 출처로 함께 제공하면, 연구자는 정확한 정보를 얻을 수 있고, 관련 자료를 추가로 탐색할 수 있다. 컴퓨팅 자원 문제를 해결하기 위해서 Scientific LLMOps는 클라우드 기반 자원 관리 시스템과 분산 학습 기능을 제공할 수 있다. 클라우드 기반 자원 관리 시스템을 통해 필요한 시간만큼 자원을 할당받을 수 있고, 사용량에 따라 자원을 유연하게 조정할 수 있다. 그리고 분산 학습 기능을 통해 대규모 데이터를 병렬로 처리하여 학습 시간을 크게 단축할 수 있고, 자원을 적절하게 분배하여 효율적인 학습을 가능하게 한다. Scientific LLMOps는 과학기술 분야의 특수성을 반영하여 설계된 플랫폼으로, FASTA, PDB 등 과학 분야에서 널리 사용되는 데이터 형식을 처리할 수 있는 기능을 제공할 수 있다[23]. 이와 함께, 통합된 하나의 시스템에서 데이터 전처리부터 모니터링까지 전 과정을 자동화된 시스템으로 제공함으로써 연구자가 LLM에 대한 전문지식이나 기술 역량이 없어도 각자의 연구 분야에 특화된 LLM을 운영할 수 있도록 연구자를 지원할 수 있다.

MLOps, LLMOps, Scientific LLMOps를 비교 정리하면 〈표 1〉과 같다. MLOps는 상대적으로 작은 머신러닝 모델을 운영할 수 있는 환경을 제공한다. 주로 구조화되고 정제된 데이터를 처리할 수 있는 기능들을 제공하고, 모델 학습, 배포, 모니터링 등 모델 운영과 관련된 일련의 작업을 통합적으로 지원한다. 이러한 시스템은 머신러닝 모델이 금융, 마케팅, 제조 등 다양한 산업 분야에서 업무의 효율성을 개선하고 안정적으로 동작할 수 있도록 한다. LLMOps와 Scientific LLMOps는 수십억 개의 파라미터를 가진 대규모 언어 모델을 운영하기 위한

〈표 1〉 MLOps, LLMOps, Scientific LLMOps 비교 표

	MLOps	LLMOps	Scientific LLMOps
주요 목적	상대적으로 작은 머신러닝 모델 개발 및 활용	수십억 개의 파라미터를 가진 LLM 개발 및 활용	과학기술 분야 데이터에 특화된 LLM 개발 및 활용
데이터	구조화되고 정제된 데이터 처리. 데이터 전처리 과정이 비교적 단순	비구조화된 대규모 텍스트 데이터 처리. 복잡한 전처리 과정 필요	과학 문헌, 그래프, 수식 등 다양한 형태의 비구조화된 데이터 처리. 전문 지식과 함께 복잡한 전처리 과정 필요
역할 및 주요 기능	피쳐 엔지니어링, 모델 학습, 배포, 모니터링	사전 학습된 LLM에 특정 분야 데이터를 추가로 fine-tuning 진행	사전 학습된 LLM에 과학기술 분야 데이터를 활용하여 fine-tuning 진행
활용 분야	금융, 마케팅, 제조 등 다양한 산업 분야에서 분류, 회귀 등의 특정 태스크 중심	자연어 처리, 텍스트 생성, AI 챗봇 등 언어 기반의 애플리케이션	논문 요약, 실험 데이터 분석 등 과학연구 분야의 연구 지원 AI 개발 및 지원 서비스

환경을 제공한다. 두 시스템 모두 비구조화된 대규모 텍스트 데이터를 전처리할 수 있는 기능을 지원하고, fine-tuning, 모니터링 등 LLM을 운영하기 위해 필요한 작업을 체계적으로 관리할 수 있다. 이를 통해 자연어 처리, 텍스트 생성, AI 챗봇 등 언어 기반 애플리케이션에서 LLM이 안정적으로 동작할 수 있는 기반을 마련한다. Scientific LLMOps는 기존의 LLMOps의 기능을 확장하여 연구자들이 과학기술 분야에 특화된 LLM을 손쉽게 제작 또는 활용할 수 있는 환경을 제공해야 한다. 과학 문헌, 그래프, 수식 등 다양한 형태의 데이터 전처리를 기능과 도메인 지식을 반영한 데이터 증강 및 학습 도구를 통해, LLM이 과학기술 데이터를 효과적으로 학습하고 해당 분야에 특화된 모델로 발전할 수 있도록 해야 한다. 이를 통해 연구 보조자로서 연구자를 지원할 수 있는 시스템을 구축할 수 있다. Scientific LLM은 연구의 생산성을 향상시킬 수 있는 중요한 도구이며, Scientific LLMOps를 통해 더 많은 연구자들이 LLM이라는 도구를 연구에 활용할 수 있도록 지원함으로써, 연구 혁신과 과학기술 발전에 기여할 수 있을 것이다.

III. AIDE: KISTI Scientific LLMOps

한국과학기술정보연구원 초거대AI연구단에서는 '24년도부터 과학기술 분야의 다양한 AI 수요에 대응하기 위해 Scientfic LLMOps에 대한 연구개발을 수행하고 있다. AIDE(AI Development Environment, 에이드)라고 명명된 이 플랫폼을 기반으로 복잡한 LLM 기술 개발을 간소화하여 과학기술 분야에 필요한 데이터 구축, 모델 개발 및 적용을 지원한다. 과학기술 정보 특화 LLM 개발 및 확산을 목표로 과학기술 분야의 다양한 문헌 및 전문 데이터의 전처리, 모델 학습, 모델 평가 및 운영까지 LLM 개발 전주기를 통합 관리하고, 자동화하는 기술 및 서비스 시스템을 연구개발하고 있다.

1. AIDE 플랫폼의 개발 목적

과학기술 분야 지원 Scientific LLMOps 구축의 첫 단계로, KISTI에서는 〈그림 6〉과 같이 거대언어모델 연구와 활용의 과정을 통합하여 사용자 요구에 맞는 LLM 개발을 효과적으로 지원하기 위한 AIDE 플랫폼을 개발하고 있다.

1) LLM 전주기 프로세스 통합 관리: 데이터 전처리 기능 등과 모델 학습, Fine-Tuning, RLHF/DPO, 모델 평가 및 배포까지의 전 과정을 통합적으로 관리한다. 모델 성능 평가, 자동 하이퍼파라미터 최적화, 버전 관리를 통

〈그림 6〉 AIDE 개발 범위 및 목표

해 LLM 성능을 지속적으로 개선할 수 있도록 하며, 모델 배포와 관리를 지원한다.

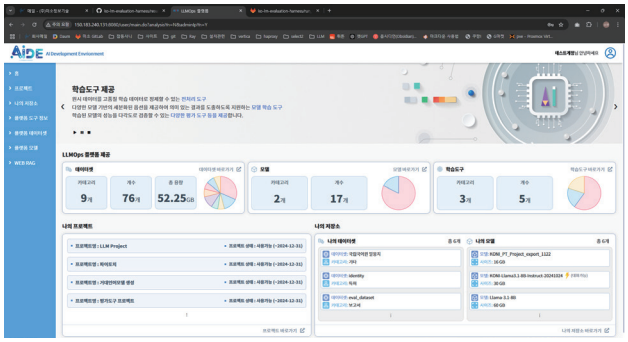
2) 과학기술 데이터 전처리 및 학습데이터 생성: 논문 및 특허 등 과학기술 분야의 데이터를 모델에 학습시키기 위해 필요한 학습데이터 생성을 지원한다. 데이터 전처리 기능(유사도 중복 제거, 개인정보 제거 등)으로 대규모 데이터에 전처리 수행을 지원한다.

3) 대규모 GPU 활용을 지원하는 분산 컴퓨팅 환경: 딥러닝 기반의 모델 훈련 및 대규모 데이터의 전처리를 위해 필요로 하는 분산 병렬 컴퓨팅 자원을 제공한다. Kubernetes 기반 Auto-Scaling 및 컨테이너화 기술을 활용하여 대규모 데이터와 컴퓨팅 자원을 효율적으로 관리하고 초거대AI 연구 생산성을 극대화한다. Ray 등과 같은 오픈소스 프레임워크를 활용하여 대규모 분산 학습, 모델 서빙, 하이퍼파라미터 튜닝 등을 지원한다.

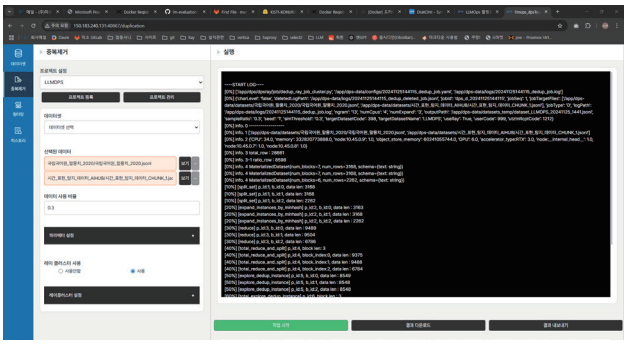
4) 웹 기반 통합 LLM 개발 환경 제공: 인공지능 경험이 부족한 일반 사용자도 간단히 데이터와 모델을 선택하는 것 만으로 고도화된 LLM 개발 과정을 활용할 수 있도록 직관적이고 접근성이 높은 환경을 제공한다. 또한, 사용자 등록, 데이터 및 모델 관리, 학습 및 평가 프로세스 모니터링, 로그 알림 및 대시보드 기능을 통해 다중 사용자 중심의 운영 및 관리 환경을 구축한다.

2. 주요 기능

AIDE는 Scientific LLM 개발자가 소유한 과학 데이터의 전처리를 수행하고, 공개된 다양한 LLM 모델을 활용하여 맞춤형 LLM 모델을 개발 및 활용할 수 있는 환경을 지원한다. AIDE를 활용해 다음의 데이터 전처리, 모델



〈그림 7〉 AIDE 사용자 메인 화면

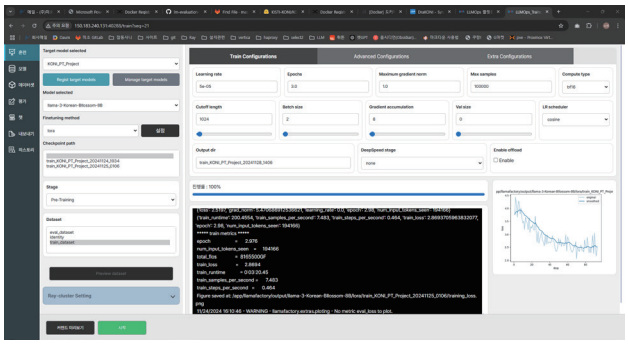


〈그림 8〉 데이터 전처리 기능

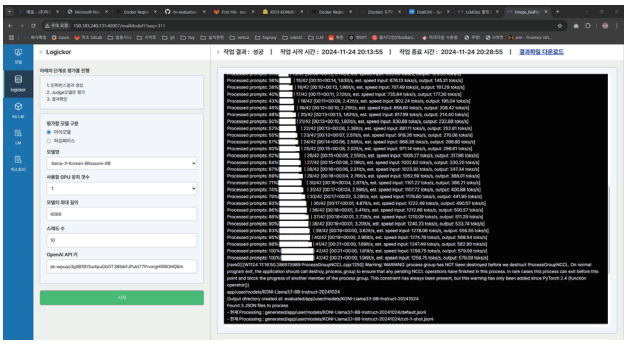
학습, 모델 평가 등 주요 기능을 웹 기반 유저 인터페이스로 사용할 수 있다.

• 데이터 전처리: AIDE 플랫폼의 데이터 전처리 도구는 EleutherAI/dps 기반으로 구현되어 있으며, 웹 애플리케이션 형태로 제공된다[24]. 유사 데이터 중복 제거, 필터링 조건 설정, 특정 패턴 검색 및 제거 등과 같은 전처리 관련 작업을 직관적인 인터페이스를 통해 사용할 수 있다. 기존 파일 형태의 수동 처리 방식을 데이터베이스 형태의 자동 처리 방식으로 개선하여 데이터 전처리 과정 및 결과를 로그로 저장 및 확인 등 이력 관리가 가능하게 하였다. AIDE의 데이터 전처리 기능은 분산처리 기반 대규모 데이터 환경에서 자동화와 최적화를 구현하여 연구자들이 데이터 품질과 처리 효율성을 극대화해 LLM 연구에 집중할 수 있도록 하였다[25][26].

• 모델 학습: 거대언어모델의 훈련 지원을 위해 오픈소스 기반 다양한 기능을 통합한 학습 환경을 제공하여 사용자가 모델 학습에 있어 사전학습 기능 외 다양한 학습 전략을 지원받고, 웹 기반 인터페이스를 활용하여 쉽게 설정하고 활용할 수 있다. 모델 학습 과정에서는 Fine-Tuning, PPO, DPO, ORPO, Pre-Training 등의 다양한 학습 방법론을 처리할 수 있다. 사용자는 학습 데이터셋과 모델을 선택하여 학습을 수행하거나 Learning Rate, Epochs, Batch Size 등 하이퍼파라미터를 수정하여 학습 과정을 세부적으로 조정하여 학습을 수행할 수 있다. 학습 완료 후 각 모델 별 사전학습 가중치 파일을 별도의 데이터베이스에 관리하여 안정적이고 체계적인 학습 환경을 보장한다. 플랫폼은 Ray Framework를 기반으로 분산 병렬 처리를 구현하여, 대규모 모델 학습 시



〈그림 9〉 모델 훈련 기능



〈그림 10〉 모델 평가 기능

에도 높은 성능과 효율성을 유지한다. 사용자는 플랫폼에 등록된 모델 외 개인 모델 또는 HuggingFace 기반 모델을 선택하여 맞춤형 학습을 수행할 수 있으며, 학습 결과는 플랫폼에 저장되어 다른 연구자와 공유할 수 있다. 학습 완료된 모델의 응답 성능을 바로 확인할 수 있는 Chat 기능도 지원되어 사용자가 해당 모델의 성능을 가늠해 볼 수 있다. 그 외 AIDE는 모델 학습 시 부가 설정 기능을 제공하기 때문에 옵티마이저 선택, 임베딩 레이어 크기 조정, 시퀀스 샘플 압축, 실험 로그 기록 등 다양한 실험 조건을 확인 및 반복할 수 있다. 이러한 기능은 사용자의 요구에 따라 미리 설정된 모듈로 제공할 수 있게 LLaMA-Factory 오픈소스 기반으로 개발하고 있으며[27], 모델 개발 과정의 복잡성을 줄이기 위해 루틴화할 수 있는 기능을 중심으로 모델 학습 기능을 간소 · 자동화할 계획이다.

• 모델 평가: AIDE는 생성형 거대언어모델의 성능을 체계적이고 다각도로 평가하기 위해 LogicKor, ko-lm-evaluation-harness, lm-evaluation-harness와 같은 3종의 평가 도구를 활용할 수 있게 개발하였다[28][29][30]. 이들 도구는 각각의 특징을 살려 객관적이고 다양한 평가 지표를 제공하며 모델의 품질을 가늠해볼 수 있는 기능을 지원하는데, AIDE 플랫폼은 이들 평가 도구를 통합하여 제공함으로써 연구자들이 해당 LLM의 성능을 각 평가 도구 별로 분석하고, 모델의 강점과 약점을 효과적으로 파악할 수 있게 한다. 마찬가지로 웹 기반 인터페이스를 통해 평가 설정과 실행이 가능하며, 평가 결과는 시각화 및 보고서 형태로 제공되어 사용자들이 모델 개선 방향을 빠르게 도출할 수 있다.

이와 같이 AIDE는 데이터 전처리, 모델 학습, 모델 평가 등 과학기술 특화 LLM 개발의 전 과정을 체계적이고 효율적으로 지원하여 연구자가 고품질의 LLM을 개발할 수 있도록 한다. 데이터 전처리 기능은 대규모 데이터를 효과적으로 관리 및 고품질 데이터셋 생성하고, 모델 학습 기능은 오픈소스 기반 다양한 학습 기법과 분산 병렬 처리 환경을 제공해 최적의 성능을 가진 모델을 학습할 수 있도록 하였다. 모델 평가 기능은 정밀한 성능 분석과 시각화된 결과를 통해 모델의 품질을 객관적으로 판단할 수 있게 한다. 이런 기능 활용을 통해 연구자는 작업을 효율적으로 수행하고 플랫폼의 생산성과 연구 품질을 극대화할 수 있고, 연구자가 필요로 하는 도구와 자원을 제공하여 KISTI Scientific LLMOps를 위한 기반을 마련하고자 한다[31][32].

IV. Scientific LLMOps의 과제

Scientific LLMOps는 과학기술 분야의 인공지능 기술 확산을 극대화하기 위한 핵심 기반으로서 앞서 많은 연구 개발과 함께 이를 실제 연구 현장에 안정적으로 정착시키기 위해서는 해결해야 할 많은 과제들을 안고 있으며, 주요한 도전 과제는 다음과 같다.

1) 도메인 특화 LLM 개발 지원: 과학기술은 물리학, 화학, 생명과학, 공학 등 다양한 학문 분야로 이루어져 있으며, 각 분야는 고유한 데이터 구조와 문제 해결 방식을 가지고 있으며, 분야에 더 특화된 전문가 AI가 필요할 것이다. 도메인별로 맞춤형 LLM을 개발하기 위해서는 고품질의 특화된 데이터셋이 필요하며, 이를 지속적

으로 검증하고 업데이트할 수 있는 체계가 마련되어야 한다. 또한, 특정 연구 주제에 맞게 모델을 신속하게 Fine-tuning할 수 있는 유연한 학습 환경과, 도메인 간 데이터와 모델의 상호 운용성을 높이는 표준화된 인터페이스도 중요하다.

2) 학습 데이터 품질과 신뢰성 관리: 과학기술 분야의 연구는 정확성과 신뢰성이 필수적이기 때문에, 학습 데이터의 품질이 곧 모델의 성능과 직결된다. 데이터셋을 자동으로 검증하고 오류를 탐지할 수 있는 기능이 필요하며, 생성된 답변이나 분석 결과에 대한 출처 명시와 신뢰도 평가가 이루어져야 한다. 이를 통해 연구자들이 제공된 정보를 신뢰할 수 있고, 필요시 출처를 추적해 추가 검토가 가능해야 한다. 특히 방대한 비구조적 데이터를 다루는 과학 분야의 특성상, 자동화된 데이터 검증 및 전처리 도구 개발이 필수적이다.

3) 자동화된 모델 개발 및 활용 체제: LLM을 학습하고 운영하기 위해서는 막대한 컴퓨팅 자원이 필요하다. 이를 해결하기 위해 클라우드 기반의 인프라를 적극 활용해 자원을 유연하게 관리하고, 고성능 GPU 및 TPU를 최적화하는 분산 학습 기술이 필요하다. 또한, 사용자 경험(UX) 개선은 LLMOps의 실질적 활용성을 높이는 중요한 요소다. 또한, 연구자들이 복잡한 AI 기술 및 훈련과정에 대한 깊은 이해없이도 직관적으로 플랫폼을 사용할 수 있어야 하며, No-Code/Low-Code 환경을 통해 누구나 쉽게 모델을 학습하고 배포할 수 있도록 지원해야 한다. 또한, 대시보드와 시각화 도구를 통해 학습 및 평가 결과를 직관적으로 확인할 수 있는 기능이 필요하며, 연구자들 간의 협업과 데이터 및 모델 공유를 지원하는 커뮤니티 기능도 강화해야 한다.

V. 전망과 결론

대규모 언어 모델(LLM)의 등장은 과학기술 분야에 새로운 패러다임을 열고 있다. 이제 우리는 단순히 데이터를 분석하거나 분류하는 수준을 넘어, AI가 방대한 과학 문헌과 실험 데이터를 스스로 이해하고, 논리적 추론을 통해 창의적인 해결책을 제안하는 시대를 맞이하고 있다. 하지만 이 혁신이 실제 연구 현장에서 실질적 가치를 발휘하기 위해서는 도메인 특화 모델 개발, 데이터 품질 관리, 대규모 컴퓨팅 자원 효율화, 사용자 경험 개선, 그리고 윤리적 운영 체계가 유기적으로 결합되어야 한다. 이러한 복잡한 도전 과제를 해결하기 위해 Scientific LLMOps는 단순한 기술 운영 도구를 넘어 과학기술 연구의 새로운 기반 인프라로 자리 잡을 것으로 기대된다.

Scientific LLMOps는 데이터 전처리부터 모델 학습, 평가, 배포, 모니터링까지의 모든 단계를 하나의 플랫폼에서 통합 관리한다. 특히 과학기술 분야의 특수성을 반영한 데이터 처리, Fine-tuning 환경, 사용자 친화적 인터페이스, 그리고 클라우드 및 고성능 컴퓨팅 자원의 효율적 관리 기능을 통해 연구자들이 기술적 장벽 없이 AI를 활용할 수 있도록 돕는다. 더 나아가 연구 재현성을 높이고, 신뢰할 수 있는 결과를 도출하며, 협업과 데이터 공유를 촉진하는 체계를 제공함으로써 연구의 생산성과 효율성을 극대화할 수 있다. 이는 단순한 기술적 도구를 넘어 연구자의 사고방식과 작업 방식을 근본적으로 변화시키는 힘을 가지고 있다.

앞으로 Scientific LLMOps는 AI 기반 과학기술 혁신의 중추적 역할을 수행하며, 국가와 글로벌 기술 경쟁력을 견인하는 핵심 엔진이 될 것이다. 이 플랫폼은 단순한 기술 도입을 넘어 과학기술 연구자들이 더 깊이 있는 탐구와 창의적 문제 해결에 집중할 수 있도록 지원할 것이다. AI 기술은 이제 더 이상 미래의 기술이 아니라, 현재 우리가 직면한 문제들을 해결하는 필수적인 동반자가 되고 있다. Scientific LLMOps가 더 많은 연구자와 기관의 LLM 요구에 부응하기 위한 AI 생산공장으로의 역할이 크게 된다.

참 고 문 헌

[1] Georgios Symeonidis, et al. "MLOps – Definitions, Tools and Chllenges", IEEE, 2022

[2] Yan Liu, et al. "Building A Platform for Machine Learning Operations from Open Source Frameworks", IFAC-PapersOnLine, 2020

[3] Dominik Kreuzberger, et al. "Machine Learning Operations

(MLOps): Overview, Definition, and Architecture", IEEE, 2023

[4] Khaild Salama et al. "Practitioners guide to MLOps: A framework for continuous delivery and automation of machine learning", Google Cloud, 2021

[5] Nipuni Hewage, et al. "Machine Learning Operations: A Survey on MLOps Tool Support", arXiv: 2202.10169, 2022

[6] Megha Sinha, et al. "LLMOps: Definitions, Framework and Best Practices", ICECET, 2024

[7] Richard Shan, et al. "Enterprise LLMOps: Advancing Large Language Models Operations Practice", IEEE, 2024

[8] IBM, "What are large language model operations(LLMOps)?", 2023

[9] Hazal Simsek, "LLMOPs vs MLOPs: Discover the Best Choice for You", AIMultiple research, 2024

[10] Long Ouyang, et al. "Training langugage models to follow instructions with human feedback", NeurIPS, 2022

[11] Shukang Yin, et al. "A Survey on Multimodal Large Language Models", IEEE, 2024

[12] Tim Leers, "Fron MLOps to LLMOps–what's the difference?", dataroot, [Online]. Available: https://dataroots.io/blog/from-mlops-to-llmops-whats-the-difference, [Accessed: 24-Dec-2024].

[13] 김나혜, "메디허브 AI 기반 24시간 의료 상담 서비스 '닥터나이트' 출시", 중앙일보, 2024

[14] 김재수, "과학기술정보 특화 한국어 LLM 공학박사 수준의 질의응답 가능해질 것", KISTI, 2024

[15] 손영주. "AI가 읽어주는 'ScienceON 논문 요약 서비스' 개시" KISTI, 2020

[16] Bishwaranjan Bhattacharjee, et al. "IBM and NASA build language models to make scientific knowledge more accessible", IBM, 2024

[17] Chris Lu "The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery", arXiv: 2408.06292, 2024

[18] CSIRO, "Artificial intelligence for science", Australia's National Science Agency, 2022

[19] Songlin Yu, "Large-language models: The game-changers for materials science research", ELSEVIER, 2024

[20] Sihang-Li, et al. "SciLitLLM: How to Adapt LLMs for Scientific Literature Understanding", arXiv: 2408.15545, 2024

[21] Ke Wang, et al. "A Survey on Data Synthesis and Augmentation for Large Language Models", arXiv: 2410.12896, 2024

[22] 이은빈, 배호, "A Survey on the Latest Research Trends in Retrieval-Augmented Generation", TKIPS, 2024

[23] Yu Zhang, "A Comprehensive Survey of Scientific Large Language Models and Their Applications in Scientific Discovery", EMNLP, 2024

[24] EleutherAI/dps, [Online]. Available: https://github.com/EleutherAI/dps, [Accessed: 30-Dec-2024]

[25] Rae-young Jang, et al. "Improving Efficiency in LLM Data Preprocessing based on Distributed Computing Frameworks", ICCC, 2024

[26] Ja-hyun Ku, et al. "Boosting LLM Performance through LLM-Driven Data Quality Assessment", ICCC, 2024

[27] Yaowei Zheng, et al. "LLAMAFACTORY: Unified Efficient Fine-Tuning of 100+ Language Models", arXiv:2403.13372, 2024

[28] instructkr/LogicKor, [Online]. Available: https://github.com/instructkr/LogicKor, [Accessed: 30-Dec-2024]

[29] Beomi/ko-lm-evaluation-harness, [Online]. Available: https://github.com/Beomi/ko-lm-evaluation-harness, [Accessed: 30-Dec-2024]

[30] EleutherAI/lm-evaluation-harness, [Online]. Available: https://github.com/EleutherAI/lm-evaluation-harness, [Accessed: 30-Dec-2024]

[31] Gun-woo Lee, et al. "Design of an Agent for Scientific Research and Development", ICCC, 2024

[32] Ryong Lee, et al. "Streamlining Domain-Specific AI Model Development An LLMOps Platform for Automated Model Creation and Deployment", ICCC, 2024

이 용

- 2003년 9월 일본 교토대학 정보학과 박사
- 2003년 10월 ~ 2008년 6월 삼성종합기술원 연구원
- 2008년 7월 ~ 2011년 3월 일본 효고현립대 연구교수
- 2011년 4월 ~ 2013년 9월 일본 NICT 연구원
- 2013년 10월 ~ 현재 한국과학기술정보연구원 책임연구원
- 2020년 1월 ~ 현재 UST 한국과학기술정보연구원 캠퍼스 응용AI 전공 교수

〈관심 분야〉
LLM, LLMOps, Multimodal LLM, Agent, AGI, Humanoid

구자현

- 2006년 8월 충남대학교 컴퓨터공학 박사
- 2007년 1월 ~ 2008년 8월 충남대 소프트웨어연구소 전임연구원
- 2008년 9월 ~ 2010년 7월 미국 애리조나대 Postdoc.
- 2010년 8월 ~ 2013년 2월 KAIST 전산학과 연구교수
- 2013년 5월 ~ 2014년 11월 한국전자통신연구원 선임연구원
- 2014년 12월 ~ 현재 한국과학기술정보연구원 책임연구원
- 2018년 9월 ~ 현재 한국정보과학회 데이터소사이어티 상임이사
- 2020년 1월 ~ 12월 정보과학회지 편집위원
- 2020년 9월 ~ 현재 UST KISTI 스쿨 응용AI전공 교수
- 2023년 ~ 현재 과기부 미래국방기술전략 AI/무인자율 분과위원

〈관심 분야〉
Large Language Model(LLM), Natural Language Processing, Deep Learning, Scalable Computing

장래영

- 2013년 2월 한남대학교 공과대학 컴퓨터공학 석사
- 2018년 8월 한남대학교 공과대학 컴퓨터공학 박사
- 2019년 9월 ~ 현재 한국과학기술정보연구원 선임기술원

〈관심 분야〉
LLMOps, Distributed Data Processing, Vision-based AI Models, Auto-scaling for AI Workloads, Cross-Environment Compatibility

이건우

- 2020년 2월 KAIST 전산학부 박사
- 2020년 4월 ~ 11월 한국교통대학교 연구교수
- 2020년 12월 ~ 현재 한국과학기술정보연구원 선임연구원

〈관심 분야〉
Artificial Intelligent, Indoor Positioning, Large Language Model (LLM), Prompt Engineering, LLM Agent System

이경하

- 2006년 8월 충남대학교 컴퓨터공학 박사
- 2007년 1월 ~ 2008년 8월 충남대 소프트웨어연구소 전임연구원
- 2008년 9월 ~ 2010년 7월 미국 애리조나대 Postdoc.
- 2010년 8월 ~ 2013년 2월 KAIST 전산학과 연구교수
- 2013년 5월 ~ 2014년 11월 한국전자통신연구원 선임연구원
- 2014년 12월 ~ 현재 한국과학기술정보연구원 책임연구원
- 2018년 9월 ~ 현재 한국정보과학회 데이터소사이어티 상임이사
- 2020년 1월 ~ 12월 정보과학회지 편집위원
- 2020년 9월 ~ 현재 UST KISTI 스쿨 응용AI전공 교수
- 2023년 ~ 현재 과기부 미래국방기술전략 AI/무인자율 분과위원

〈관심 분야〉
Large Language Model(LLM), Natural Language Processing, Deep Learning, Scalable Computing

영상 초해상화 인공지능 모델의 양자화 기술

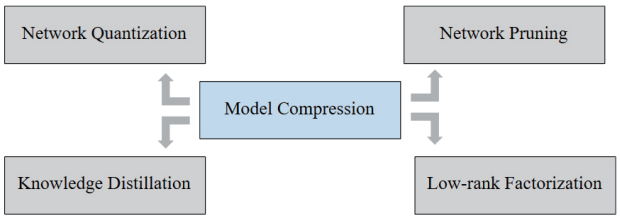
Ⅰ. 서 론

이미지 초해상화 (Image Super-Resolution, SR) 기술은 저해상도 (Low-Resolution, LR) 이미지를 고해상도 (High-Resolution, HR) 이미지로 복원하는 기술로, 다양한 산업과 연구 분야에서 중요한 역할을 한다. 예를 들어, 의료 영상에서 사용되는 MRI와 CT 데이터는 해상도가 제한적인 경우가 많아, SR 기술을 활용하면 병변의 세부 정보를 더욱 명확히 관찰할 수 있어 진단 정확도를 높일 수 있다. 또한, 오래된 영상이나 저해상도로 스트리밍되는 비디오의 품질을 개선해 사용자 경험을 향상시키거나, 위성 사진과 감시 카메라 영상과 같은 데이터를 고해상도로 복원하여 다양한 분석 작업에 활용할 수 있다. SR 기술은 단순히 이미지를 확대하는 데 그치지 않고, 원본 데이터의 한계를 극복하며 여러 응용 분야에서 새로운 가치를 창출하고 있다.

SR 기술은 여러 접근 방식을 통해 지속적으로 발전해 왔다. 초기에는 주변 픽셀 값을 기반으로 이미지를 보간하여 해상도를 높이는 보간법 (interpolation)이나, 이미지의 통계적 특징과 구조를 분석해 고해상도를 예측하는 예측 기반 기법이 주로 사용되었다. 최근에는 딥러닝과 결합된 접근 방식이 주목받고 있다. 신경망을 활용해 저해상도와 고해상도 이미지 간의 복잡한 매핑을 학습하는 기법이 높은 성능을 보이고 있다. 특히, 합성곱 신경망 (Convolutional Neural Network, CNN)은 다층 네트워크 구조를 통해 저해상도 입력과 고해상도 출력 간의 비선형적 관계를 효과적으로 모델링하여 보다 정확한 결과를 생성할 수 있다. 이처럼 SR 기술은 제한적인 자원에서도 고품질 이미지를 생성할 수 있는 강력한 도구로 자리 잡았으며, 딥러닝과의 결합을 통해 앞으로도 높은 성장 가능성을 가진 핵심 기술로 평가받고 있다.

그러나 딥러닝 기반의 SR 기술은 뛰어난 성능을 보여주지만 높은 계

홍채은
서울대학교



〈그림 1〉 모델 경량화를 위한 연구

산 비용을 요구하는 복잡한 네트워크 구조에 의존하고 있다. 예를 들어, 딥러닝 기반 SR 모델의 대표적인 모델 중 하나인 EDSR[1]의 경우 파라미터 수가 약 43M개로 1080p 이미지 한 장의 해상도를 2배 키울 때 2080Ti GPU 기준 약 2~3초 걸린다. 해상도가 큰 이미지가 입력될 경우 계산량은 보다 커지며, 모바일 기기에서는 이러한 계산량을 처리하기 어렵다. 또한, 목표로 하는 이미지의 해상도가 높아질수록 계산 비용이 기하급수적으로 증가하며, 최근 4K 및 8K 영상에 대한 수요 증대로 인해 SR 기술의 연산 부담은 이전보다 훨씬 커지고 있다. 이러한 특성은 모바일 디바이스나 엣지 컴퓨팅 환경에서 SR 기술의 활용을 제한하는 주요 요인 중 하나다. 이 같은 계산 복잡성을 완화하기 위해 다양한 모델 경량화 기술이 제안되고 있다 〈그림 1〉. 직접 새로운 가벼운 모델 구조를 설계하는 방법 외에도 기존에 존재하는 모델에 경량화 기술을 적용하는 네트워크 양자화 (network quantization), 네트워크 가지치기 (network pruning), 지식전이 (knowledge distillation), 저차원 행렬 분해 (low-rank factorization) 등의 방법이 있다. 그 중에서도 네트워크 양자화는 모델 추론 속도와 메모리 사용량을 줄이는 유망한 대안으로 주목받고 있다.

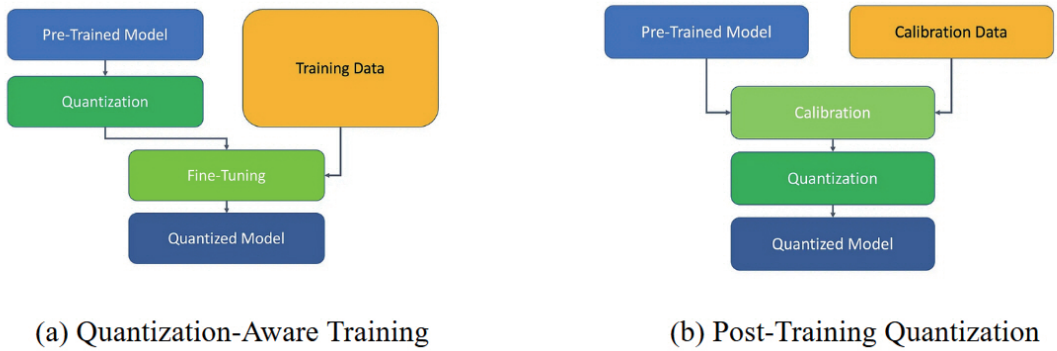
양자화는 신경망의 가중치와 활성화 값을 높은 정밀도 (floating-point 32-bit, FP32)에서 더 낮은 정밀도 (INT8, INT4 등)로 변환하여 연산 효율성을 높이는 기법이다. 기존의 연산을 낮은 정밀도의 연산으로 대체하여 메모리 사용량은 줄어들고 처리 속도는 크게 향상된다는 이점이 있다. 그러나 양자화를 통해 정밀도가 낮아지면 표현 가능한 숫자의 범위가 줄어들고 이에 따라 정보 손실이 발생하며 이는 네트워크의 성능 저하로 이어질 위험이 있다. 따라서, 양자화 연구들의 근본적인 목표는 기존 모델의 성능을 유지하면서 계산 비용과 메모리 요구사항을 줄이는 데 있다.

이에 따라서 SR 문제에서도 높은 품질의 이미지를 생성하면서도 경량화된 모델을 제공하기 위해 다양한 양자화 기법이 활발히 연구되고 있다. 특히, 엣지 디바이스와 같은 리소스 제약 환경에서도 SR 기술을 적용하기 위해 양자화는 필수적인 기술로 자리 잡았다.

본 기고문에서는 딥러닝 기반 이미지 초해상화 알고리즘에 양자화를 적용하는 최신 기술 동향을 조사하고, 각 기법의 장단점과 한계를 분석한다. 이를 통해, 양자화를 적용한 SR 기술의 가능성을 체계적으로 정리하고, 향후 연구 방향을 제시하고자 한다.

Ⅱ. SR 모델의 양자화

이미 네트워크 양자화는 이미지 분류와 같은 고수준 시각 작업 (high-level vision task)에서 모델 예측의 정확



(a) Quantization-Aware Training (b) Post-Training Quantization

〈그림 2〉 양자화 기법의 학습 방식에 따른 구분

도를 크게 손상시키지 않으면서 계산 비용을 줄이는 데 효과적임이 입증되었다. 하지만 보다 복잡한, 픽셀 단위의 예측을 해야 하는 이미지 초해상화 문제를 위한 네트워크를 낮은 비트로 양자화할 경우, 성능이 크게 손상되는 문제가 발생하며, 이는 여전히 도전적인 과제로 남아 있다.

이러한 문제를 해결하기 위해 다방면의 연구가 진행되고 있다. SR 모델의 양자화 기법 연구는 학습 방식, 적용 범위, 정밀도 수준에 따라 구분할 수 있다. 이 장에서 각 구분에서 구체적으로 어떤 연구가 이루어지고 있는지 기술하고자 한다.

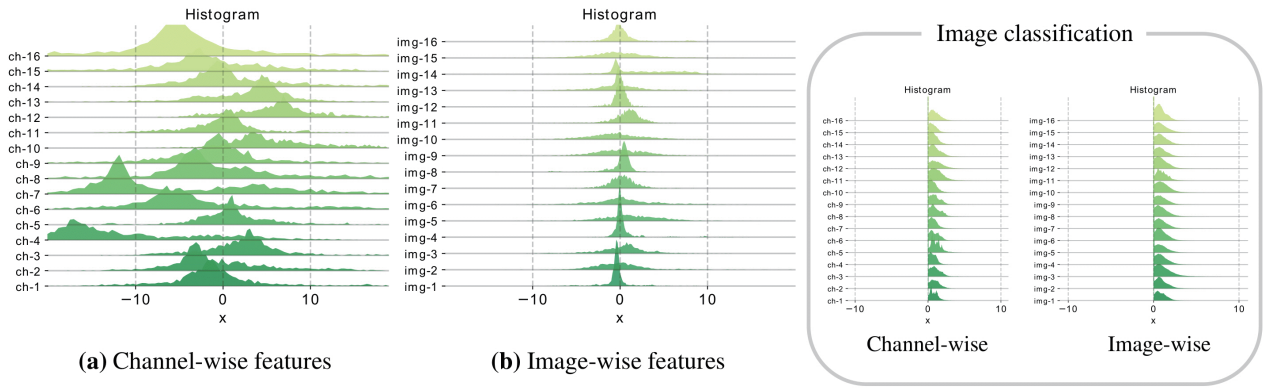
1. 학습 방식에 따른 구분

모델의 양자화는 학습 방식에 따라 크게 Quantization-Aware Training (QAT)과 Post-Training Quantization (PTQ)로 나눌 수 있다 〈그림 2〉.

Quantization-Aware Training (QAT)

QAT는 모델 학습 과정에서 양자화를 시뮬레이션하여, 모델이 양자화로 인한 정밀도 손실을 학습 중에 보정할 수 있도록 하는 방식이다. 학습 중 가중치와 활성화 값을 낮은 정밀도의 데이터 타입으로 변환한 후 연산을 수행하고, 역전파(backpropagation)는 FP32의 정밀도를 유지한 상태로 진행해 그라디언트 정보의 손실을 방지한다. QAT는 양자화된 모델의 성능을 높은 수준으로 유지할 수 있다는 장점이 있지만, 학습 과정이 복잡하고 시간이 오래 걸리며, 학습 과정에서 추가적인 학습 데이터와 하드웨어 자원이 요구될 수 있다.

SR 모델에서의 양자화 연구는 대부분 QAT를 활용하는 연구가 이루어졌다. 이는 SR 모델이 갖는 활성화 값 분포의 고유한 특성 때문으로, SR 모델의 활성화 값 분포는 네트워크의 층마다 크게 다르며, 비대칭적인(asymmetric) 분포를 갖는 동시에 outlier가 존재하는 것으로 알려져 있다. 이러한 분포적 특성은 양자화를 어렵게 만드는 주요 요인이다. 양자화 과정에서 활성화 값 또는 가중치가 적절히 표현되기 위해서는 양자화 함수에 사용되는 스케일링 팩터(scaling factor)와 제로 포인트(zero point)를 정확하게 설정하는 것이 중요하다. 그러나 비대칭적이고 outlier가 존재하는 분포에서는 스케일링 팩터를 설정하기가 어렵다. 예를 들어, 구현이 간단하여 이미지 분류에서 흔히 쓰이는 최대값 기반 양자화 함수(max quantization function)를 그대로 SR 모델에 적용할 경우, outlier의 존재로 인해 양자화 범위가 왜곡될 수 있다. 이러한 왜곡은 분포 내 실제로 많은 값이 밀집한 범위를 세밀하게 표현하지 못하게 만들며, 이는 SR 모델에서 큰 성능 저하로 이어질 가능성이 높다. 이를 해결하기 위해 SR 모델의 양자화 연구 중 초기 연구에서 제안된 PAMS[2]는 모델의 각 층의 스케일링 팩터를 학습 가능한 파라미터로 설정하여, 층별로 outlier를 고려한 적절한 양자화 값을 학습하도록 설계되었다. 이는 기존의 이미지 분류 모델에서 성공적으로 적용되었던 양자화 기법을 SR 모델에 직접 적용한 방법과 비교했을 때 큰 성능 향상을 달성하며 SR 모델에 특화된 양자화 기법의 가능성을 보



(a) Channel-wise features **(b)** Image-wise features

〈그림 3〉 채널별, 이미지별 상이한 SR 모델 (EDSR)의 활성화 값 분포

여주었다.

그러나, SR 모델의 활성화 값 분포는 네트워크의 층에 따라 다른 것 뿐 아니라, 입력 데이터에 따라 크게 변화하는 특성이 있다〈그림 3b〉. 따라서 설정된 양자화 범위가 새로운 입력 데이터에 잘 맞지 않아 정보 손실이 커지는 문제가 여전히 남아 있었다. 이를 개선하기 위한 연구인 DDTB[3]는 추론 시 양자화 범위를 동적으로 업데이트하는 모듈을 제안하여, 입력 데이터 분포 변화에 보다 유연하게 대응할 수 있도록 했다.

또 다른 연구로 ODM[4]은 학습 과정에서 입력 데이터에 따라 크게 다른 활성화 값 분포를 정규화하는 손실 함수를 제안했으며, 이 손실 함수는 기존의 SR 모델을 학습할 때 사용하는 복원 손실 함수와 협력적으로 작동하도록 설계되었다. 이를 통해 활성화 값 분포를 균일하게 만들고, 양자화로 인한 손실을 줄이는 데 성공했다.

마지막으로 QuantSR[5]은 SR 모델의 특성에 맞춰 QAT 알고리즘을 개선한 사례로, 기존 QAT의 backward 과정에서 이전 층의 그라디언트를 그대로 전파하는 Straight-through Estimator (STE)의 한계를 지적하며, forward와 backward 과정에서 추가적인 변환 함수를 도입해 양자화로 인한 정보 손실을 줄일 수 있음을 밝혔다. 이러한 연구들은 QAT를 기반으로 SR 모델에서의 양자화 성능을 개선하기 위한 다양한 접근법을 보여준다.

Post-Training Quantization (PTQ)

PTQ는 학습이 완료된 FP32 모델에 양자화를 적용하여 경량화된 모델을 생성하는 방식이다. 모델 학습이 끝난 후에 양자화 작업이 수행되고 추가 학습이 아닌 소량의 캘리브레이션 데이터만 사용해 활성화 값의 동적 범위를 관측하여 양자화 함수의 스케일링 팩터와 제로 포인트를 정한다. PTQ는 기존 학습된 모델을 바로 양자화할 수 있으며 간단한 구현과 빠른 처리 속도를 제공하며, 소량의 캘리브레이션 데이터만 사용하기 때문에 추가적인 데이터 준비가 필요하지 않다는 장점이 있다. 그러나 학습과정이 없다보니 FP32 모델 대비 성능 저하가 클 수 있으며, 특히 SR과 같은 민감한 작업에서는 성능 감소가 더욱 두드러질 수 있다. 또한, 네트워크 구조에 따라 특정 양자화 기법의 효과가 크게 다를 수 있다는 단점도 있다.

따라서 SR 모델에서 PTQ를 적용한 사례는 현재까지는 많지 않다. 최근 연구 중 하나인 PTQ4SR[6]은 소량의 라벨이 없는 캘리브레이션 이미지만을 사용해 기존 학습된 SR 모델을 양자화하는 방법을 제안했다. 이 연구에서는 SR 모델의 활성화 값 분포를 히스토그램으로 변환한 후 각 구간의 밀도를 분석하여 최종 양자화 범위를 설정하여, 양자화 범위에 outlier가 포함되지 않도록 하는 기법을 사용했다. 해당 연구는 기존의 학습 기반의 양자화(QAT) 방법보다 성능은 낮지만, 추가 학습 없이도 기존 모델을 경량화할 수 있는 가능성을 보여주었다. PTQ는 간단한 모델 경량화와 빠른 추론 속도가 요구되는 실시간 시스템에서 특히 유용하기 때문에, 향후 후속 연구를 통해 활용 가능성이 더욱 높아질 것으로 기대된다.

2. 적용 범위에 따른 구분

모델에 같은 양자화 함수를 적용하는 범위에 따라 양자화 기법은 채널별 양자화, 레이어별 양자화, 그룹별 양자화, 텐서별 양자화 등으로 구분할 수 있다. 그러나 이미지 초해상화 (SR) 모델에 적용되는 주요 기법은 채널별 양자화 (Channel-wise Quantization)와 레이어별 양자화 (Layer-wise Quantization)로 나뉜다. 이 두 접근법은 각각의 장단점과 특성에 따라 SR 모델에 서로 다른 영향을 미친다.

채널별 양자화 (Channel-wise Quantization)

채널별 양자화는 모델의 각 채널에 대해 독립적인 스케일링 팩터와 제로 포인트를 계산하여 양자화를 수행하는 방식이다. DAQ[7] 연구에 따르면, SR 모델의 활성화 값 분포는 입력 데이터뿐 아니라 채널별로도 크게 다르다는 것이 밝혀졌다〈그림 3〉. 이러한 특성으로 인해, 채널별 양자화는 각 채널의 고유한 데이터 분포를 반영할 수 있어 SR 모델에서 정밀도 손실을 최소화할 수 있다. 특히 데이터 범위가 레이어 내 채널마다 크게 상이한 SR 모델에서는 채널별 양자화가 더욱 효과적인 결과를 보인다.

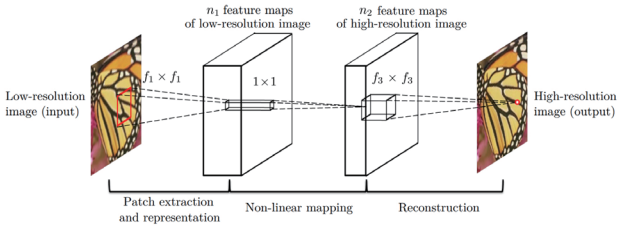
그러나 채널별 양자화는 각 채널에 대해 별도의 스케일링 팩터와 제로 포인트를 계산하고 저장해야 하므로 추가

적인 메모리와 계산 비용이 발생한다는 단점이 있다. 이를 해결하기 위해 DAQ는 양자화에 쓰이는 파라미터인 스케일링 팩터와 제로 포인트를 한 번 더 양자화하는 기법을 제안하여 연산 부담을 줄이고자 했다. 하지만 일부 하드웨어 가속기는 단일 스케일과 제로 포인트를 사용하는 방식을 가정하기 때문에, 채널별 양자화를 지원하지 않는 경우가 있다. 이러한 하드웨어 제한으로 인해 SR 모델에서 채널별 양자화의 활용이 제한되는 사례가 종종 발생한다. 결과적으로, SR 모델에서 채널별 양자화는 고정 밀도를 요구하는 작업에 적합하지만, 메모리와 계산 비용 증가, 하드웨어 호환성 문제로 인해 다양한 환경에서 활용되기는 어렵다고 볼 수 있다.

레이어별 양자화 (Layer-wise Quantization)

레이어별 양자화는 모델의 각 레이어 전체에 대해 단일 스케일링 팩터와 제로 포인트를 사용하는 방식이다. 이 방식은 스케일과 제로 포인트를 레이어 단위로만 계산하기 때문에 관리가 간단하고 계산 비용이 낮다. 채널별 양자화에 비해 계산이 단순하고 메모리 효율적이며, 대부분의 하드웨어에서 기본적으로 지원되므로 호환성이 높다는 장점이 있다.

레이어별 양자화는 데이터 분포가 채널 간에 유사한 경우, 특히 이미지 분류나 객체 탐지와 같은 비교적 간단한 작업에서 효과적으로 작동한다. 하지만, 채널별 데이터 범위가 크게 다른 경우에는 단일 스케일링 팩터가 데이터의 세밀한 분포를 반영하지 못해 정보 손실이 발생할 가능성이 크다. 이러한 정보 손실 문제는 SR과 같은 고성능



〈그림 4〉 CNN기반 SR 모델의 일반적인 구조

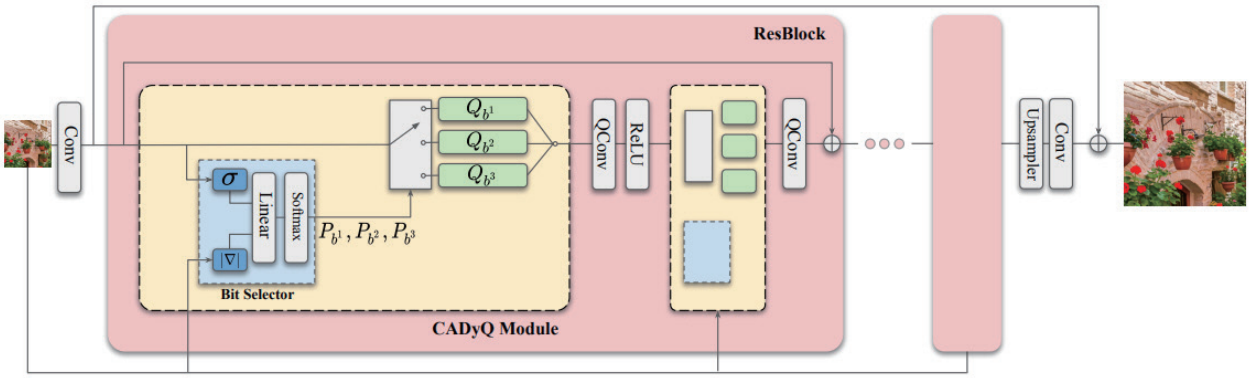
작업에서 특히 심각하게 나타난다.

SR 모델에서 레이어별 양자화를 적용할 때 이러한 한계를 극복하기 위한 다양한 기술이 도입되었다. 예를 들어, ODM[4]은 학습 과정에서 활성화 값의 채널 간 분산을 줄이기 위해 정규화 손실 함수를 추가로 도입하였다. 이를 통해 채널별 데이터 분포 차이를 완화하면서도 최종 초해상화 성능을 유지할 수 있었다.

최근 모바일 환경과 엣지 디바이스에서의 추론 수요가 증가함에 따라, SR 모델에서도 빠른 추론 속도와 메모리 효율성이 점점 더 중요해지고 있다. 이러한 요구를 충족하기 위해, 레이어별 양자화는 간단한 구현과 하드웨어 호환성 측면에서 여전히 강력한 선택지로 남아 있다. 따라서 레이어별 양자화와 관련된 연구는 앞으로도 지속적으로 이루어질 것으로 기대된다.

3. 정밀도 수준에 따른 구분

양자화에서 정밀도 (precision, bit-width) 수준을 어떻게 설정하는지에 따라 고정 정밀도(Fixed Precision), 혼합 정밀도(Mixed Precision), 적응형 정밀도(Adaptive Precision)로 나눌 수 있다. 각 방식은 정밀도와 효율성



〈그림 5〉 CADyQ의 적응형 정밀도 양자화 방식

의 균형을 유지하며 다양하게 활용되고 있다.

고정 정밀도 (Fixed Precision)

고정 정밀도는 모델의 모든 가중치와 활성화 값을 동일한 비트 폭(예: INT8, INT4)으로 양자화하는 방식이다. 모델의 모든 레이어에서 동일한 정밀도를 사용하며, 모든 연산이 동일한 정밀도로 이루어지기 때문에 계산이 단순하고 최적화가 용이하다. 특히 INT8과 같은 고정 정밀도는 NVIDIA Tensor Cores, ARM CPU 등 대부분의 하드웨어 가속기에서 잘 지원된다.

그러나 모든 레이어에서 동일한 비트 폭을 사용하기 때문에, 레이어별로 양자화에 민감한 정도가 다른 경우 정밀도가 부족해 정보 손실이 발생할 수 있다. 예를 들어, 네트워크의 특정 레이어가 더 높은 정밀도를 요구하는 경우에도 이를 반영하기 어려운 한계가 있다.

SR 모델 역시 더 높은 정밀도를 요구하는 레이어가 따로 있다. SR 모델은 주로 특징 추출(feature extraction), 비선형 매핑(non-linear mapping), 복원(reconstruction)의 세 단계로 구성된다 〈그림 4〉. 이 중 계산량은 여러 레이어가 쌓인 비선형 매핑 단계에 집중되어 있으며, 특징 추출과 복원 단계는 양자화에 상대적으로 민감하다. 따라서 대부분의 SR 양자화 기법에서는 특징 추출 모듈과 복원 모듈은 FP32로 유지하고, 비선형 매핑 모듈을 고정된 비트 폭으로 양자화하는 방식을 채택한다.

반면, FQSR[8]은 SR 모델의 모든 모듈(특징 추출, 비선형 매핑, 복원)을 양자화하는 방법을 제안한다. 이 기법은 가중치와 활성화 값뿐만 아니라, 많은 메모리를 차지하는 ResNet 구조의 Skip Connection까지 양자화한다. FQSR은 기존 SR 양자화 기법보다 성능이 다소 낮지만, 네트워크의 모든 레이어가 동일한 정밀도로 양자화되어 있기 때문에 하드웨어 최적화가 용이하다는 장점이 있다.

혼합 정밀도 (Mixed Precision)

혼합 정밀도는 네트워크 내에서 레이어 또는 모듈마다 다른 비트 폭을 사용하는 방식이다. 이는 정밀도에 민감한 부분에는 높은 비트 폭을, 덜 민감한 부분에는 낮은 비트 폭을 적용하여 정보 손실을 최소화하면서 계산 효율성을 극대화한다.

예를 들어, Super-Resolution Model Quantized in Multi-Precision[9] 연구에서는 Perception Index 값을 통해 양자화 민감도를 평가하고, 높은 민감도를 가진 모듈에는 높은 비트 폭을, 덜 민감한 모듈에는 낮은 비트 폭을 적용했다. 이를 통해 계산 비용을 줄이면서도 SR 모델의 성능을 유지할 수 있었다.

혼합 정밀도 방식은 메모리와 계산 자원을 효율적으로 분배할 수 있는 장점이 있지만, 레이어 또는 모듈별로 정밀도 민감도를 분석하고 적절한 비트 폭을 설정하는 과정이 필요하다. 또한, 특정 하드웨어는 고정 정밀도만 지원하므로 호환성 검증이 필수적이다.

적응형 정밀도 (Adaptive Precision)

적응형 정밀도는 네트워크 실행 중 입력 데이터나 레이어 특성에 따라 동적으로 비트 폭을 조정하는 방식이다. 이 방식은 입력 데이터의 통계적 특성(예: 값의 분포)에 따라 정밀도를 유연하게 조정할 수 있어, 다양한 데이터와 환경에 적응한다.

예를 들어, SR 모델의 입력 이미지가 구조적 디테일이 적은 하늘과 같은 이미지일 때는 더 낮은 비트 폭으로 양자화하고, 건물처럼 세부 정보가 많은 이미지일 때는 높은 정밀도를 유지하는 방식이 가능하다. 최근 연구에서 제안된 CADyQ[10]와 CABM[11]은 입력된 이미지에 따라 각 레이어의 비트 폭을 동적으로 할당하는 모듈을 제안하여 적응형 정밀도를 구현하였다 〈그림 5〉. 해당 모듈은 여러 비트 폭 후보에 대해 확률값을 예측하도록 학습되고, 추론 시에는 가장 높은 확률값을 갖는 비트 폭으로 양자화된 값이 선택되는 방식이다.

적응형 정밀도는 SR 문제처럼 다양한 데이터가 입력되며 중요도가 동적으로 변하는 작업에서 높은 성능과 효율성을 동시에 유지할 수 있는 강력한 접근법이다. 하지만, 동적으로 정밀도를 변경하기 위해 추가적인 제어 로직과 분석이 필요하며, 실행 중 정밀도 변경으로 인해 추가 계산 비용이 발생한다는 단점이 있다.

최근에 제안된 AdaBM[12]은 이러한 한계를 극복하기 위해 설계된 방법으로, 동적 비트 할당에 필요한 학습 과

정을 최소화하고, 추론 시 추가 계산 비용도 줄이는 알고리즘이다. 각 레이어마다 동적 비트 폭 할당 모듈을 사용하는 것이 아닌, 입력 이미지의 평균 그라이던트 변화량을 계산하여 이를 기반으로 네트워크 내 모든 연산의 비트 폭을 한 번에 업데이트하는 방식을 사용한다. 추론 전 기존 모델의 가중치 값을 이용해 할당된 각 레이어의 비트 폭이 한 번에 업데이트 되기 때문에 추론 비용, 학습 비용 모두 절감할 수 있다.

적응형 정밀도 방식의 다양한 연구들은 동적으로 정밀도를 변경할 수 있는 하드웨어와 결합하여 더욱 효율적인 추론을 가능하게 한다. 그러나 기존 하드웨어에서 적응형 정밀도가 완전히 지원되지 않을 수 있다는 점도 고려해야 한다. 따라서 SR 모델을 추론하고자 하는 하드웨어에 따라 정밀도 설정 방식을 신중히 선택해야 한다.

III. 전망과 결론

최근 몇 년간 이미지 초해상화(SR)분야에서 네트워크 효율성을 목표로 한 양자화 연구가 활발히 이루어졌다. 특히 SR 모델의 복잡한 구조와 높은 계산 비용으로 인해 8비트, 심지어는 초저비트 (2, 3비트) 양자화 기술이 주목받아 왔다. 이러한 연구는 계산 비용과 메모리 사용량을 크게 줄이는 데 기여했지만, SR 모델의 특수한 특성으로 인해 기존의 양자화 방법들이 SR 모델에 적용될 때 성능 저하가 발생하는 한계가 확인되었다. 이에 따라 최근에는 네트워크 층 및 입력 이미지 패치별 적응형 비트폭 할당, 입력 이미지 별 적응형 양자화 함수 등의 기술이 개발되며, SR 모델에 특화된 양자화 기술이 발전해 왔다.

최종 전망으로는 다음과 같은 방향에서 추가적인 연구가 기대된다. SR에도 최근 딥러닝의 트렌드인 대규모 모델, 즉 초대형 생성 모델 (generative models) 및 다모달 모델(multi-modal models)을 적용하는 최신 연구들이 이루어지고 있는데, 이러한 거대 모델에서도 양자화를 통해 효율성을 확보하는 연구가 필요하다. 거대 모델은 주로 초거대 데이터셋으로 학습되고 이 데이터셋이 공개되지 않는 경우도 많기에, 데이터 접근이 제한된 환경에서의 양자화 성능을 높이는 기법 개발이 요구된다. 또한, 다양한 하드웨어 및 애플리케이션의 요구에 따라 적응적으로 설계된 양자화 기술에 대한 연구도 필요할 것이다. 효율성과 성능의 균형을 맞춘 양자화 기술이 SR 뿐 아니라 다양한 이미지 복원 문제의 해결책으로 자리잡을 것으로 기대한다.

참 고 문 헌

[1] Bee Lim, et al. “Enhanced Deep Residual Networks for Single Image Super-Resolution”, CVPR Workshops 2017

[2] Huixia Li, et al. “PAMS: Quantized Super-Resolution via Parameterized Max Scale”, ECCV 2020

[3] Yunshan Zhong, et al. “Dynamic Dual Trainable Bounds for Ultra-Low Precision Super-Resolution Networks”, ECCV 2022

[4] Cheeun Hong, et al. “Overcoming Distribution Mismatch in Quantizing Image Super-Resolution Networks”, ECCV 2024

[5] Haotong Qin, et al. “QuantSR: Accurate Low-Bit Quantization for Efficient Image Super-Resolution”, NeurIPS 2023

[6] Zhijun Tu, et al. “Toward Accurate Post-Training Quantization for Image Super Resolution”, CVPR 2023

[7] Cheeun Hong, et al. “DAQ: Channel-Wise Distribution-Aware Quantization for Deep Image Super-Resolution Networks”, WACV 2022

[8] Hu Wang, et al. “Fully Quantized Image Super-Resolution Networks”, ACMMM 2021

[9] Jingyu Liu, et al. “Super-Resolution Model Quantized in Multi-Precision”, Electronics, 10(17):2176, 2021

[10] Cheeun Hong, et al. “CADyQ: Content-Aware Dynamic Quantization for Image Super-Resolution”, ECCV 2022

[11] Senmao Tian, et al. “CABM: Content-Aware Bit Mapping for Single Image Super-Resolution Network with Large Input”, CVPR 2023

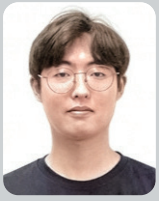
[12] Cheeun Hong, et al. “AdaBM: On-the-Fly Adaptive Bit Mapping for Image Super-Resolution”, CVPR 2024

홍채은

• 2020년 2월 서울대학교 전기정보공학부 학사
• 2020년 3월 ~ 현재 서울대학교 전기정보공학부 석박통합과정

〈관심 분야〉
Computer Vision, Image Restoration, Efficient AI, Network Quantization

구조적 관점에서의 영상 초해상화 인공지능 모델의 경량화 기법 분석


손명준
고려대학교


이홍재
고려대학교


정승원
고려대학교

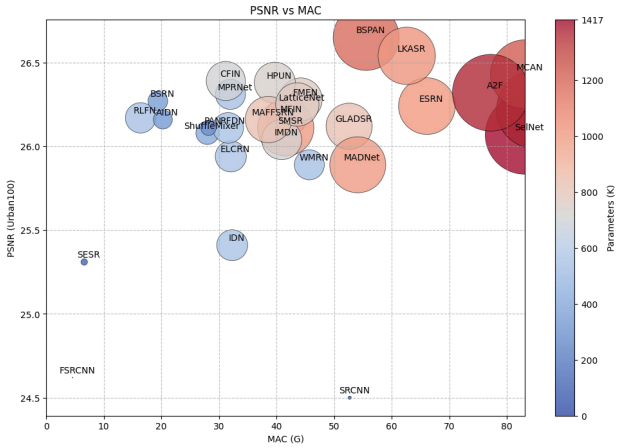
Ⅰ. 서 론

초해상화(Super Resolution, SR) 기술은 저해상도 (Low-Resolution, LR) 영상을 입력받아 대응하는 고해상도 (High Resolution, HR) 영상을 생성하는 기술이다. 특히 한 장의 영상만으로 초해상화를 실행하는 단일 영상 초해상화 (Single Image Super Resolution, SISR)의 연구가 주로 진행된다. 일반적으로 확대(upscale), 보간 (interpolation)과 같은 다양한 용어로 불리기도 하는데, 얼굴 인식, 의료 영상, 위성 영상 분석 등 여러 분야에서 중요한 역할을 한다.

최근 SISR 기술은 딥러닝 기반의 연구에 힘입어 크게 진보했으며, 높은 해상도의 결과물을 생성하는 데 있어 뛰어난 성능을 보여준다. 하지만, 이러한 성능 향상에는 네트워크 구조의 복잡도와 모델 크기의 증가가 뒤따르고, 계산량과 메모리 요구사항 역시 급격히 늘어났다. 〈그림 1〉은 공개된 SR 모델의 총 연산량과 Urban100[17] 데이터셋에 대한 성능을 나타낸 차트로, 우상단으로 가면서 모델의 성능과 연산량이 함께 커지는 경향을 확인할 수 있다. 이러한 이유로, 연산 자원이 한정적인 모바일 기기와 같은 환경에서 딥러닝 기반의 SR 모델을 실행하는 것이 매우 어렵다는 문제가 발생하였다. 이를 해결하기 위해 높은 품질의 복원을 유지하면서도 비교적 가벼운 모델을 개발하는 SR 모델의 경량화 기법에 대한 연구가 계속되었다.

모델 경량화에 대한 연구는 크게 네 가지의 범주로 분류할 수 있다. 가지치기 (Pruning), 양자화 (Quantization), 지식 증류 (Knowledge Distillation), 경량 구조 설계 및 탐색이 그것이다. 각 기법에 대한 개략적인 설명은 아래와 같다.

가지치기: 모델의 불필요한 연결이나 뉴런을 제거하여 네트워크를



〈그림 1〉 Super Resolution Task에서, 각 모델의 MAC에 대한 Urban100 [17] 데이터셋의 PSNR을 나타낸 공 차트. 공의 크기는 파라미터의 수를 의미함. 벤치마크 결과는 Gendy et al. [15]를 참고하였음.

축소하는 방식이다. 학습이 완료된 모델의 가중치의 중요도를 각각 분석하여 모델 성능에 대한 기여가 낮은 가중치를 제거함으로써 파라미터와 계산량을 줄이는 것을 목표로 한다.

양자화: 모델 가중치와 활성화 값의 표현을 32비트에서 8비트 또는 그 이하로 줄이는 방식이다. 이를 통해 메모리 사용량과 총 비트 연산량을 줄이고, 가속 연산을 지원하는 하드웨어에서 더 높은 연산 효율을 달성할 수 있다.

지식 증류: 고성능의 교사 (Teacher) 모델로부터 지식을 추출하여 경량화된 학생 (Student) 모델로 전달하는 방식이다. 일반적으로 교사 모델은 학생 모델보다 크고 복잡한 구조를 가지는데, 학생 모델은 교사 모델이 생성한 출력 분포를 학습하여 단순한 구조로 교사 모델의 동작과 성능을 따라가도록 학습한다.

경량 구조 설계 및 탐색: 처음부터 효율적인 네트워크 블록이나 모듈을 설계하거나, 효과적인 모델의 구조를 찾아가도록 학습시키는 등의 방법을 통해 모델의 구조 자체를 경량화하는 연구 또한 계속 이어져왔다. 특히 NTIRE (New Trends in Image Restoration and Enhancement)에서 꾸준히 진행되고 있는 Challenge on Efficient Super-Resolution과 같은 대회에서 이러한 목적을 가진 수많은 연구의 다양한 모델 구조와 연산 기법을 확인할 수 있다.

위의 네 가지 기법 중에서 가지치기와 양자화는 모델 자체에 변형을 주기 때문에 뒤따르는 문제가 발생한다. 예를 들어 양자화를 진행하는 경우, 높은 비트 표현 (32비트의 부동소수점)에서 낮은 비트 표현 (8비트의 정수 등)으로 모델의 파라미터를 변환하게 되면 양자화 과정에서 발생하는 오류로 인해 모델의 성능이 떨어진다. 이러

〈표 1〉 SR 모델의 경량화 연구의 파라미터 수, 실행 시간, 연산량과 BSD100 [16], Urban100 [17] 데이터셋에 대한 PSNR을 정리한 표. AdderSR [1]과 GhostSR [7]은 기준(baseline) 모델로 EDSR [18]을 사용하였음.

method	params(k)	runtime(ms)	MAC(G)	FLOPs(G)	BSD100(dB)	Urban100(dB)
CARN [8]	1592	39.96	65.02	130.04	27.58	26.07
IMDN [5]	715	20.56	23.295	46.59	27.56	26.04
RFDN [10]	550	20.4	13.55	27.1	27.57	26.11
RLFN [9]	543	16.41	9.85	19.7	27.6	26.17
ECBSR [12]	3.7	–	0.105	0.21	26.89	24.41
RepSR [4]	3.7	–	0.105	0.21	26.89	24.42
DIPNet [13]	2400	14.77	7.45	14.9	27.59	26.16
AdderSR [1]	–	–	5449	10898	27.58	26.33
SR-LUT [11]	–	15	–	–	26.33	23.68
GhostSR [7]	4309	–	1448	2896	27.73	26.6
AsConvSR [6]	2350	3.91	4.53	9.06	28.6	26.5

한 현상은 SISR과 같이 작은 픽셀 값의 변화에 민감한 작업에 특히 치명적이다. 또한 TPU와 같이 낮은 비트의 정수 연산을 지원하는 하드웨어가 없다면, 양자화–복호화 과정을 거쳐야 하는 특성으로 인해 오히려 기존 방식보다 느릴 수 있다.

가지치기의 경우, 제거되어야 하는 레이어, 블록 등이 모델에 골고루 위치하지 않아 경량화 이후의 모델 구조가 비균형적으로 변화할 수 있는데, 이는 하드웨어의 병렬 연산 능률을 저하시키는 원인이 된다. 특히, 희소성(Sparsity)을 활용하지 않는 기존 하드웨어에서는 성능 이득이 거의 없을 수 있다. 또한 잘못된 기준으로 레이어 혹은 가중치의 중요도를 분석하게 되면 모델의 성능에 상당한 악영향을 줄 수 있다. 게다가 가지치기 후 모델의 성능을 회복하기 위해 추가적으로 미세 조정을 해야 하는데, 이 과정에서 추가적인 시간과 연산 자원이 요구된다.

이러한 각각의 기법의 단점을 고려해보면, 경량 모델 구조의 설계를 통한 경량화 방식이 매력적으로 보인다. 때문에 수많은 연구들이 이러한 접근을 토대로 경량 SR 모델을 설계하고자 노력하였다. 본 고에서는 특히 주목할 만한 몇 가지 논문의 동기와 기술 세부 사항을 소개하고, 앞으로 후속 연구들이 다루어야 할 과제를 제시하고자 한다.

II. Preliminaries

SISR 모델의 경량화 연구들을 살펴보기 앞서, 해당 논문들을 이해하기 위해 필요한 몇 가지 요소를 짚어보고 넘어갈 필요가 있다. 우선, 연구 결과의 타당성을 수치의 형태로 보여줄 수 있는 여러 평가 지표에 대해 살펴본다. 주로 사용되는 지표에는 크게 결과 영상의 품질을 평가하는 PSNR (Peak Signal–to–Noise Ratio), SSIM (Structural Similarity Index)과, 모델의 총 연산량을 나타내는 FLOPs (Floating Point Operations)와 MAC (Multiply–Accumulate Operation)이 널리 사용되고, 추가로 모델의 총 파라미터 수 또한 모델 경량화의 주요한 지표로 사용되기도 한다. 각 지표의 정의와 특성은 아래와 같다.

PSNR: 저해상도 영상으로부터 재구성된 영상과 고해상도의 정답 영상 간의 픽셀 차이를 로그 스케일로 계산하는 지표로, 값이 높을수록 품질이 좋음을 의미한다.

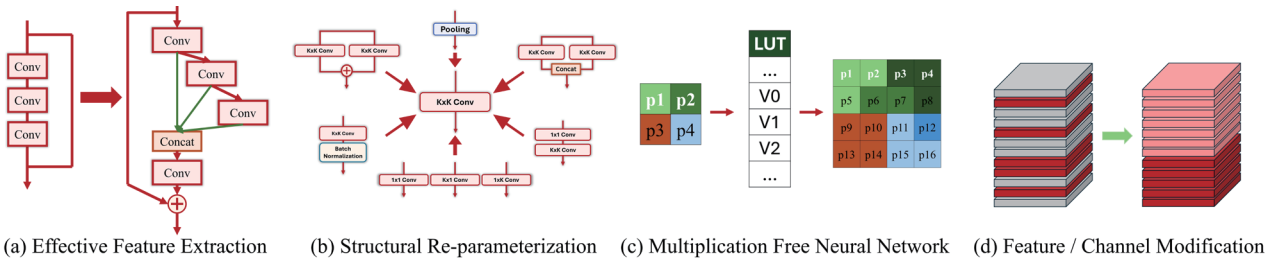
SSIM: 구조적 유사성, 명암, 대비를 기반으로 사람의 시각과 유사하게 영상 품질을 평가한다. PSNR보다 시각적 품질의 유사성을 더 잘 반영하는 특성이 있다. 역시 값이 높을수록 품질이 좋음을 의미한다.

FLOPs: 모델이 수행하는 부동소수점 연산의 총량을 나타낸다. 이 수치가 낮을수록 모델이 적은 연산을 수행함을 의미한다. 초당 수행 가능한 부동소수점 연산 횟수를 나타내는 FLOPS(Floating Point Operations per Second)와 다른 평가 지표임에 주의하여야 한다.

MAC: 곱셈과 덧셈의 총 횟수를 나타내는 지표이다. FLOPs는 곱셈과 덧셈 연산의 횟수를 각각 세어 2번의 연산으로 간주하지만, MAC은 곱셈과 덧셈을 묶어 1번의 연산으로 계산한다. 따라서, 일반적으로 FLOPs는 MAC의 약 2배에 해당하는 값으로 추정할 수 있다.

둘째로, 일반적으로 SR 모델에 사용되는 용어 또는 구조의 정의와 역할을 정리하면 다음과 같다.

합성곱: 합성곱(Convolution)은 입력 영상에서 특징을 추출하기 위한 핵심 연산으로, 필터(filter), 또는 커



(a) Effective Feature Extraction (b) Structural Re-parameterization (c) Multiplication Free Neural Network (d) Feature / Channel Modification

〈그림 2〉 구조적 관점에서 모델 경량화를 달성하기 위한 대표적인 네 가지 전략

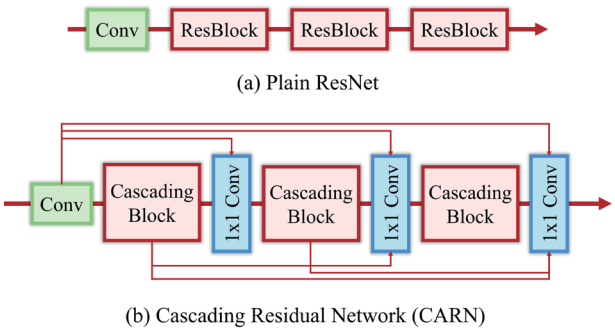
널(kernel)을 영상에 슬라이딩하며 연산하는 방식이다. 합성곱 층(convolution layer)은 경계선, 텍스처와 같은 영상의 국소적 특징을 효과적으로 학습하며, SISR에서 저해상도 영상의 고해상도 복원을 위한 필수적인 구성 요소다.

잔차 블록: 잔차 블록(Residual Block)은 네트워크의 각 계층이 입력 데이터를 그대로 전달하는 건너뛰기 연결(skip connection)을 포함하는 구조이다. 잔차 블록은 기울기 소실 문제를 완화하고, 학습 과정을 안정화하기 위해 사용한다.

순람표: 순람표(Lookup table, LUT)는 사전 계산된 결과를 테이블 형태로 저장하고, 이를 참조하여 연산을 수행하는 방식이다. 순람표를 사용하는 것으로 실시간 실행 환경에서 계산량을 줄이고 속도를 높이는 데 큰 이득을 취할 수 있다.

III. 구조적 관점에서의 영상 초해상화 인공지능 모델의 경량화 기법

구조적 수준에서 SR 모델의 경량화에 대한 연구는 기존 모델에 대한 깊이 있는 분석을 기반으로 다양한 접근 방법을 시도하였다. 그 중에서 몇 가지 주요한 아이디어를 정리해보자면, 다음과 같다. 첫째로 기존의 잔차 블록 구조에서 더 적은 파라미터를 이용해 더 의미 있는 특징을 추출할 수 있는 구조를 설계해 모델의 크기를 줄이면서 동시에 성능을 유지하는 것을 목표로 하는 연구들이 있다. 둘째로, 학습 과정에서는 복잡한 모델 구조를 이용해 다양한 특징을 학습하고, 추론 과정에서 복잡한 모델을 간단한 모델로 변형해 사용하는 기법이 존재한다. 셋째로, 순람표를 이용하거나, 곱셈 연산을 덧셈 연산으로 바꾸는 방법 등을 통해 모델의 연산 과정에서 곱셈 연산을 최대한 배제하는 방법이 제시되었다. 마지막으로, 특징 맵(Feature map) 또는 채널(Channel)을 변형하여 경량화 모델을 구성하고자 한 연구가 존재한다.



〈그림 3〉 Cascading Residual Network의 구조

1. 효과적인 잔차 블록 설계 방법에 관한 연구

본 절에서는 기존의 잔차 블록 구조를 변형해 보다 의미 있는 특징을 잘 잡아내면서 동시에 기존의 구조보다 더 적은 연산량, 더 빠른 실행 시간을 가지는 모델 구조에 대한 연구를 살펴본다.

1.1. 단계적 잔차를 이용한 네트워크

Ahn et al.[8]은 기존 SR 모델들의 한계를 실험적으로 분석하며 Cascading Residual Network (CARN)를 설계하였다. 연산량이 지나치게 많아 모바일 장치에서의 실행이나 실시간 실행에 적합하지 않은 기존 모델의 문제를 해결하기 위해 단계적 연결(cascading connection)과 Efficient Residual Block (ERB)을 도입하여 성능과 효율성을 동시에 개선할 방법을 고안했다.

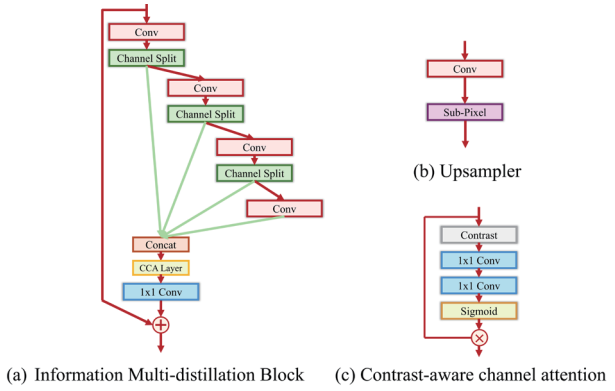
단계적 잔차 네트워크(Cascading Residual Network) 구조는, 다단계 잔차 연결을 통해 다양한 레벨의 특징을 효율적으로 결합하는 것을 목적으로 한다. 단계적 연결은 네트워크의 중간 층에서 생성된 특징을 상위 층으로 전달해 다양한 층에서 추출된 정보를 통합한다. 이를 통해 저레벨 특징(feature)과 고레벨 특징이 모두 반영되어 복잡한 세부 정보를 쉽게 복원할 수 있게 되었다. 이러한 설계는 고주파 정보를 보존하면서도 연산량 증가를 최소화해 초해상도 성능을 크게 향상시킨다.

추가적으로, 효율성을 극대화하기 위해 Ahn et al. 은 1x1 합성곱과 그룹 합성곱 (Group Convolution)을 활용하였다. 1x1 합성곱은 각 채널의 정보를 독립적으로 처리하며, 입력 채널의 수를 줄이거나 늘리기 위한 연산으로 사용된다. 이를 통해 채널 간 불필요한 연산을 줄이고 모델의 연산량을 최적화한다. 그룹 합성곱은 입력 채널을 여러 그룹으로 나누어 각 그룹이 독립적으로 연산을 수행

하도록 하여 연산량을 감소시킨다. 예를 들어, 입력 채널이 64개이고 이를 4개의 그룹으로 나누면, 각 그룹에서 수행하는 연산은 1/4로 줄어든다. 이러한 방식은 연산량을 줄이는 동시에 주요 특징을 효과적으로 학습할 수 있도록 돕는다.

네트워크의 업샘플링(upsampling) 방식 또한 CARN의 중요한 설계 요소 중 하나다. 기존 모델은 초기 단계에서 입력 영상을 고해상도로 변환한 후 처리하기 때문에 중간 연산량이 매우 크다. 반면 CARN은 대부분의 연산을 저해상도 상태에서 수행하고, 마지막 단계에서만 고해상도로 변환하는 방식을 채택했다. 이를 통해 처리 속도를 크게 높이면서도 고해상도 영상을 생성할 수 있다. 이러한 접근법은 효율성과 성능의 균형을 맞추는 데 크게 기여한다.

CARN은 다양한 요구를 충족하기 위해 두 가지 버전으로 제공되는데, CARN-M은 CARN의 설계를 기반으로 하여, 모바일 환경과 같은 자원 제한적 환경에서도 높은 성능을 유지할 수 있도록 CARN을 더욱 경량화한 버전이다. CARN-M은 Efficient Residual Block (ERB)을 보다 효율적으로 재구성하여 연산량과 파라미터 수를 줄였다. 특히, ERB 내에서 1x1 합성곱과 그룹 합성곱의 사용을 최적화하여 각 연산 단계의 부담을 줄였다. 기존 CARN에서는 더 많은 채널을 처리하기 위해 추가적인 연산이 필요했던 반면, CARN-M은 그룹 합성곱에서 그룹 크기를 세밀하게 조정하고, 레이어의 깊이를 조절하여 연산량을 최소화하면서도 특징 학습 능력을 유지하도록 설계되었다.



〈그림 4〉 IMDN [5]에서 제시한 Information Multi-distillation Block과 Upsampler의 구조

CARN-M의 또 다른 주요 특징은 네트워크 구조의 단순화와 재귀적 설계를 결합한 점이다. 네트워크의 반복적인 패턴을 활용해 같은 모듈을 재사용함으로써, 파라미터 수를 줄이고 메모리 사용량을 낮췄다. 또한, 업샘플링 과정에서도 최적화를 통해 고해상도 영상을 생성하는 데 필요한 자원을 최소화했다. 이러한 설계는 특히 배터리 소모를 줄이고 실시간 처리를 가능하게 하여 모바일 장치에서 활용성을 크게 높였다. CARN-M은 실험적으로 기존 대규모 모델과 유사한 품질의 영상을 생성하면서도 처리 속도를 크게 향상시켜, 실제 응용에서 효과적임을 입증하였다. CARN의 영상 처리 성능과 연산량은 〈표 1〉에서 확인할 수 있다.

1.2. 정보의 다중 증류를 이용한 네트워크

정보 다중 증류 네트워크(Information Multi-distillation Network, IMDN)[5]은 정보 다중 증류 블록(Information Multi-distillation Block, IMDB)을 이용해 효율적인 연산을 수행하고자 하였다. IMDB로 이루어진 IMDN은 입력 특징을 다단계로 나누어 점진적으로 정제하며 필요한 정보를 선별적으로 학습하는 방식을 채택하였다. 이를 통해 네트워크는 필요한 정보만을 집중적으로 학습하여 연산 효율을 극대화할 수 있었다.

IMDB 내부에서는 채널 분리 기법이 활용되는데, 이 기법은 입력 채널을 두 부분으로 나누어 일부는 바로 출력으로 연결하고, 나머지는 다음 단계에서 추가적으로 처리되도록 설계되었다. 이렇게 함으로써 각 단계에서의 연산량을 줄이고, 필요한 정보를 보다 정교하게 추출할 수 있다. 마지막으로 정제된 특징들은 통합되어 최종 출력으로 연결된다.

또한 IMDN은 대조 기반 채널 어텐션(Contrast-Aware Channel Attention, CCA) 기법을 도입하여 네트워크 성능을 한층 더 향상시켰다. 이 기법은 각 특징의 상대적 중요도를 학습하여 중요한 특징에 더 큰 가중치를 부여하고 덜 중요한 특징은 억제하도록 설계되었다. 이를 통해 SR 작업에서 필요한 세부 정보를 효과적으로 복원할 수 있었다.

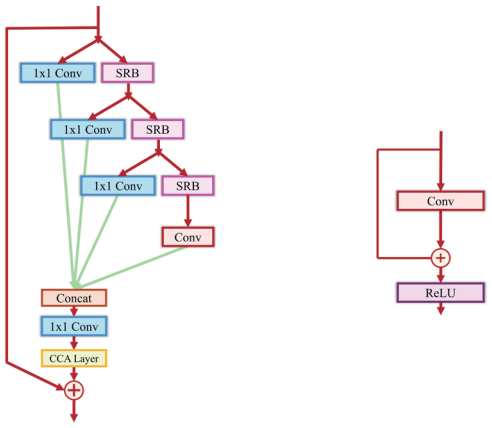
그리고 IMDN은 기존의 많은 SR 모델들이 사용하던 업샘플링 방식을 개선하였다. 기존 모델들은 초기 단계에서 영상을 고해상도로 변환한 뒤 연산을 수행하였지만, IMDN은 대부분의 연산을 저해상도 상태에서 수행한 뒤 마지막 단계에서만 고해상도로 변환하였다. 이러한 접근법은 연산량을 크게 줄이고, 고품질의 출력을 유지하는 데 기여하였다.

IMDN은 재귀적 설계나 과도한 네트워크 깊이에 의존하지 않는다. 기존 경량화 모델들은 재귀적 구조를 사용하여 파라미터 수를 줄였으나, 이는 모델의 표현력을 저하시킬 위험이 있었다. 반면 IMDN은 독립적인 모듈로 구성된 설계를 채택하여 네트워크의 복잡성을 줄이는 동시에 높은 수준의 성능을 유지하였다.

마지막으로 IMDN은 적응형 크롭핑 전략을 도입하여 입력 영상을 블록 단위로 처리하는 방식을 채택하였다. 이를 통해 네트워크의 메모리 사용량을 줄이고 처리 속도를 높일 수 있었다. 이 전략은 다양한 입력 크기에 대해 유연하게 작동할 수 있도록 하여 실시간 응용에서의 활용성을 높였다. 결과적으로 IMDN은 PSNR과 SSIM 같은 품질 지표에서 기존 대규모 모델과 유사한 성능을 보이면서도 연산량과 메모리 사용량을 크게 줄이는 데 성공하였다. 벤치마크 결과는 〈표 1〉에서 확인할 수 있다.

1.3. 잔차 특징의 증류를 이용한 네트워크

잔차 특징 증류 네트워크(Residual Feature Distillation Network, RFDN)[10]은 IMDN을 변형 및 개선하여 개발된 모델이다. IMDN은 IMDB를 기반으로 경량화와 성능 간의 균형을 목표로 했지만, RFDN은 이를 더욱 간소화하고 최적화하여 계산 효율성과 복원 품질을 동시에 개선하고자 설계되었다. RFDN은 연산량을 줄이기 위해 IMDN에서 사용된 대조 기반 채널 어텐션을 생략하고, 더 단순하고 직관적인 특징 정제 접근 방식을 채택하였다. 또한, 특징 추출과 융합 단계에서의 구조를 재설계하여 초해상화 품질을 유지하면서도 경량화된 모델로 작동할 수 있도록 했다. 이러한 개선은 모바일 기기와 같은 제한된 환경에서도 실시간 성능을 제공할 수 있도록 목표를 맞췄다.



(a) Residual feature distillation block (b) Shallow residual block

〈그림 5〉 RFDN [10]에서 제시한 Residual feature distillation block과 Shallow residual block의 구조

특징 증류 연결(Feature Distillation Connection , FDC)은 기존 IMDN에서 정보 증류 메커니즘을 대체하며 더 경량화되고 유연한 구조를 제공한다. FDC는 입력과 출력 사이에서 특징 정보를 더 효과적으로 전달할 수 있는 구조적 개선 사항을 포함한다. 특히, FDC는 1x1 및 3x3 합성곱 계층을 통해 특징을 추출하고 정제하여 매개변수 수를 줄이면서도 특징 표현력을 유지한다. 또한, 항등 연결(identity connection)을 사용하여 네트워크 학습 중 정보 손실을 줄이고 효율성을 높이는 역할을 한다.

항등 연결은 입력과 출력 사이에 직접적인 연결을 추가함으로써 학습 중 정보가 더 잘 전달되도록 돕는 방식이다. 이를 통해 정보 손실을 최소화하고 네트워크의 학습 안정성을 강화한다. 이러한 연결은 특히 경량화 모델에서 중요한 역할을 하며, 성능 향상을 지원하는 핵심적인 요소로 작용한다.

얕은 잔차 블록(Shallow Residual Block, SRB)은 3x3 합성곱, 활성화 함수, 그리고 항등 연결로 구성된다. SRB는 기존의 IMDN 구조에서 채널 분할과 결합 과정을 단순화하면서도, 깊은 잔여 연결의 장점을 경량 네트워크에서도 활용할 수 있게 한다. 이러한 설계를 통해 SRB는 기존의 단순 합성곱 계층보다 더 세밀한 학습을 가능하게 하며, 추가적인 매개변수를 도입하지 않고도 성능을 향상시킨다. 이를 통해 RFDN은 더 많은 잔여 학습 경로를 활용하여 특징 학습의 효율성을 극대화할 수 있다.

RFDN은 초기 특징 추출 계층, 다수의 잔차 특징 증류 블록(Residual feature distillation block, RFDB), 특징 융합 계층, 그리고 최종 재구성 계층으로 구성된다. 초기 계층은 입력 영상을 기반으로 기본 특징을 추출하며, RFDB는 이를 점진적으로 정제한다. IMDN이 CCA를 활용하여 특징을 융합한 것과 달리, RFDN은 단순한 특징 융합 방식을 통해 연산량을 줄였다. 최종적으로 모든 특징을 결합하고 고해상도 영상을 생성한다. 이러한 구조는 계산량을 줄이면서도 초해상화 품질을 극대화한다.

RFDN은 다양한 벤치마크 데이터셋에서 기존 모델을 능가하는 성능을 보였다. 〈표 1〉에서 RFDN과 IMDN을 비교했을 때, 여러 데이터셋에서 PSNR이 더 나은 결과를 달성하면서도 매개변수 수와 연산량이 더 적었다. 이는 RFDN이 IMDN의 구조적 장점을 유지하면서도, 설계를 간소화하여 경량화와 성능의 균형을 더욱 잘 유지한다는 것을 보여준다.

1.4. 잔차 지역 특징에 집중한 네트워크

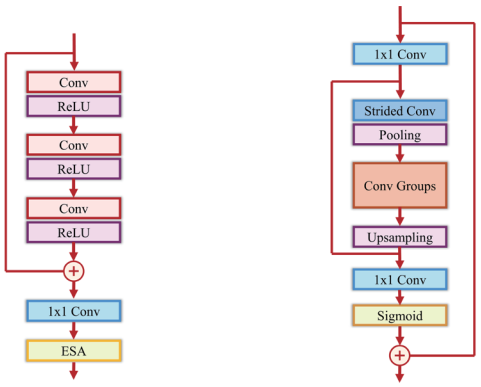
Kong et al.[9]은 RFDN에서 도입된 잔여 특징 증류 블록(RFDB)을 변형하여 효율성을 더욱 향상시킨 잔차 지역 특징 네트워크(Residual Local Feature Network, RLFN)을 설계하였다. 이 모델은 복잡한 특징 연결 방식을 단순화하면서도 연산 속도와 품질 간의 균형을 유지하기 위해 설계되었다. 기존 RFDB에서 채널 분리 및 결합 방식으로 인해 발생했던 하드웨어 비호환성을 해결하기 위해, RLFN은 잔차 지역 특징 블록(Residual Local Feature Block ,RLFB)을 도입하여 연산량은 줄이고 성능은 유지하는 데 중점을 두었다.

RLFB는 RFDB에서 특징 증류 연결을 제거하고, 대신 간단히 합성곱에 잇달아 오는 ReLU[19] 구조를 채택하여 지역적 특징을 학습하도록 설계되었다. 이 방식은 여러 단계의 지역적 특징의 정제를 통해 입력 특징을 보존하며 동시에 성능 저하 없이 처리 속도를 높였다. 또한, RFDB에서 사용된 향상된 공간 어텐션(Enhanced Spatial Attention, ESA)을 최적화하여 불필요한 합성곱 계층을 줄임으로써 추가적인 속도 향상을 달성하였다. 이러한 구조 변경은 RLFN이 IMDN과 RFDN과 같은 이전 모델 대비 더욱 효율적인 실행 속도를 제공하는 데 기여하였다.

학습 과정에서 사용한 손실 함수(Loss function)에도 변화를 주었다. 기존 SR 연구에서 자주 사용되는 L1 손실 외에도 개선된 대조 손실(contrastive loss)을 도입하였다. 대조 손실은 얕은 계층에서 추출된 특징이 PSNR을 높이려고 노력하는 모델의 세부 묘사를 개선하는 데 중요하다는 점에 착안하여 설계되었다. 이를 위해 VGG-19[20] 기반의 기존 특징 추출기를 재설계하여, 더 많은 세부 정보와 텍스처를 포착할 수 있는 구조를 채택하였다. 이 덕분에 학습 과정에서 더 의미있는 loss를 활용하여 성능을 향상시켰다.

또한, RLFN은 다단계 웜스타트(multi-stage warm-start) 학습 전략을 제안하였다. 초기 단계에서는 모델을 처음부터 학습한 후, 이후 단계에서는 이전 단계에서 학습된 가중치를 활용하여 학습을 이어가는 방식이다. 이 접근법은 학습 수렴 속도를 높이고 복원된 영상의 품질을 높이는 데 기여하였다. 특히, 높은 배율 확대(x4)와 같은 복잡한 초해상화 작업에서도 효과적으로 작동함을 실험적으로 입증하였다.

NTIRE 2022 Challenge on efficient super-resolution[21]에서 RLFN은 뛰어난 성능으로 1위를 차지하며 상용화 가능성을 입증하였다. 챌린지에서 사용된 RLFN-cut 모델은 RLFB의 개수를 줄이고 채널 수를 최적화하여 더욱 간결한 구조를 제공하였다. 이를 통해 높은 품질의 영상을 빠른 속도로 복원할 수 있는 가능성



(a) Residual local feature block (b) Enhanced Spatial Attention

〈그림 6〉 RLFN [9]에서 제시한 Residual local feature block과 Enhanced Spatial Attention의 구조

을 제시하였다. 벤치마크 결과는 〈표 1〉에서 확인할 수 있다.

2. 구조적 재매개변수화를 이용한 모델 경량화에 관한 연구

본 절에서는, 구조적 재매개변수화(Structural Re-parameterization)을 이용해 추론 연산시의 모델의 연산량을 줄인 연구를 소개하고자 한다. 구조적 재매개변수화는 Ding et al.[2]에서 처음 명명된 기법으로, 학습 과정에서는 복잡한 모델 구조를 이용해 모델이 다양한 패턴을 학습하도록 유도하고, 추론 과정에서는 단순한 모델 구조를 이용해 연산량에서 이득을 취하는 기법이다. Ding et al.은 [2]에서 몇 가지 주요 합성곱 연산이 실질적 동일성을 유지한 채로 보다 단순한 합성곱 연산으로 변경할 수 있음을 증명하였다. 이를 이용해 학습 과정에서 사용한 복잡한 구조의 모델을 수학적으로 동일하되 연산량 자체는 크게 감소한 모델로 변형해 추론 단계에서 사용해 실질적으로 실행 시간과 연산량에서 이득을 취할 수 있다.

2.1. 테두리 정보에 집중한 합성곱 블록을 이용해 실시간 초해상화를 달성한 연구

테두리 기반 합성곱 블록을 이용한 초해상화 네트워크(Edge-oriented Convolution Block for Super-Resolution, ECBSR)는 모바일 기기에서 실시간 초해상화를 달성하기 위해 설계된 경량화 네트워크이다. ECBSR은 초해상화 작업에서 성능과 효율성을 동시에 고려한 새로운 접근 방식을 도입하였다. 특히, 테두리 기반 합성곱 블록(Edge-oriented Convolution Block, ECB)은 재파라미터화 기법을 활용하여 다중 경로에서 학습된 복잡한 연산을 단일 3×3 합성곱으로 병합한다. 이를 통해 성능 저하 없이 추론 속도를 크게 향상시켰다. 이 모델은 모바일 장치의 제한된 하드웨어 자원에서도 높은 초해상화 품질을 유지하면서도 실시간 처리가 가능하도록 설계되었다. ECBSR의 또 다른 주요 목표는 모델의 간결성을 유지하면서도 다양한 기기와 환경에서 최적의 성능을 발휘할 수 있는 유연성을 제공하는 것이다.

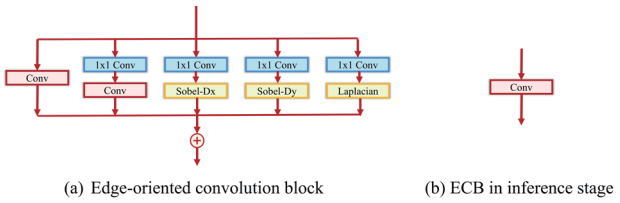
ECB는 ECBSR의 핵심 구성 요소로, 네 가지 주요 연산 경로를 포함한다. 첫 번째는 기본 3×3 합성곱으로, 이는 영상의 저해상도 입력에서 기본적인 특징을 추출한다. 두 번째는 채널 확장 및 축소(conv-expand-and-squeeze) 연산으로, 이는 입력 채널을 확장하여 모델의 표현력을 증가시키고 다시 축소하여 계산 효율성을 유지한다. 세 번째로는 Sobel 필터를 활용한 1차 미분 연산을 통해 수평 및 수직 방향의 엣지 정보를 효과적으로 추출하며, 마지막으로 라플라시안(Laplacian) 필터를 활용한 2차 미분 연산은 더 정밀한 테두리 정보를 제공하고 노이즈에 강한 특성을 나타낸다. 이러한 네 가지 경로는 학습 과정에서 독립적으로 최적화되며, 재파라미터화 과정을 통해 효율적으로 결합되어 단일 3×3 합성곱으로 병합된다.

기본 3×3 합성곱은 SR 작업에서 핵심적인 특징을 추출하며, 다른 연산 경로의 기초 역할을 한다. 이 연산은 일반적인 영상을 처리하기 위한 기본 성능을 제공하며, 초기 입력 데이터를 기반으로 신호를 정제하고 전달하는데 중요한 역할을 한다. 단순하지만 안정적인 구조로 인해, 이 경로는 모든 네트워크에서 공통적으로 사용된다. ECBSR에서는 이 경로를 통해 전체 모델의 신뢰성과 일관성을 유지하도록 설계되었다.

또한 채널 확장 및 축소 연산을 이용해 학습의 효과를 높이고자 하였다. 이 연산은 입력 데이터의 채널 수를 확장하고 다시 축소함으로써 모델의 표현력을 극대화하는 연산으로, 채널 확장 과정은 더 넓은 특징 공간을 탐색할 수 있도록 하며, 축소 과정은 효율성을 유지하면서도 불필요한 계산을 줄이는 역할을 한다. 이를 통해 모델이 복잡한 패턴과 텍스처를 효과적으로 학습할 수 있도록 하며, 추가적인 계산 부담 없이 데이터의 풍부한 정보를 보존하였다.

소벨(Sobel) 필터를 활용한 1차 미분 연산은 테두리 정보를 효과적으로 추출하기 위한 중요한 도구이다. 수평 및 수직 방향을 강조하여 영상의 구조적 세부 사항을 잘 포착하며, 특히 텍스처가 풍부한 영역에서 SR 성능을 크게 향상시킨다.

라플라시안 필터를 활용한 2차 미분 연산은 소벨 필터에서 한 단계 더 나아가, 더 정밀하고 노이즈에 강한 엣지



〈그림 7〉 ECBRN [12]에서 사용한 Edge-oriented convolution block과 추론 과정에서 ECB의 변환 형태

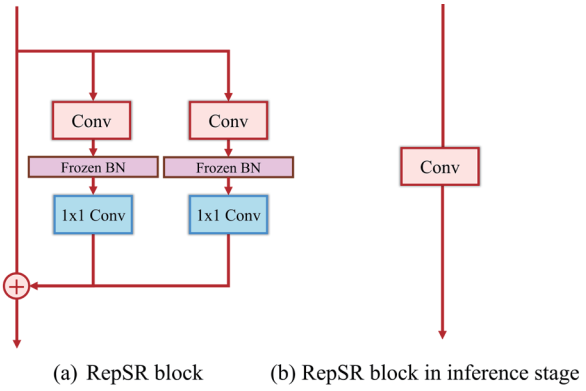
정보를 제공한다. 이 연산은 영상의 미세한 디테일을 복원하는 데 중요한 역할을 하며, 특히 부드러운 영역에서 텍스처 정보를 보존하면서도 경계 부분의 선명도를 높인다. 이 과정은 SR 작업에서 전반적인 시각적 품질을 개선하는 데 기여한다.

ECBSR은 모델의 이름에 포함된 "M4C8" 또는 "M4C16"과 같은 표기법으로 모델 구성의 복잡도를 나타낸다. 여기서 "M"은 모델 내부에서 사용된 ECB의 개수를, "C"는 각 ECB에서 사용된 채널의 수를 의미한다. 예를 들어, "M4C8"은 4개의 ECB를 포함하며, 각 ECB는 8개의 채널을 활용한다는 것을 나타낸다. 이러한 표기법은 모델의 크기와 연산량을 직관적으로 표현하며, 사용자가 다양한 환경에 맞게 모델을 조정할 수 있도록 돕는다. 따라서 "M4C8"과 같은 경량 모델은 실시간 초해상화가 요구되는 모바일 디바이스에서 효율적으로 작동할 수 있는 적합한 선택지이다.

ECBSR은 제한된 하드웨어 환경에서도 초해상화 작업을 실시간으로 처리할 수 있는 성능을 입증하였다. ECBSR-M4C8 및 ECBSR-M4C16과 같은 경량 모델은 스냅드래곤 865 및 디멘시티 1000+ 환경에서 높은 처리 속도를 기록하며, 540p에서 1080p로의 업스케일링 작업을 안정적으로 수행하였다. 벤치마크 데이터셋(Set5, Urban100 등)에서 PSNR 및 SSIM 측면에서 우수한 성능을 보여주었으며, 다양한 복잡도 수준에서 모델의 확장 가능성을 입증하였다. 다른 모델과의 성능 비교는 〈표 1〉에서 확인할 수 있다.

2.2. 배치 정규화의 변형과 구조적 재매개변수화를 이용한 VGG 형태의 초해상화 네트워크

RepSR[4]은 구조적 재매개변수화와 배치 정규화(Batch



〈그림 8〉 RepSR [4]에서 사용한 (a) RepSR block과 (b) 추론 과정에서의 변환 형태

Normalization, BN)[22]를 활용하여 효율적인 VGG 스타일[20] 의 SR 네트워크를 설계한 모델이다. 이 모델은 훈련 단계에서는 다중 경로 구조를 활용하고, 추론 단계에서는 단일 3×3 합성곱으로 병합하여 성능과 효율성을 극대화한다. 특히 배치 정규화를 재도입하여 훈련 과정에서 비선형성을 제공함으로써 기존 BN 문제를 해결하고 최종 성능을 향상시켰다. 이는 ECBSR과 같은 기존 구조적 재매개변수화를 적용한 모델에 비해 더 간결하면서도 효과적인 설계를 특징으로 한다.

VGG 스타일 네트워크는 3×3 합성곱 계층과 활성화 함수로 구성된 단순한 구조를 기반으로 한다. 이러한 구조는 딥러닝 모델의 병렬 처리와 하드웨어 가속을 최대화할 수 있어 모바일 환경에서 특히 적합하다. 단순성과 유연성을 갖춘 VGG 스타일 네트워크는 초해상화 작업에서 널리 활용되고 있다.

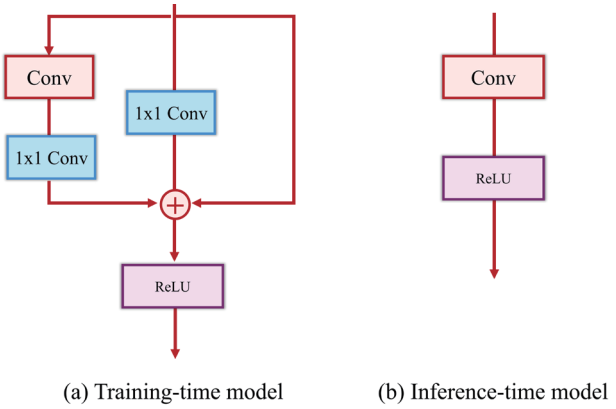
배치 정규화는 신경망의 학습 안정성을 높이고 수렴 속도를 향상시키는 기법이다. BN은 각 계층에서 입력 데이터를 정규화하여 내부 공변량 이동을 줄이고, 이를 통해 훈련 과정에서 더 나은 일반화와 최적화를 가능하게 한다. 그러나 SR 작업에서는 배치 정규화가 성능 저하와 아티팩트(artifact)를 유발할 수 있어 종종 모델 구조에서 탈락되었다.

고정된 배치 정규화(Frozen BN)는 기존 배치 정규화에서 발생하는 아티팩트를 해결하기 위해 설계된 전략이다. 일반적인 배치 정규화는 훈련 중에는 미니배치 통계를 사용하고, 추론 중에는 전체 데이터 통계를 사용하는데, 이

과정에서 발생하는 불일치가 SR 작업에서 아티팩트를 유발한다. 고정된 배치 정규화는 훈련 초기에는 미니배치 통계를 사용하다가 훈련 후반부에는 전체 데이터 통계로 전환하여 이러한 문제를 해결한다. 이를 통해 배치 정규화의 비선형성을 유지하면서도 추론 시 발생할 수 있는 통계 불일치를 제거하였다. RepSR은 이 전략을 도입하여 기존 SR 네트워크보다 더 안정적이고 효율적인 성능을 보여주었다.

RepSR 블록은 두 개의 채널 확장-축소(expand-and-squeeze) 경로와 하나의 잔여 경로(clean residual path)로 구성된다. 잔여 경로는 BN 없이 설계되어 특징 정보를 효율적으로 전달하며, 다른 두 경로는 3×3 및 1×1 합성곱과 배치 정규화를 활용해 채널 확장을 수행한다. 이러한 설계는 넓은 특징 표현력과 빠른 수렴 속도를 동시에 제공한다. 고정된 배치 정규화 전략은 훈련 후반에 미니배치 통계에서 전체 데이터 통계로 전환하여 BN에서 발생하는 아티팩트를 방지하며, ECBSR이 소벨 및 라플라시안 필터를 사용하는 복잡한 설계를 대체하는 데 효과적이었다.

RepSR은 다양한 벤치마크 데이터셋에서 ECBSR, IMDN, RFDN 등과 비교하여 뛰어난 성능을 입증하였다. 특히 Urban100 및 Set14와 같은 데이터셋에서 PSNR 및 SSIM 결과가 기존 모델을 능가하였으며, 경량 모델부터 대규모 모델까지 확장 가능한 설계를 보여주었다. 이는 RepSR이 실제 모바일 환경에서 효율적인 초해상화 네트워크의 새로운 표준을 제시함을 의미한다.



〈그림 9〉 DIPNet [13]에서 사용한 (a) 학습 과정에서의 잔차 블록과 (b) 추론 과정에서의 변환 형태

결론적으로, RepSR은 재파라미터화 기법과 배치 정규화를 결합하여 초해상화 네트워크 설계에서 중요한 진전을 이루었다. 이 모델은 간결성과 성능의 균형을 통해 ECBSR 대비 효율성과 품질을 동시에 제공하며, 초해상화 작업의 새로운 기준을 수립하였다.

2.3. 효율적인 학습 전략과 구조적 재매개변수화를 이용한 초해상화 네트워크

DIPNet[13]은 성능과 속도의 균형을 맞추기 위해서는 경량화된 구조와 더불어 효율적인 학습 전략 또한 중요하다는 아이디어로부터 구체화되었다. 이 모델은 지식 증류, 점진적 학습, 반복적 가지치기와 같은 다단계 학습 전략을 통해 경량 네트워크가 고품질의 초해상화 영상을 생성할 수 있도록 설계되었다. DIPNet은 NTIRE 2023 Challenge on Efficient Super-Resolution 에서 가장 빠른 추론 시간을 기록하며 성능을 입증하였다.

Yu et al. 은 재매개변수화 잔차 특징 블록(Reparameterization Residual Feature Block, RRFB)을 도입하여 모델의 효율성을 높이려 노력하였다. 이 블록은 RFDN과 RLFN에서 사용된 RLFB를 기반으로 확장된 구조로, 훈련 중에는 복잡한 구조로 최적화된 성능을 제공하지만, 추론 시에는 단순한 3×3 합성곱으로 병합되어 계산 효율성을 유지한다. 이는 ECBSR의 소벨 및 라플라시안 필터 기반 설계와 비교하여 간결한 구조로 더 나은 성능을 제공한다.

DIPNet의 또 다른 핵심 요소는 모델 기반 정답 향상 전략(model guided ground-truth enhancement)이다. 기존 SR 모델이 원본 HR 영상을 학습하는 것과 달리, DIPNet은 향상된 HR 영상을 사용하여 더 풍부한 고주파 정보를 학습하도록 설계되었다. 이를 통해 제한된 학습 능력을 가진 경량 모델도 더 나은 텍스처와 세부 정보를 복원할 수 있다. 이러한 접근은 RepSR에서 활용된 고정된 배치 정규화와 달리, GT 데이터를 직접 향상시켜 학습 데이터를 강화하는 방식으로 차별화된다.

다중 앵커 지점 증류(Multi-anchor Feature Distillation)는 DIPNet의 또 다른 특징이다. 이는 다양한 깊이의 네트워크에서 추출된 특징을 이용하여 경량 학

생 네트워크의 학습을 지도하는 방식으로, 각 블록에서 추출된 특징을 사용하여 학습 성능을 극대화한다. 이 방식은 ECBSR 및 RepSR의 단순 증류 전략과 비교해 더 세부적인 학습 지도를 제공한다.

또한 DIPNet은 반복적 가지치기(Iterative Pruning)를 통해 모델의 매개변수 수를 줄이고 경량화를 극대화하였다. 이 방법은 L2 필터 가지치기를 사용하여 모델의 비효율적인 파라미터를 제거하며, 각 단계 이후에는 모델을 다시 미세 조정하여 성능을 유지한다. 이는 단일 단계 가지치기보다 PSNR 성능을 더 잘 유지할 수 있음을 실험적으로 입증하였다.

벤치마크 실험에서 DIPNet은 ECBSR, RFDN, RLFN 등 기존 모델과 비교하여 뛰어난 성능을 보여주었다. 특히, Urban100 및 Set5와 같은 데이터셋에서 유사한 PSNR 및 SSIM 성능을 기록하면서도 추론 속도와 모델 크기 면에서 월등한 효율성을 제공하였다. 이는 DIPNet이 실제 모바일 및 엣지 디바이스 환경에서 활용하기 적합한 SR 네트워크임을 의미한다.

3. 곱셈 연산이 없는 모델 구조에 관한 연구

곱셈 연산이 없는 모델에 대한 연구는 곱셈 연산을 제거, 혹은 대체하여 효율성과 에너지 소비를 개선한 신경망 구조를 찾는 방향으로 진행되었다. 이 접근법은 곱셈 대신 덧셈, 시프트, 논리 연산을 활용하거나 Look-up table(LUT)을 이용해 계산 비용을 크게 줄이는 것을 목적으로 한다. 이 절에서는 덧셈 연산만을 이용한 SR 네트워크와 LUT를 이용한 SR 모델의 연구에 대해 설명한다.

3.1. 덧셈 연산만을 이용한 초해상화 네트워크

AdderNet[2]은 곱셈 연산이 많은 에너지와 계산 비용을 소모한다는 점에서 출발하였다. 합성곱 기반 신경망(Convolutional Neural network, CNN)에서 사용되는 합성곱 연산은 입력 특징과 필터 간의 유사도를 계산하는 데 대규모의 곱셈 연산을 필요로 한다. AdderNet은 이를 개선하기 위해 L1 노름 기반의 유사도 계산 방식을 제안하였다. 이 접근법은 절대 차이를 통해 유사도를 측정하여 곱셈을 덧셈으로 대체함으로써 계산 복잡도를 줄이고 에너지 효율성을 높였다. 또한, AdderNet은 기존 합성곱 기반 신경망과 유사한 성능을 유지하면서도 연산 비용을 대폭 절감하였다.

AdderNet은 L1 노름을 활용한 필터와 입력 간의 유사도 계산을 기반으로 설계되었다. 이 과정에서 덧셈 연산이 주요 계산 방식으로 사용되며, 기존 합성곱 기반 신경망의 합성곱 연산을 대체한다. AdderNet은 이러한 계산 방식을 최적화하기 위해 특별히 설계된 역전파(backpropagation) 알고리즘을 도입하였다. 기존 CNN에서의 역전파는 필터와 입력 간의 곱셈 기반 유사도를 바탕으로 그래디언트(Gradient)를 계산하지만, AdderNet은 절대 차이에 기반한 L1 노름 연산을 사용하기 때문에 그래디언트를 계산하는 방식이 달라진다. AdderNet의 역전파 알고리즘은 필터와 입력 간의 차이값에 대해 부호(sign) 정보를 활용하여 그래디언트를 계산하며, 이를 통해 필터의 업데이트 방향을 정확히 결정한다. 또한, AdderNet은 적응형 학습률(adaptive learning rate) 전략을 도입하여 각 계층의 필터 크기에 맞춰 학습률을 조정함으로써 학습 안정성과 수렴 속도를 높였다. 이러한 개선은 에너지 효율성을 극대화하면서도 기존 합성곱 기반 신경망과 유사한 성능을 유지할 수 있게 한다. 이로 인해 네트워크의 수렴 속도를 향상시키고, 성능 저하 없이 에너지 효율성을 극대화하였다.

AdderSR[1]은 AdderNet의 계산 효율성을 초해상화(SR) 문제에 적용하기 위해 설계되었다. AdderNet은 SR 작업에서 항등 매핑(identity mapping)과 고주파 정보 복원의 어려움이 있었다. 이는 AdderNet의 L1 노름 기반 계산 방식이 입력 데이터와 출력 데이터 간의 동일성을 유지하는 데 어려움을 겪기 때문이다. SR 작업에서는 입력 영상의 저해상도 데이터와 고해상도 데이터 간의 구조적 유사성을 보존해야 하지만, AdderNet은 차이값 기반의 계산 특성상 이러한 유사성을 자연스럽게 학습하기 어렵다. 또한, 고주파 정보 복원은 텍스처와 디테일을 강조하기 위해 필터가 미세한 차이를 감지할 수 있어야 하는데, AdderNet의 절대 차이 계산 방식은 이를 효과적으로 처리하지 못한다는 한계가 있었다. 이를 해

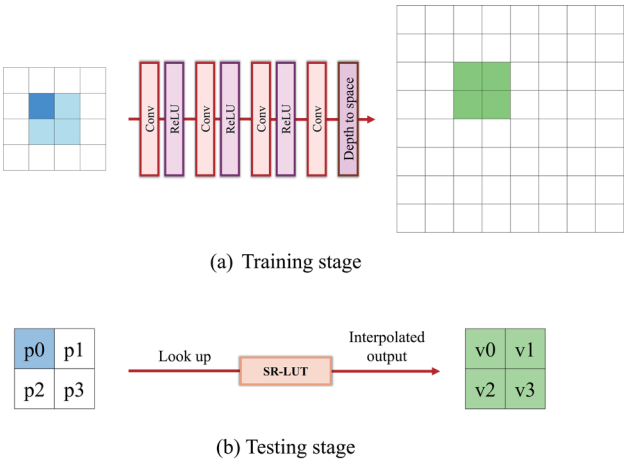
결하기 위해 AdderSR은 셀프 숏컷(self-shortcut) 기법을 도입하였다. self-shortcut은 입력 데이터를 출력과 직접 연결하여 AdderNet이 정보 손실 없이 원본 데이터를 보존하도록 돕는다. 이 기법은 각 AdderNet 계층에서 출력이 입력 데이터를 포함하도록 설계되어, 항등 매핑을 효과적으로 구현한다. 이를 통해 네트워크가 텍스처와 색상 정보를 유지하면서도 고주파 디테일을 보존할 수 있게 하였다.

또한, AdderSR은 고주파 정보를 강조하고 노이즈를 제거하기 위해 학습 가능한 전력 활성화 함수(Learnable Power Activation)를 추가하였다. 이 활성화 함수는 출력 특징을 조정하고 텍스처 및 색상 복원을 강화하기 위해 설계되었다. 구체적으로, 이 함수는 출력 특징값의 부호를 유지하면서, 값의 크기를 제곱 또는 루트 형태로 조정한다. 활성화 함수의 학습 가능한 파라미터는 높은 주파수 영역을 강조하거나 낮은 주파수 정보를 부드럽게 조정할 수 있도록 한다. 예를 들어, 파라미터 값이 1보다 클 경우 고주파 영역을 더욱 선명하게 강조하고, 1보다 작을 경우 노이즈 제거와 부드러운 신호 처리가 가능하다. 이러한 유연성은 AdderSR이 다양한 초해상화 작업에서 적응적으로 작동할 수 있도록 돕는다. 이를 통해 AdderNet의 기존 한계를 극복하며, 초해상화 작업에서도 높은 품질의 영상을 생성할 수 있게 하였다.

AdderSR은 기존 합성곱 신경망 기반 초해상화 모델과 비교하여 PSNR 및 SSIM 지표에서 유사한 성능을 기록하였다. 특히, Set5, Set14, Urban100과 같은 데이터셋에서 에너지 소비를 약 2.5배 줄이면서도, 시각적으로 유사한 품질의 영상을 복원하였다. 이는 AdderSR이 에너지 효율성과 고품질 복원을 동시에 달성할 수 있음을 입증한다.

3.2. 순람표를 이용한 초해상화 모델

Jo et al.은 [11]에서 SR 작업에 적합한 순람표를 구성하는 기법을 제시했다. 이 방법은 복잡한 계산 대신 사전 계산된 값을 순람표에 저장하고, 테스트 시 이를 검색하여 빠르게 고해상도 영상을 생성한다. 이러한 접근은 특히 계산 자원이 제한된 모바일 디바이스 환경에서 실시간



〈그림 10〉 SR-LUT [11]의 학습 과정과 추론 과정

성능을 제공할 수 있는 가능성을 열어주었다.

SR-LUT는 작은 수용 영역(Receptive Field, RF)을 가진 SR 네트워크를 훈련시켜 순람표로 출력 값을 전이하는 과정을 거친다. 순람표에서 사전 계산된 값을 검색하는 과정에서 입력 값이 정확히 순람표에 존재하지 않는 경우가 많기 때문에, 보간(interpolation)이 필요하다. 이는 순람표에 저장된 샘플링된 값들을 기반으로 입력 값에 대한 근사 출력을 생성하는 과정으로, 순람표 기반 시스템에서 필수적인 단계이다. 훈련된 네트워크는 입력 픽셀의 작은 패치(예: 2×2, 3×3)를 기반으로 출력을 생성하며, 이 결과는 순람표에 저장된다. 테스트 단계에서는 순람표에서 해당 값을 검색하고, 다차원 보간(interpolation), 특히 4D tetrahedral interpolation을 이용해 최종 출력을 생성한다.

순람표의 크기를 제한하기 위해 SR-LUT는 균등 샘플링 전략을 사용한다. 균등 샘플링은 입력 값의 원래 범위를 일정 간격으로 나누어 순람표에 저장할 값의 수를 줄이는 방식이다. 순람표는 8비트 입력 값을 기준으로 하면 최대 64GB 크기에 이를 수 있지만, 균등 샘플링 기법을 이용해 순람표의 크기를 1.274MB 까지 줄일 수 있었다. 순람표에 포함되지 않은 입력 값은 다차원 보간법을 사용하여 처리된다. 보간법 중에서도 4차원 테트라헤드럴 보간(tetrahedral interpolation)은 높은 정확도를 제공하면서도 계산 복잡도를 낮추는 데 효과적이다. 테트라헤드럴 보간은 다차원 공간에서의 입력 값들 사이의 관계를

고려하여, 가장 인접한 네 개의 점으로 이루어진 삼각뿔의 무게중심을 계산하는 방식으로 동작한다. SR-LUT는 각 점의 기여도를 가중치로 계산하여 최종 값을 생성하였고, 이러한 기법을 이용해 정확도가 높은 보간 결과를 얻을 수 있었다.

Jo et al.은 실제 모바일 기기에서 SR-LUT를 실험하여 기술의 효용성을 입증하였다. 특히, RF 크기가 2×2인 모델은 삼성 갤럭시 S7에서 15ms만에 SR 작업을 완료하였는데 이는 바이큐빅(bicubic) 보간보다 빠르면서 동시에 더 높은 성능을 보여주었다. 이 외에도 다른 데이터셋에 대해서도 바이큐빅 보간보다 더 빠른 실행 시간과 함께 더 높은 성능을 보였으며, 희소 코딩(Sparse Coding) 기반 및 깊은 신경망(Deep Neural network, DNN) 기반 SR 방법과 비교했을 때, SR-LUT는 더 적은 메모리를 사용하면서도 경쟁력 있는 품질을 유지하였다.

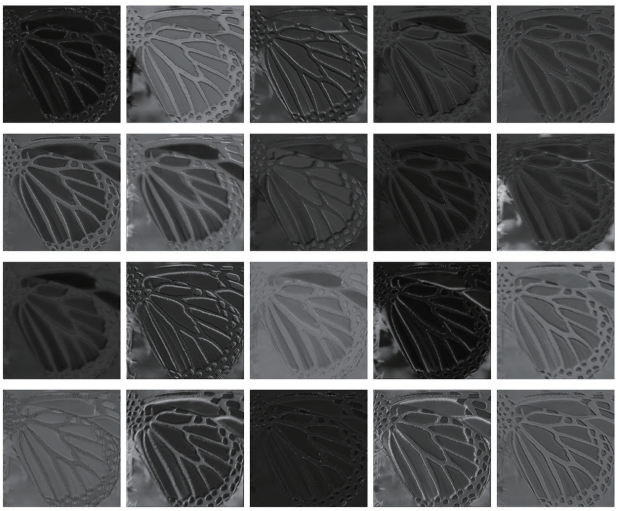
4. 특징, 필터 합성에 관한 연구

Han et.al.은 [14]에서 CNN의 특징 맵 사이에 매우 유사한 형태를 가지는 특징이 종종 등장함을 확인하였다. 〈그림 11〉을 통해 합성곱 신경망 기반의 SR 모델에서 나타나는 특징 맵의 중복을 확인할 수 있다. 이러한 중복은 모델이 높은 성능을 유지하는 데 기여할 수 있지만, 동시에 계산량과 매개변수를 증가시켜 효율적인 연산이 불가능하게 만든다. 이 중복된 특징 맵들을 직접 생성하지 않고, 덧셈과 같이 연산 비용이 적은 방법으로 생성하는 접근을 통해 모델의 연산량을 줄이고 성능을 유지시켰다. 이러한 기법을 SR 모델에 적용한 연구가 바로 GhostSR[7]이다.

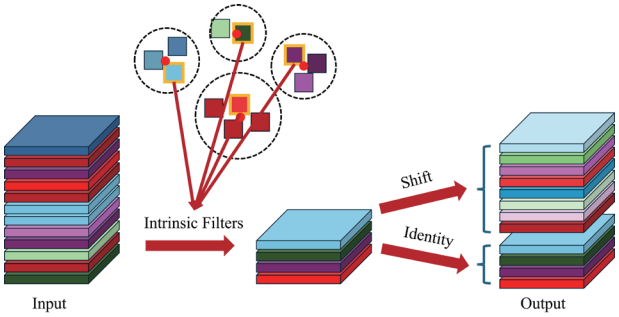
4.1. 합성곱의 고유 특징을 이용해 추가적인 특징을 합성하는 연구

Nie et al.은 GhostNet의 개념을 SR 작업에 적용하여, 값비싼 합성곱 연산 대신 시프트(shift) 연산을 활용하여 유령 특징(ghost feature)을 생성하는 GhostSR 기법을 제시하였다. GhostSR의 핵심 아이디어는 특징 중복(feature redundancy)을 효과적으로 처리하는 것이다. 특징 중복을 해결하기 위해, 우선 입력 특징 맵을 클러스터링하여 유사한 특징을 그룹화하였다. 하나의 클러스터에 대해 중심점에 가장 가까운 특징을 고유 특징(intrinsic feature)으로 정의하였다. 이렇게 얻은 고유 특징으로부터 시프트 연산과 같은 저비용 연산을 통해 유령 특징을 생성한다. 이렇게 얻은 고유 특징과 유령 특징을 결합하여 최종 출력을 생성함으로써, 중복 특징의 유용한 부분은 보존하고 불필요한 연산은 제거하는 균형 잡힌 설계를 실현한다.

이 모델은 검벨-소프트맥스(Gumbel-Softmax) 기법을 도입하여 시프트 연산의 방향성을 학습 가능하게 만들었다. 학습 과정에서 각 계층은 최적의 시프트 방향을 동적으로 학습하며, 이를 통해 유령 특징의 품질을 높인다. 또한, 클러스터링 기법을 사용하여 사전 학습된 모델의



〈그림 11〉 합성곱 연산 기반의 SR 모델인 Real-ESRGAN [24]의 2번째 블록의 마지막 레이어에서 생성된 특징 맵의 일부. 한 레이어 안에 비슷한 특징이 다수 존재한다.



〈그림 12〉 GhostSR [7]의 유령 특징 합성 과정

필터를 분석하고, 고유 특징과 유령 특징을 구분하는 데 활용한다. 이러한 접근은 기존 필터의 정보를 최대한 활용하여 모델 성능을 유지하면서도 경량화를 달성한다.

GhostSR은 Set5, Set14, Urban100 등 다양한 벤치마크 데이터셋에서 검증되었다. 실험 결과, EDSR, RDN, IMDN 등 기존 모델 대비 FLOPs를 최대 46%, 추론 시간을 42%까지 감소시키면서도 PSNR 및 SSIM 성능에서 거의 동일한 결과를 보여주었다. 이는 GhostSR이 경량화와 성능 유지 간의 균형을 효과적으로 달성했음을 입증한다.

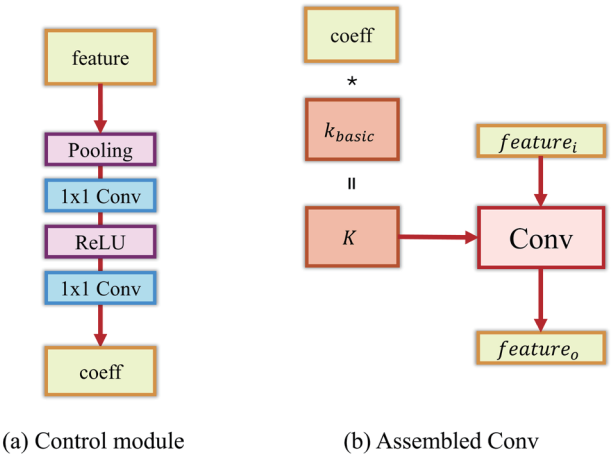
GhostSR의 또 다른 주요 기여는 시프트 연산이 고주파 정보를 강화하고 필터의 수용 영역을 확장하는 데 기여한다는 점이다. 시프트 연산은 텍스처와 디테일 복원에서 중요한 고주파 정보를 증폭시키고, 특징 맵 간의 공간적 정보를 더 넓게 결합할 수 있도록 한다. 이러한 특성은 초해상화 작업의 품질을 더욱 향상시키는 데 기여한다.

GhostSR은 깊이 방향 합성곱(depth-wise convolution)과 비교했을 때, GPU 환경에서 더 빠른 추론 속도를 제공한다. 깊이 방향 합성곱은 메모리 접근 비용이 높아 실제 실행 속도가 느리지만, 시프트 연산은 메모리 접근을 최소화하고 간단한 연산만으로도 유사한 품질을 제공한다. 이러한 차이는 GhostSR이 모바일 및 임베디드 디바이스에서 실용적으로 활용될 수 있음을 보여준다.

4.2. 조합 합성곱을 이용한 실시간 초해상화 연구

AsConvSR은 고해상도(HD, FHD, UHD) 영상과 동영상을 실시간으로 처리하기 위해 설계된 초해상화(SR) 네트워크이다. 이 모델은 기존 네트워크에서 사용된 복잡한 구조를 제거하고, 단순한 연산과 경량화된 설계로도 높은 성능을 유지하도록 설계되었다. 특히, AsConvSR은 픽셀 역셔플(pixel-unshuffle)과 픽셀 셔플(pixel-shuffle) 연산을 네트워크의 시작과 끝에 사용하여 입력 영상의 크기를 효율적으로 조정하며, 전체 연산량을 줄이는 데 기여한다.

AsConvSR은 조합 합성곱(Assembled convolution)이라는 새로운 구조를 도입하여 입력 영상의 다양한 패턴과 텍스처에 맞는 최적의 커널을 생성한다. 컨트롤 모듈(Control module)은 입력 특징을 분석하여 각 채널별로 최적의 커널 계수를 계산한 뒤 이를 조합하여 최적화된 커널을 구성한다. 이 접근 방식은 기존의 동적 합성곱(dynamic convolution)보다 세밀한 조정과 높은 유연성을 제공하며, SR 작업에서 성능과 효율성을 동시에 달성한다.



〈그림 13〉 AsConvSR [6]의 (a) 컨트롤 모듈과 (b) 조합 합성곱

조합 합성곱은 분할 정복의 원칙을 활용하여 설계되었다. 이 접근법은 입력 영상을 다양한 패치로 나누고, 각 패치의 텍스처 복잡도와 패턴에 따라 최적의 커널을 동적으로 적용하는 방식으로 작동한다. 예를 들어, 복잡한 텍스처가 있는 영역에서는 더 세밀하고 고주파 정보를 복원할 수 있는 커널을 사용하고, 단순하고 평탄한 영역에서는 간단한 연산만을 수행하는 커널을 적용한다. 이러한 접근은 전체 연산량을 줄이면서도 영상 복원의 품질을 유지하거나 향상시키는 데 효과적이다.

모델 설계에서는 중간 건너뛰기 연결을 제거하고 네트워크 구조를 단순화하였다. 대신 글로벌 건너뛰기 연결만을 남겨 네트워크의 안정성을 유지하면서도 계산량을 크게 줄였다. 또한, 반복 확대(repeat upscaling) 방식을 도입하여 고해상도 영상을 효율적으로 복원하도록 설계되었다. 이 방식은 바이큐빅 보간과 같은 기존 방법에 비해 더 빠르며, 유사한 품질의 영상을 제공한다.

AsConvSR은 NTIRE 2023 Real-Time Super-Resolution Challenge[23]에서 우수한 성적을 거두며

성능을 입증하였다. 이 모델은 다양한 데이터셋(Set5, Set14, Urban100 등)에서 실험적으로 검증되었으며, 기존 SR 모델들과 비교하여 FLOPs와 매개변수를 대폭 줄이면서도 PSNR과 SSIM에서 경쟁력 있는 성능을 보여주었다. 특히, AsConvSR은 기존 모델들이 성공하지 못했던 1080p 입력에 대한 실시간 처리가 가능하다는 강점을 보였다.

IV. 전망과 결론

영상 초해상화(SR) 모델의 경량화 연구는 최근 몇 년간 실질적인 진전을 이루었으며, 다양한 혁신적 접근법이 등장하고 있다. AdderNet과 GhostNet과 같은 기법은 불필요한 계산을 제거하고, Shift 연산이나 덧셈 연산을 활용해 에너지 효율성을 대폭 개선하였다. 또한, SSR-LUT와 같은 LUT 기반 접근은 미리 계산된 값을 활용해 계산 부담을 최소화하는 방향으로 설계되었으며, AsConvSR과 DIPNet은 고효율의 동적 연산을 통해 모델 경량화와 성능 간의 균형을 맞추고 있다. 이러한 연구들은 초해상화 작업에서 성능과 속도, 에너지 효율성의 균형을 목표로 발전하고 있다.

앞으로 SR 모델의 경량화는 실시간 처리와 더 높은 품질의 영상 생성을 목표로 더욱 발전해야 한다. 이를 위해 경량화된 구조의 최적화를 넘어 데이터 특성에 적응하는 동적 모델링과 하드웨어 친화적인 설계가 필요하다. 특히, 고해상도 비디오와 같은 대규모 데이터 처리에서 효율성을 높이기 위해 하드웨어 가속과 네트워크 구조 간의 협력이 필수적이다. 또한, 경량화 모델이 다양한 응용 분야에 적용될 수 있도록 일반화를 고려한 연구와 함께, 사용자 경험을 향상시키기 위한 주관적 품질 평가를 포함한 다각적인 검증이 중요하다. 이러한 방향은 SR 모델이 모바일 및 엣지 환경뿐만 아니라, 고해상도 콘텐츠 생성의 주요 도구로 자리 잡는 데 기여할 것이다.

참고문헌

[1] D. Song, Y. Wang, H. Chen, C. Xu, C. Xu, and D. Tao, "AdderSR: Towards Energy-Efficient Image Super-Resolution," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[2] H. Chen, Y. Wang, C. Xu, B. Shi, C. Xu, Q. Tian, and C. Xu, "AdderNet: Do We Really Need Multiplications in Deep Learning?" Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[3] X. Ding, X. Zhang, J. Han, and G. Ding, "Diverse Branch Block: Building a Convolution as an Inception-like Unit," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[4] X. Wang, C. Dong, and Y. Shan, "RepSR: Training Efficient VGG-style Super-Resolution Networks with Structural Re-Parameterization and Batch Normalization," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2022.

[5] Z. Hui, X. Gao, Y. Yang, and X. Wang, "Lightweight Image Super-Resolution with Information Multi-Distillation Network," Proceedings of the ACM International Conference on Multimedia (ACM MM), 2019.

[6] J. Guo, X. Zou, Y. Chen, Y. Liu, J. Hao, J. Liu, and Y. Yan, "AsConvSR: Fast and Lightweight Super-Resolution Network with Assembled Convolutions," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2023.

[7] Y. Nie, K. Han, Z. Liu, C. Liu, and Y. Wang, "GhostSR: Learning Ghost Features for Efficient Image Super-Resolution," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022.

[8] N. Ahn, B. Kang, and K.-A. Sohn, "Fast, Accurate, and Lightweight Super-Resolution with Cascading Residual Network," Proceedings of the European Conference on Computer Vision (ECCV), 2018.

[9] F. Kong, M. Li, S. Liu, D. Liu, J. He, Y. Bai, F. Chen, and L. Fu, "Residual Local Feature Network for Efficient Super-Resolution," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2022.

[10] J. Liu, J. Tang, and G. Wu, "Residual Feature Distillation Network for Lightweight Image Super-Resolution," Proceedings of the European Conference on Computer Vision (ECCV), 2020.

[11] Y. Jo and S. J. Kim, "Practical Single-Image Super-Resolution Using Look-Up Table," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

[12] X. Zhang, H. Zeng, and L. Zhang, "Edge-oriented Convolution Block for Real-time Super-Resolution on Mobile Devices," Proceedings of the ACM International Conference on Multimedia (ACM MM), 2021.

[13] L. Yu, X. Li, Y. Li, T. Jiang, Q. Wu, H. Fan, and S. Liu, "DIPNet: Efficiency Distillation and Iterative Pruning for Image Super-Resolution," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[14] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More Features from Cheap Operations," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[15] G. Gendy, G. He, and N. Sabor, "Lightweight image super-resolution based on deep learning: State-of-the-art and future directions," Information Fusion, vol. 94, pp. 284-310, 2023.

[16] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2001.

[17] J.-B. Huang, A. Singh, and N. Ahuja, "Single image super-resolution from transformed self-exemplars," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[18] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced deep residual networks for single image super-resolution," Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017.

[19] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," Proceedings of the 27th International Conference on Machine Learning (ICML), 2010.

[20] K. Simonyan, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[21] Y. Li, Y. Zhang, R. Timofte, L. Van Gool, L. Yu, Y. Li, ... and X. Wang, "NTIRE 2023 challenge on efficient super-resolution: Methods and results," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[22] S. Ioffe, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," arXiv preprint arXiv:1502.03167, 2015.

[23] M. V. Conde, E. Zamfir, R. Timofte, D. Motilla, C. Liu, Z. Zhang, ... and N. Lee, "Efficient deep models for real-time 4k image super-resolution: NTIRE 2023 benchmark and report," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2023.

[24] Wang, X., Xie, L., Dong, C., & Shan, Y. "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data." Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021.

손명준

• 2024년 2월 고려대학교 공과대학 전기전자공학부 학사
• 2024년 3월 ~ 현재 고려대학교 공과대학 전기전자공학과 석사과정

〈관심 분야〉
Computer Vision, Efficient Neural Network, Image Restoration

이홍재

- 2022년 2월 경희대학교 전자정보대학 전자공학과 학사
- 2022년 3월 ~ 현재 고려대학교 공과대학 전기전자공학과 석박사통합과정

〈관심 분야〉
Computer Vision, Efficient Neural Network, Image Restoration

정승원

- 2005년 8월 고려대학교 공과대학 전기전자공학부 학사
- 2011년 2월 고려대학교 공과대학 전기전자공학과 박사
- 2011년 9월 ~ 2012년 2월 고려대학교 정보통신기술연구소 연구교수
- 2012년 3월 ~ 2014년 2월 삼성종합기술원 전문연구원
- 2014년 3월 ~ 2020년 2월 동국대학교 멀티미디어공학과 조교수
- 2020년 3월 ~ 2020년 8월 고려대학교 전기전자공학과 조교수
- 2020년 9월 ~ 현재 고려대학교 전기전자공학과 부교수

〈관심 분야〉
Computer Vision, Efficient Neural Network, Image Restoration, Medical Image Analysis, Neural Networks for ISP, Video Processing

THE INSTITUTE OF ELECTRONICS AND INFORMATION ENGINEERS

논문지 논문목차

전자공학회 논문지 제 62권 2호 발행

통신 분야

[통신]

반도체 분야

[SoC 설계]

컴퓨터 분야

[융합컴퓨팅]

인공지능 신호처리 분야

[영상 신호처리]

- 트리 구조를 이용한 가시성 그래프 기반 레이더 스캔패턴 분류 기법
  김지현, 권순영, 최혁재, 김제안, 김형남
- 연속적 과거 정보를 활용한 비디오 기반 객체 탐지 성능 향상 기법
  류준환, 김민기, 최병인

[음향 및 신호처리]

- 심리음향지표기반 드론 블레이드 파손 고장 진단
  정민철, 김영원

[영상 이해]

- 백본 모델을 이용한 여객 수하물 내 위험화물 검출 성능 비교 연구
  정지욱

시스템 및 제어 분야

[지능로봇]

- 토양 데이터 활용 강화학습 기반 굴착 궤적 생성 알고리즘
  신민규, 조준형, 김민수, 정소이
- Q-러닝을 적용한 행동 트리 기반 굴착 단위 작업 계획 알고리즘
  신민규, 조준형, 김민수, 정소이

산업전자 분야

[반도체재료 및 VLSI 설계]

- CT 분석 방법을 이용한 실리콘 고무 소켓의 저항 증가 분석
  백현빈, 김문정

[신호처리 및 시스템]

- 카메라 변위를 이용한 AR 3D HUD 고스트 영상 측정 방법
  정영주

국내외학술행사안내

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

〉〉2025년 3월

일 자	학술대회명	개최장소	홈페이지/연락처
03.01. - 03.08	2025 IEEE Aerospace Conference	Big Sky, Montana, USA	https://www.aeroconf.org/
03.01. - 03.05	2025 IEEE International Symposium on High Performance Computer Architecture (HPCA)	Las Vegas, Nevada, USA	https://hpca-conf.org/
03.02. - 03.05	2025 IEEE Underwater Technology (UT)	Taipei, Taiwan	https://ut2025.org/
03.03. - 03.07.	2025 IEEE IAS Electrical Safety Workshop (ESW)	Jacksonville, Florida, USA	https://www.ewh.ieee.org/cmte/ias-esw/
03.03. - 03.06.	2025 20th ACM/IEEE International Conference on Human-Robot Interaction (HRI)	Melbourne, Australia	https://humanrobotinteraction.org/2025/
03.04. - 03.07.	2025 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)	Montreal, Quebec, Canada	https://conf.researchr.org/home/saner-2025
03.04. - 03.06.	2025 International Conference on Automation and Computation (AUTOCOM)	Dehradun, India	https://autocom.org.in/
03.05. - 03.07.	2025 International Conference on Emerging Smart Computing and Informatics (ESCI)	Pune, India"	https://esciioit.org/
03.06. - 03.08.	2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)	Gwalior, India"	https://iatmsi.iiitm.ac.in/
03.07. - 03.09.	2025 International Conference on Power Electronics and Electric Drives (PEED)	Dali, China	http://www.icpeed.com/
03.07. - 03.08.	2025 3rd International Conference on Disruptive Technologies (ICDT)	Greater Noida, India	https://www.glbitm.org/icdt-2025/
03.07. - 03.09.	2025 9th International Conference on Green Energy and Applications (ICGEA)	Singapore, Singapore	https://www.icgea.org/index.html
03.07. - 03.08.	2025 3rd International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)	Tumakuru, India"	http://icsses.sit.ac.in/
03.09. - 03.12.	2025 9th IEEE Electron Devices Technology & Manufacturing Conference (EDTM)	Hong Kong, Hong Kong	https://edtm2025.com/
03.10. - 03.12.	2025 International Conference on Machine Learning and Autonomous Systems (ICMLAS)	Prawet, Thailand "	http://icmlas.com/
03.11. - 03.14.	2025 28th Conference on Innovation in Clouds, Internet and Networks (ICIN)	Paris, France	https://www.icin-conference.org/
03.12. - 03.14.	2025 Systems of Signals Generating and Processing in the Field of on Board Communications	Moscow, Russia	http://media-publisher.ru/en/2025-on-board/
03.14. - 03.16.	2025 14th International Conference on Educational and Information Technology (ICEIT)	Guangzhou, China	https://www.iceit.org/index.html
03.14. - 03.16.	2025 IEEE 8th Information Technology and Mechatronics Engineering Conference (ITOEC)	Chongqing, China	http://www.itoec.org/
03.15.	2025 IEEE Integrated STEM Education Conference (ISEC)	Princeton, New Jersey, USA	https://ewh.ieee.org/conf/stem/
03.16. - 03.20.	2025 IEEE Applied Power Electronics Conference and Exposition (APEC)	Atlanta, Georgia, USA	https://apec-conf.org/
03.16. - 03.20.	2025 IEEE Symposium on Computational Intelligence in Artificial Life and Cooperative Intelligent Systems Companion (ALIFE-CIS Companion)	Trondheim, Norway"	https://ieee-ssci.org/

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

일 자	학술대회명	개최장소	홈페이지/연락처
03.30. - 04.03.	2025 Optical Fiber Communications Conference and Exhibition (OFC)	San Francisco, California, USA	https://www.ofcconference.org/en-us/home/
03.30. - 04.03.	2025 IEEE International Reliability Physics Symposium (IRPS)	Monterey, California, USA	https://www.irps.org/
03.31. - 04.04.	2025 IEEE Conference on Software Testing, Verification and Validation (ICST)	Napoli, Italy	https://conf.researchr.org/home/icst-2025
03.31. - 04.04.	2025 IEEE 22nd International Conference on Software Architecture (ICSA)	Odense, Denmark	https://conf.researchr.org/home/icsa-2025/
03.31. - 04.04.	2025 IEEE 22nd International Conference on Software Architecture Companion (ICSA-C)	Odense, Denmark	https://conf.researchr.org/home/icsa-2025
〉〉2025년 4월			
04.02. - 04.04.	2025 IEEE International Conference on Cybernetics and Innovations (ICCI)	Chonburi, Thailand	http://icci2025.smc-thailand.org/
04.03. - 04.05.	2025 2nd International Conference on Trends in Engineering Systems and Technologies (ICTEST)	Ernakulam, India	https://ictest.mec.ac.in/
04.04. - 04.06.	2025 International Conference on Advancement in Communication & Computing in Technology (INOACC)	Bengluru, India	https://inoacc.in/
04.04. - 04.06.	2025 27th International Conference on Enterprise Information Systems (ICEIS)	Porto, Portugal	https://iceis.scitevents.org/
04.04. - 04.06.	2025 20th International Conference on Evaluation of Novel Approaches to Software Engineering (ENASE)	Porto, Portugal	https://enase.scitevents.org/
04.05. - 04.06.	2025 Devices for Integrated Circuit (DevIC)	Kalyani, India	https://edu.ieee.org/in-kgec/devic-2025-home-page/
04.05. - 04.06.	2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI)	Michigan, USA	http://www.icmiconf.com/
04.06. - 04.11.	ICASSP 2025 - 2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)	Hyderabad, India	https://2025.ieeeicassp.org/
04.06. - 04.11.	2025 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)	Hyderabad, India	https://2025.ieeeicassp.org/
04.07. - 04.10.	2025 IEEE International systems Conference (SysCon)	Montreal, Quebec, Canada	https://ieeesystemscouncil.org/event/conference/ieee-international-systems-conference
04.08. - 04.10.	2025 IEEE International Workshop on Integrated Power Packaging (IWIPP)	Tuscaloosa, Alabama, USA	http://iwipp.org/
04.08. - 04.09.	2025 IEEE Texas Symposium on Wireless and Microwave Circuits and Systems (WMCS)	Waco, Texas, USA	https://texassymposium.org/
04.08. - 04.10.	2025 7th International Youth Conference on Radio Electronics, Electrical and Power Engineering (REEPE)	Moscow, Russia	https://reepe.mpei.ru/IEEE/Pages/default.aspx
04.09. - 04.11.	2025 IEEE Conference on Secure and Trustworthy Machine Learning (SaTML)	Copenhagen, Denmark	https://satml.org/
04.10. - 04.11.	2025 International Workshop on Integrated Nonlinear Microwave and Millimetre-Wave Circuits (INMMIC)	Torino, Italy	https://www.inmmic.org/
04.12.	2025 IEEE International Conference on Robotics and Technologies for Industrial Automation (ROBOTHIA)	Kuala Lumpur, Malaysia	https://sites.google.com/view/robothia-malaysia/home
04.13. - 04.15.	2024 International Conference on IT Innovation and Knowledge Discovery (ITIKD)	Manama, Bahrain	http://itikd.ahlia.edu.bh/
04.13. - 04.17.	2025 IEEE Custom Integrated Circuits Conference (CICC)	Boston, Massachusetts, USA	https://www.ieee-cicc.org/
04.14. - 04.17.	2025 26th International Vacuum Electronics Conference (IVEC)	Rotterdam, Netherlands	https://atpi.eventsair.com/ivec-2025
04.14. - 04.17.	2025 IEEE Silicon Photonics Conference (SiPhotonics)	London, United Kingdom	http://www.ieee-siphotonics.org/
04.14. - 04.17.	2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI)	Houston, Texas, USA	https://biomedicalimaging.org/2025/
04.14. - 04.15.	2025 IEEE Wireless and Microwave Technology Conference (WAMICON)	Cocoa Beach, Florida, USA	https://www.ieeewamicon.org/
04.15. - 04.17.	2025 IEEE International Microwave Biomedical Conference (IMBioC)	Kaohsiung, Taiwan	https://www.imbioc2025.org/site/page.aspx?pid=901&sid=1585&lang=en
04.15. - 04.19.	2025 10th Asia Conference on Power and Electrical Engineering (ACPEE)	Beijing, China	https://acpee.net/

일 자	학술대회명	개최장소	홈페이지/연락처
04.16. - 04.17.	2025 2nd International Conference on Microwave, Antennas & Circuits (ICMAC)	Islamabad, Pakistan	http://icmac.seecs.edu.pk/
04.16. - 04.17.	2025 11th International Conference on Web Research (ICWR)	Tehran, Iran	https://iranwebconf.ir/
04.18. - 04.20.	2025 11th International Conference on Control, Automation and Robotics (ICCAR)	Kyoto, Japan	https://www.iccar.org/
04.18. - 04.20.	2025 6th International Conference on Electrical, Electronic Information and Communication Engineering (EEICE)	Shenzhen, China	https://www.eeice.net/
04.18. - 04.20.	2025 International Conference on Sensor-Cloud and Edge Computing System (SCECS)	Zhuhai, China	https://www.scecs.org/
04.20. - 04.23.	2025 IEEE Conference on Technologies for Sustainability (SusTech)	Los Angeles, California, USA	https://ieee-sustech.org/
04.22. - 04.25.	2025 IEEE Global Engineering Education Conference (EDUCON)	London, United Kingdom	https://educon-conference.org/
04.22. - 04.25.	2025 IEEE 18th Pacific Visualization Conference (PacificVis)	Taipei City, Taiwan	https://pacificvis2025.github.io/pages/index.html
04.22. - 04.24.	2025 IEEE 5th International Conference on Digital Twins and Parallel Intelligence (DTPI)	Atlanta, Georgia, USA	http://2025.ieee-dtpi.org/
04.22. - 04.24.	2025 IEEE International Conference on RFID (RFID)	Atlanta, Georgia, USA	https://2025.ieee-rfid.org/
04.23. - 04.24.	2025 IEEE Vision, Innovation, and Challenges Summit and Honors Ceremony (VIC Summit)	Tokyo, Japan	https://corporate-awards.ieee.org/event/vic-summit-honors-ceremony-gala/
04.23. - 04.25.	2025 26th International Symposium on Quality Electronic Design (ISQED)	San Francisco, California, USA	https://www.isqed.org/
04.24. - 04.25.	2025 International Conference on Nanoelectronics, Nanophotonics, Nanomaterials, Nanobioscience & Nanotechnology (5NANO)	Ernakulam, Kerala, India	https://www.5nano2025.com/
04.24. - 04.25.	2025 13th International Workshop on Biometrics and Forensics (IWBF)	Munich, Germany	https://www.unibw.de/iwbf2025
04.25. - 04.26.	2025 IEEE Bangalore Humanitarian Technology Conference (B-HTC)	Belagavi, India	http://bhtc-2025.ieeebangalore.org/
04.25. - 04.27.	2025 10th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)	Chengdu, China	https://www.icccbda.com/index.html
04.25. - 04.27.	2025 IEEE 3rd International Conference on Control, Electronics and Computer Technology (ICCECT)	Jilin, China	http://www.iccect.com/
04.25. - 04.27.	2025 11th International Conference on Electrical Engineering, Control and Robotics (EECR)	Changzhou, China	https://eecr.org/index.html
04.25. - 04.27.	2025 International Conference on Intelligent Transportation and New Energy Technology (ITNET)	Nanning, China	https://www.ic-itnet.com/
04.26. - 05.03.	2025 IEEE/ACM 47th International Conference on Software Engineering (ICSE)	Ottawa, Ontario, Canada	http://www.icse-conferences.org/
04.28. - 05.01.	2025 IEEE Rural Electric Power Conference (REPC)	Denver, Colorado, USA	https://ieeerepc.org/
04.28. - 05.01.	2025 IEEE/ION Position, Location and Navigation Symposium (PLANS)	Salt Lake City, Utah, USA	https://www.ion.org/plans/index.cfm
04.28. - 04.30.	2025 IEEE International Conference on Consumer Technology-Europe (ICCT-Europe)	Algarve, Portugal	https://www.it.pt/Events/Event/5632
04.28. - 05.02.	2025 International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR)	Palermo, Italy	https://www.simpar2025.org/
〉〉2025년 5월			
05.01. - 05.02.	2025 International Conference on Electronics, Computing, Communication and Control Technology (ICECCC)	Bengaluru, India	https://sites.google.com/view/iceccc-2025/home
05.02.	2025 Systems and Information Engineering Design Symposium (SIEDS)	Charlottesville, Virginia, USA	http://engineering.virginia.edu/sieds
05.02. - 05.03.	2025 International Conference on Intelligent and Cloud Computing (ICoICC)	Bhubaneswar, India	http://www.icicc.org.in/
05.03. - 05.09.	2025 IEEE International Radar Conference (RADAR)	Atlanta, Georgia, USA	https://2025.ieee-radar.org/
05.04. - 05.08.	2025 IEEE 33rd Annual International Symposium on Field-Programmable Custom Computing Machines (FCCM)	Fayetteville, Arkansas, USA	https://www.fccm.org/

일 자	학술대회명	개최장소	홈페이지/연락처
05.04. - 05.09.	2025 Conference on Lasers and Electro-Optics (CLEO)	Long Beach, California, USA	https://cleoconference.org/
05.04. - 05.07.	2025 IEEE 32nd Symposium on Computer Arithmetic (ARITH)	El Paso, Texas, USA	https://www.arith2025.org/
05.04. - 05.06.	2025 IEEE 3rd International Conference on Mobility, Operations, Services and Technologies (MOST)	Newark, Delaware, USA	http://ieeemobility.org/MOST2025/index.php
05.04. - 05.08.	2025 IEEE-IAS/PCA Cement Industry Conference (IAS/PCA)	Birmingham, Alabama, USA	https://cementconference.org/
05.05. - 05.07.	2025 IEEE 28th International Symposium on Design and Diagnostics of Electronic Circuits and Systems (DDECS)	Lyon, France	https://cas.polito.it//DDECS2025/
05.05. - 05.07.	2025 IEEE Conference on Artificial Intelligence (CAI)	Santa Clara, California, USA	https://cai.ieee.org/2025/
05.05. - 05.08.	2025 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)	San Jose, California, USA	http://www.hostsymposium.org/
05.05. - 05.07.	2025 Conference on AI x Software Engineering	Laguna Hills, California, USA	https://www.aixse.org/
05.05. - 05.08.	IEEE INFOCOM 2025 - IEEE Conference on Computer Communications	London, United Kingdom	https://infocom2022.ieee-infocom.org/
05.05. - 05.09.	2025 II International Congress of Engineering with a Social Sense (CIISS)	Bogotá D.C., Colombia	https://eventos.uniminuto.edu/107918
05.06. - 05.09.	2025 11th International Conference on Engineering, Applied Sciences, and Technology (ICEAST)	Phuket, Thailand	https://iceast.kmitl.ac.th/2025/
05.06. - 05.09.	2025 IEEE 31st Real-Time and Embedded Technology and Applications Symposium (RTAS)	Irvine, California, USA	http://2025.rtas.org/
05.08. - 05.09.	2025 IEEE Build-Up Substrate Symposium (BUSS)	Milpitas, California, USA	http://ieee-buss.org/
05.09.	2025 IEEE Sensors in Spotlight (SenSiS)	Boston, Massachusetts, USA	https://2025.ieee-sensorsinspotlight.org/
05.09. - 05.11.	2025 Global Conference in Emerging Technology (GINOTECH)	PUNE, India	https://ginotech.org/
05.09. - 05.10.	2025 Intermountain Engineering, Technology and Computing (IETC)	Orem, Utah, USA	https://www.uvu.edu/cet/i-etc/
05.09. - 05.11.	2025 IEEE 2nd International Conference on Big Data Science and Engineering (ICBDSE)	Kunming, China	https://www.icbdse.org/
05.09. - 05.11.	2025 4th International Symposium on New Energy Technology Innovation and Low Carbon Development (NET-LC)	Hangzhou, China	https://www.net-lc.net/
05.09. - 05.11.	2025 5th International Conference on Mechanical, Electronics and Electrical and Automation Control (METMS)	Chongqing, China	https://www.met-ms.com/
05.09. - 05.11.	2025 10th International Conference on Control and Robotics Engineering (ICCRE)	Nagoya, Japan	https://www.iccre.org/
05.11. - 05.14.	2025 IEEE 29th Workshop on Signal and Power Integrity (SPI)	Gaeta, Italy	https://spi-workshop.org/
05.11. - 05.13	2025 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)	Ghent, Belgium	http://ispass.org/ispass2025/
05.11. - 05.14.	2025 IEEE Energy Conversion Congress & Exposition Asia (ECCE-Asia)	Bengaluru, India	https://ecceasia2025.org/
05.11. - 05.14.	2025 IEEE/IAS 61st Industrial and Commercial Power Systems Technical Conference (I&CPS)	Montreal, Quebec, Canada	https://site.ieee.org/ias-icps/main-conference/
05.11. - 05.14.	2025 29th IEEE International Symposium on Asynchronous Circuits and Systems (ASYNC)	Portland, Oregon, USA	https://asyncsymposium.org/
05.12. - 05.15.	2025 IEEE Symposium on Security and Privacy (SP)	San Francisco, California, USA	https://sp2025.ieee-security.org/
05.12. - 05.16.	2025 Joint Conference of the European Frequency and Time Forum and IEEE International Frequency Control Symposium (EFTF/IFCS)	Querétaro, Mexico	https://2025.ieee-ifcs-eftf.org/
05.12. - 05.16.	2025 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)	Sochi, Russia	https://icieam.su-ieee.ru/
05.12. - 05.14.	2025 ESA Workshop on Aerospace EMC (Aerospace EMC)	Seville, Spain	https://atpi.eventsair.com/emc-2025
05.12. - 05.16.	2025 International Wireless Communications and Mobile Computing (IWCMC)	Abu Dhabi, United Arab Emirates	https://iwcmc.net/2025/
05.12. - 05.16.	NOMS 2025-2025 IEEE Network Operations and Management Symposium	Honolulu, Hawaii, USA	https://noms2025.ieee-noms.org/
05.12. - 05.16.	2025 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF)	St. Petersburg, Russia	http://media-publisher.ru/en/about-weconf-2025/

일 자	학술대회명	개최장소	홈페이지/연락처
05.12. - 05.16.	2025 IEEE Magnetic Frontiers: Altermagnetism - New Opportunities in Magnetism	Liblice, Czech Republic	https://www.frontiers2025.cz/
05.13. - 05.14.	2025 International Conference on Military Communication and Information Systems (ICMCIS)	Oerias, Portugal	https://www.icmcis.eu/
05.14. - 05.16.	2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST)	Astana, Kazakhstan	https://sist.astanait.edu.kz/
05.14. - 05.16.	2025 5th International Conference on Artificial Intelligence and Education (ICAIE)	Suzhou, China	https://www.icaie.org/index.html
05.14. - 05.18.	2025 International Spring Seminar on Electronics Technology (ISSE)	Budapest, Hungary	https://isse-conf.eu/
05.14. - 05.16.	2025 5th International Conference on Pervasive Computing and Social Networking (ICPCSN)	Salem, India	https://icpcsn.com/
05.15. - 05.16.	2025 5th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)	Fez, Morocco	http://www.iraset.org/2025/
05.15. - 05.17.	2025 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2025 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)	Timisoara, Romania	https://info-optim.ro/
05.16. - 05.19.	2025 37th Chinese Control and Decision Conference (CCDC)	Xiamen, China	http://www.ccdc.neu.edu.cn/
05.16. - 05.18.	2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC)	Bhubaneswar, India	http://assic.info/index.php
05.16. - 05.18.	2025 IEEE 3rd International Conference on Power Science and Technology (ICPST)	Kunming, China	https://www.icpst.org/
05.16. - 05.18.	2025 The 16th International Conference on Mechanical and Intelligent Manufacturing Technologies (ICMIMT)	Cape Town, South Africa	https://www.mimt.us/
05.16. - 05.17.	2025 3rd International Conference on Data Science and Information System (ICDSIS)	Hassan, India	https://icdsis.in/
05.16. - 05.18.	2025 IEEE 8th International Electrical and Energy Conference (CIEEC)	Changsha, China	https://www.cieec.com.cn/
05.16. - 05.18.	2025 5th International Conference on Computer, Control and Robotics (ICCCR)	Hangzhou, China	https://www.icccr.org/
05.16. - 05.18.	2025 2nd International Conference on Intelligent Computing and Robotics (ICICR)	Dalian, China	http://www.icrconf.com/
05.16. - 05.18.	2025 10th International Conference on Intelligent Computing and Signal Processing (ICSP)	Xi'an, China	https://www.ic-icsp.org/
05.16. - 05.18.	2025 IEEE International Annual Conference on Complex Systems and Intelligent Science (CSIS-IAC)	Shenzhen, China	http://www.conference123.org/csisiac2025/
05.16. - 05.18.	2025 8th International Conference on Communication Engineering and Technology (ICCET)	Guangzhou, China	https://www.iccet.org/
05.16. - 05.18.	2025 4th International Conference on Electronics, Integrated Circuits and Communication Technology (EICCT)	Chengdu, China	https://www.ic-eict.net/
05.17. - 05.20.	2025 8th International Conference on Electronics Technology (ICET)	Chengdu, China	https://www.icet.net/
05.18. - 05.21.	2025 IEEE International Memory Workshop (IMW)	Monterey, California, USA	https://www.ewh.ieee.org/soc/eds/imw/
05.18. - 05.21.	2025 IEEE International Electric Machines & Drives Conference (IEMDC)	Houston, Texas, USA	https://www.iemdc.org/
05.18. - 05.21.	2025 15th International Symposium on Linear Drivers for Industry Applications (LDIA)	Daejeon, Korea (South)	http://www.ldia2025.com/
05.19. - 05.23.	2025 IEEE International Conference on Robotics and Automation (ICRA)	Atlanta, Georgia, USA	https://2025.ieee-icra.org/
05.19. - 05.22.	2025 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)	Chemnitz, Germany	https://i2mtc2025.ieee-ims.org/
05.19. - 05.23.	2025 Asia-Pacific International Symposium and Exhibition on Electromagnetic Compatibility (APEMC)	Taipei, Taiwan	https://www.apemc2025.org/site/page.aspx?pid=901&sid=1565&lang=en
05.19. - 05.21.	2025 IEEE 34st Microelectronics Design & Test Symposium (MDTS)	Albany, New York, USA	https://mdts.ieee.org/
05.19. - 05.22.	2025 10th International Conference on Fog and Mobile Edge Computing (FMEC)	Tampa, Florida, USA	https://emergingtechnet.org/FMEC2025/index.php

일 자	학술대회명	개최장소	홈페이지/연락처
05.19. - 05.22.	2025 International Conference on Modern Computing, Networking and Applications (MCNA)	Tampa, Florida, USA	http://mcna-conference.org/2025/index.php
05.19. - 05.22.	2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing (CCGrid)	Tromsø, Norway	https://site.uit.no/ccgrid2025/
05.19. - 05.22.	2025 IEEE 25th International Symposium on Cluster, Cloud and Internet Computing Workshops (CCGridW)	Tromsø, Norway	https://site.uit.no/ccgrid2025/
05.19. - 05.22.	2025 IEEE Conference on Education and Training in Optics and Photonics (ETOP)	Glasgow, United Kingdom	http://www.etop-conference.org/
05.19. - 05.22.	2025 IEEE Emerging Technology Reliability Roundtable (ETR-RT)	Prague, Czech Republic	https://cqr.committees.comsoc.org/etr-rt-2025/
05.19. - 05.22.	2025 IEEE 9th International Conference on Fog and Edge Computing (ICFEC)	Tromso, Norway	https://icfec2025.ontariotechu.ca/
05.19.	IEEE INFOCOM 2025 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)	London, United Kingdom	https://infocom2025.ieee-infocom.org/authors/call-workshop-proposals
05.19. - 05.23.	2025 IEEE 41st International Conference on Data Engineering (ICDE)	Hong Kong, Hong Kong	https://ieee-icde.org/2025/
05.20. - 05.22.	2025 IEEE 19th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG)	Antalya, Turkiye	https://cpepowereng2025.org/
05.20. - 05.22.	2025 IEEE 26th International Conference on High Performance Switching and Routing (HPSR)	Suita, Osaka, Japan	https://hpsr2025.ieee-hpsr.org/
05.20. - 05.23.	2025 22nd International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)	Bangkok, Thailand	https://ecti-con2025.eng.chula.ac.th/
05.21. - 05.23.	2025 Third International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)	Trichy, India	http://icaiss.com/
05.21. - 05.23.	2025 26th International Radar Symposium (IRS)	Hamburg, Germany	https://www.dgon-irs.org/home
05.23. - 05.25.	2025 International Conference in Advances in Power, Signal, and Information Technology (APSIT)	Bhubaneswar, India	https://apsit.co.in/
05.23. - 05.25.	2025 6th International Conference for Emerging Technology (INCET)	BELGAUM, India	http://www.incet.org/
05.23. - 05.25.	2025 IEEE 12th Joint International Information Technology and Artificial Intelligence Conference (ITAIC)	Chongqing, China	http://www.itaic.org/
05.23. - 05.25.	2025 2nd International Symposium on New Energy Technologies and Power Systems (NETPS)	Hangzhou, China	https://www.netps.org/
05.23. - 05.25.	2025 IEEE International Conference on Electrical Energy Conversion Systems and Control (IEECSC)	Chongqing, China	https://ieee-ieecsc.net/2025/
05.23. - 05.26.	2025 8th International Conference on Artificial Intelligence and Big Data (ICAIBD)	Chengdu, China	https://www.icaibd.org/
05.23. - 05.25.	2025 6th International Conference on Control, Communication and Computing (ICCC)	Thiruvanathapuram, India	https://iccc2025.cet.ac.in/
05.23. - 05.25.	025 IEEE 5th International Conference on Electronic Technology, Communication and Information (ICETCI)	Changchun, China	http://www.icetci.org/
05.23. - 05.24.	2025 7th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (ICHORA)	Ankara, Turkiye	http://www.ichoracongress.com/
05.23. - 05.26.	2025 IEEE International Workshop on Radio Frequency and Antenna Technologies (iWRF&AT)	Shenzhen, China	https://rfat.szu.edu.cn/rfat2025/
05.23. - 05.25.	2025 6th International Conference on Computing, Networks and Internet of Things (CNIOT)	Shanghai, China	http://www.cniot-conf.net/
05.23. - 05.25.	2025 IEEE 5th International Conference on Computer Communication and Artificial Intelligence (CCAI)	Haikou, China	http://www.ccai.net/
05.23. - 05.26.	2025 17th International Conference on Bioinformatics and Biomedical Technology (ICBBT)	Hangzhou, China	https://www.icbbt.org/index.htm
05.23. - 05.25.	2025 6th International Conference on Computer Vision, Image and Deep Learning (CVIDL)	Ningbo, China	https://www.cvidl.org/
05.24. - 05.25.	2025 IEEE 15th Symposium on Computer Applications & Industrial Electronics (ISCAIE)	Penang, Malaysia	https://www.iscaie.org/home
05.25. - 05.28.	2025 IEEE International Symposium on Circuits and Systems (ISCAS)	London, United Kingdom	https://iscas2025.org/lander

일 자	학술대회명	개최장소	홈페이지/연락처
05.26. - 05.31.	2025 IEEE 75th Electronic Components and Technology Conference (ECTC)	Dallas, Texas, USA	https://www.ectc.net/
05.26. - 05.30.	2025 IEEE 19th International Conference on Automatic Face and Gesture Recognition (FG)	Tampa/Clearwater, Florida, USA	https://fg2025.ieee-biometrics.org/
05.26. - 05.29.	\2025 IEEE International Conference on Machine Learning for Communication and Networking (ICMLCN)	Barcelona, Spain	https://icmlcn2025.ieee-icmlcn.org/
05.26. - 05.28	2025 IEEE 5th International Conference on Human-Machine Systems (ICHMS)	Abu Dhabi, United Arab Emirates	http://ieee-ichms.org/2025/
05.26. - 05.28	2025 IEEE European Technology and Engineering Management Summit (E-TEMS)	Bruges, Belgium	https://etems.digital/2025/
05.26. - 05.29.	2025 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)	Vancouver, British Columbia, Canada	https://attend.ieee.org/ccece-2025/
05.27. - 05.30.	2025 24th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)	Dallas, Texas, USA	https://www.ieee-itherm.net/
05.27. - 05.30.	2025 International Conference on Military Technologies (ICMT)	Brno, Czech Republic	https://icmt2025.cz/
05.27. - 05.30.	2025 12th International Conference on Information Technology (ICIT)	Amman, Jordan	https://icit.zuj.edu.jo/Home/
05.28. - 05.30.	2025 7th International Conference on Inventive Material Science and Applications (ICIMA)	Namakkal, India	https://icoimsa.com/2025/
05.28. - 05.29.	2025 IEEE Zooming Innovation in Consumer Technologies Conference (ZINC)	Novi Sad, Serbia	https://www.gozinc.org/
05.28. - 05.30.	2025 IEEE Medical Measurements & Applications (MeMeA)	Chania, Greece	http://www.memea2024.ieee-ims.org/
05.29. - 05.30.	2025 IEEE Engineering Reliable Autonomous Systems (ERAS)	Worcester, Massachusetts, USA	https://erasrobotics.org/
05.29. - 05.31	2025 IEEE International Conference on Electro Information Technology (eIT)	Valparaiso, Indiana, USA	https://eit-conference.org/eit2025/
05.30. - 06.01.	2025 International Conference on Artificial Intelligence and Digital Ethics (ICAIDE)	Guangzhou, China	https://www.icaide.net/
〉〉2025년 6월			
06.01. - 06.05.	2025 37th International Symposium on Power Semiconductor Devices and ICs (ISPSD)	Kumamoto, Japan	https://www.ispsd2025.com/
06.02. - 06.05.	2025 IEEE 101st Vehicular Technology Conference (VTC2025-Spring)	Oslo, Norway	https://events.vtsociety.org/vtc2025-spring/
06.02. - 06.05.	2025 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)	Duisburg, Germany	https://edas.info/web/cogsima2025/home.html
06.02. - 06.04.	2025 IEEE Vehicular Networking Conference (VNC)	Porto, Portugal	https://vnc2025.ieee-vnc.org/
06.02. - 06.05.	2025 26th IEEE International Conference on Mobile Data Management (MDM)	Irvine, California, USA	https://mdm2025.github.io/
06.02. - 06.06.	2025 MIPRO 48th ICT and Electronics Convention	Opatija, Croatia	http://www.mipro.hr/
06.03. - 06.06.	2025 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)	Poznan, Poland	https://www.eucnc.eu/
06.03. - 06.06.	2025 25th International Conference on Process Control (PC)	Štrbské Pleso, Slovakia	https://www.process-control.sk/
06.03. - 06.04.	2025 International Conference on Smart Computing, IoT and Machine Learning (SIML)	Surakarta, Indonesia	https://siml.ums.ac.id/2025/
06.03. - 06.06.	2025 IEEE Wireless Power Technology Conference and Expo (WPTCE)	Rome, Italy	https://ieee-wptce.org/
06.03. - 06.07.	2025 IEEE International Parallel and Distributed Processing Symposium (IPDPS)	Milano, Italy	http://www.ipdps.org/
06.04. - 06.06.	2025 IEEE 55th International Symposium on Multiple-Valued Logic (ISMVL)	Montreal, Quebec, Canada	https://mvl.jpn.org/ISMVL2024/
06.04. - 06.06.	IEEE EUROCON 2025 - 21st International Conference on Smart Technologies	Gdynia, Poland	https://2025.ieee-eurocon.org/
06.04. - 06.06.	2025 6th International Conference in Electronic Engineering & Information Technology (EEITE)	Chania, Greece	https://eeite.hmu.gr/
06.04. - 06.06.	2025 IEEE Seventh International Conference on DC Microgrids (ICDCM)	Tallinn, Estonia	https://taltech.ee/en/ieee-icdcm-2025/

일 자	학술대회명	개최장소	홈페이지/연락처
06.04. - 06.06.	2025 Tenth Conference on Lighting (Lighting)	Sozopol, Bulgaria	https://conference.nko.bg/
06.05. - 06.08.	2025 IEEE International Conference on Predictive Control of Electrical Drives and Power Electronics (PRECEDE)	Nanjing, China	http://www.precede2025.com/
06.06. - 06.07.	2025 International Conference on Emerging Trends in Industry 4.0 Technologies (ICETI4T)	Navi Mumbai, India	https://iceti4t.siesgst.edu.in/
06.06. - 06.08.	2025 4th International Symposium on Robotics, Artificial Intelligence and Information Engineering (RAIIE)	Zhenjiang, China	https://www.raiie.org/
06.06. - 06.07.	2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)	Bengaluru, India	https://icicke.in/
06.06. - 06.08.	2025 IEEE International Symposium on Ethics in Engineering, Science, and Technology (ETHICS)	Evanston, Illinois, USA	https://attend.ieee.org/ethics-2025/
06.06. - 06.07.	2025 IEEE International Conference on Hydraulic Engineering Calculation and Simulation Technology (HECST)	Shanghai, China	https://www.hecstconf.com/
06.07. - 06.11.	2025 IEEE Electrical Insulation Conference (EIC)	South Padre Island, Texas, USA	https://ieee-eic.org/
06.08. - 06.12.	ICC 2025 - IEEE International Conference on Communications	Montreal, Quebec, Canada	https://icc2025.ieee-icc.org/
06.08. - 06.12.	2025 IEEE Congress on Evolutionary Computation (CEC)	Hangzhou, China	https://www.cec2025.org/
06.08. - 06.13.	2025 IEEE 53rd Photovoltaic Specialists Conference (PVSC)	Montreal, Quebec, Canada	https://ieee-pvsc.org/PVSC53/
06.08. - 06.12.	2025 Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits)	Kyoto, Japan	https://www.vlsisymposium.org/
06.09. - 06.11.	2025 IEEE International Conference on Prognostics and Health Management (ICPHM)	Denver, Colorado, USA	https://phmconf.org/
06.09. - 06.13.	2025 IEEE Conference on Norbert Wiener in the 21st Century (21CW)	Kitakyushu, Japan	https://www.21stcenturywiener.org/
06.09. - 06.11.	2025 21st International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)	Lucca, Italy	https://dcoss.org/
06.10. - 06.12.	2025 IMAPS Nordic Conference on Microelectronics Packaging (NordPac)	Copenhagen, Denmark	https://nordic.imapseurope.org/nordpac/
06.10. - 06.12.	2025 International Conference on Localization and GNSS (ICL-GNSS)	Rome, Italy	https://events.tuni.fi/icl-gnss/
06.10 - 06.13	2025 IEEE 21st International Conference on Factory Communication Systems (WFCS)	Rostock, Germany	https://wfcs25.uni-rostock.de/
06.10. - 06.14.	2025 14th Mediterranean Conference on Embedded Computing (MECO)	Budva, Montenegro	https://mecoconference.me/meco2025/
06.10. - 06.17.	2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)	Nashville, Tennessee, USA	http://cvpr2025.thecvf.com/
06.11. - 06.13.	2025 7th Global Power, Energy and Communication Conference (GPECOM)	Bochum, Germany	https://gpecom.org/2025/
06.11. - 06.13.	2025 4th International Conference on Advances in Computing, Communication, Embedded and Secure Systems (ACCESS)	Ernakulam, India	https://sites.google.com/view/access25
06.11. - 06.13.	2025 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv)	Venezia, Italy	https://www.metrolivenv.org/
06.11. - 06.13.	2025 IEEE Evolution - Life Members Conference	Boston, Massachusetts, USA	https://lifememberscconference.ieee.org/
06.12. - 06.13.	2025 IEEE/ACM 34th International Workshop on Logic and Synthesis (IWLS)	Verona, Italy	http://www.iwls.org/iwls2025/
06.13. - 06.14.	2025 International Conference on Computing Technologies (ICOCT)	Bengaluru, India	https://icoct.in/
06.13. - 06.15	2025 IEEE International Conference on Electron Devices and Solid-State Circuits (EDSSC)	Yinchuan, China	http://www.edssc2025.com/
06.13. - 06.15	2025 International Symposium on Intelligent Robotics and Systems (ISoIRS)	Chengdu, China	https://www.isoirs.org/
06.13. - 06.16.	2025 International Conference on Distance Education and Learning (ICDEL)	Kunming, China	http://www.icdel.org/index.html
06.13. - 06.14.	2025 3rd Cognitive Models and Artificial Intelligence Conference (AICCONF)	Prague, Czech Republic	http://ai-conf.com/

일 자	학술대회명	개최장소	홈페이지/연락처
06.13. - 06.15	2025 International Conference on Advanced Energy Systems and Power Electronics (AESPE)	Hangzhou, China	http://aespe.net/
06.13. - 06.15	2025 11th International Symposium on Sensors, Mechatronics and Automation System (ISSMAS)	Zhuhai, China	https://www.is-smas.net/
06.15. - 06.20.	2025 IEEE/MTT-S International Microwave Symposium - IMS 2025	San Francisco, California, USA	https://ims-ieee.org/
06.15. - 06.17.	2025 IEEE Radio Frequency Integrated Circuits Symposium (RFIC)	San Francisco, California, USA	https://rfic-ieee.org/
06.15. - 06.20.	2025 IEEE Industry Applications Society Annual Meeting (IAS)	Taipei, Taiwan	https://ias-am.ieee.org/2025/
06.15. - 06.20.	2025 IEEE 23nd International Forum on MPSoC for Software-Defined Hardware (MPSoC)	Megeve, France	http://mpsoc-forum.org/
06.16. - 06.20.	2025 IEEE 38th Computer Security Foundations Symposium (CSF)	Santa Cruz, California, USA	https://csf2025.ieee-security.org/index.html
06.16. - 06.20.	2025 10th International Conference on Smart and Sustainable Technologies (SpliTech)	Bol and Split, Croatia	https://splitech.org/
06.16. - 06.17.	2025 International Conference on Smart Learning Courses (SCME)	Hebron, Palestine	https://scme.edu.ps/
06.17. - 06.19.	2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)	Tirunelveli, India	http://icicv.org/conf2025/
06.17. - 06.20.	2025 IEEE 34th International Symposium on Industrial Electronics (ISIE)	Toronto, Ontario, Canada	https://ieee-isie-2025.org/
06.18. - 06.20	2025 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)	Dublin, Ireland	http://bmsb2025.com/
06.18. - 06.21.	2025 IEEE 13th International Conference on Healthcare Informatics (ICHI)	Rende, Italy	https://events.dimes.unical.it/ichi2025/
06.18. - 06.20	2025 IEEE 38th International Symposium on Computer-Based Medical Systems (CBMS)	Madrid, Spain	https://2025.cbms-conference.org/
06.18. - 06.20	2025 IEEE/AIAA Transportation Electrification Conference and Electric Aircraft Technologies Symposium (ITEC+EATS)	Anaheim, California, USA	https://itec-conf.com/
06.18. - 06.21.	2025 International Conference on Power Electronics Converters for Transportation and Energy Applications (PECTEA)	Jatni, India	http://conference.iitbbs.ac.in/pectea2025/
06.18. - 06.20	2025 IEEE 12th International Workshop on Metrology for AeroSpace (MetroAeroSpace)	Naples, Italy	https://www.metroaerospace.org/
06.18. - 06.20	2025 IEEE Guwahati Subsection Conference (GCON)	Itanagar, India	https://gcon2025.in/
06.19. - 06.20.	2025 1st International Conference on Radio Frequency Communication and Networks (RFCoN)	Thanjavur, India	https://sastra.edu/rfcoe23/
06.19. - 06.20.	2025 IEEE 2nd International Conference on Blockchain, Smart Healthcare and Emerging Technologies (SmartBlock4Health)	Bucharest, Romania	https://smartblock4health.upb.ro/
06.19. - 06.21.	2025 19th Conference on Electrical Machines, Drives and Power Systems (ELMA)	Sofia, Bulgaria	https://e-university.tu-sofia.bg/e-conf/?konf=227
06.20. - 06.22.	2025 5th International Conference on Intelligent Technologies (CONIT)	HUBBALI, India	https://inconf.in/
06.20. - 06.22.	2025 2nd International Conference on Digital Media, Communication and Information Systems (DMCIS)	Nanjing, China	https://www.icdmcis.org/
06.20. - 06.22.	2025 6th International Conference on Electrical Technology and Automatic Control (ICETAC)	Nanjing, China	https://www.icetac.org/
06.20. - 06.21.	2025 IEEE 2nd International Conference on Energy and Electrical Engineering (EEE)	Nanchang, China	https://www.iceeeconf.com/
06.20. - 06.22.	2025 9th International Conference on Power Energy Systems and Applications (ICoPESA)	Nanjing, China	https://www.icpesa.org/
06.20. - 06.22.	2025 IEEE International Conference on Pattern Recognition, Machine Vision and Artificial Intelligence (PRMVAI)	Loudi, China	https://www.prmvai.org/
06.20. - 06.22.	2025 IEEE 5th International Conference on Software Engineering and Artificial Intelligence (SEAI)	Fuzhou, China	https://www.seai.org/
06.20. - 06.22.	2025 10th International Symposium on Advances in Electrical, Electronics and Computer Engineering (ISAEECE)	Xi'an, China	https://www.isaeece.com/
06.21. - 06.24.	2025 23rd IEEE Interregional NEWCAS Conference (NEWCAS)	Paris, France	http://www.newcas2025.com/

일 자	학술대회명	개최장소	홈페이지/연락처
06.21. - 06.25.	2025 ACM/IEEE 52nd Annual International Symposium on Computer Architecture (ISCA)	Tokyo, Japan	http://iscaconf.org/isca2025/
06.22. - 06.26	2025 62nd ACM/IEEE Design Automation Conference (DAC)	San Francisco, California, USA	https://www.dac.com/
06.22. - 06.26.	2025 25th International Conference on the Computation of Electromagnetic Fields (COMPUMAG)	Naples, Italy	https://www.compumag2025.com/cm/
06.22. - 06.25	2025 IEEE Intelligent Vehicles Symposium (IV)	Cluj-Napoca, Romania	https://ieee-iv.org/2025/
06.22. - 06.25	2025 IEEE International Conference on Flexible and Printable Sensors and Systems (FLEPS)	Singapore	https://2025.ieee-fleps.org/
06.22. - 06.26.	2025 Joint Conference on Electrostatics (ESA)	St. Catharines, Ontario, Canada	https://electrostatics.org/annual-meeting-2/
06.22. - 06.25.	2025 IEEE 16th International Symposium on Power Electronics for Distributed Generation Systems (PEDG)	Nanjing, China	https://www.pedg2025.com/
06.23. - 06.27.	2025 Conference on Lasers and Electro-Optics Europe & European Quantum Electronics Conference (CLEO/Europe-EQEC)	Munich, Germany	https://www.cleoeurope.org/
06.23. - 06.26.	2025 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)	Naples, Italy	https://dsn2025.github.io/
06.23. - 06.26.	2025 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks - Supplemental Volume (DSN-S)	Naples, Italy	https://dsn2025.github.io/
06.23. - 06.26.	2025 55th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)	Naples, Italy	https://dsn2025.github.io/
06.23. - 06.26.	2025 21st International Conference on Intelligent Environments (IE)	Darmstadt, Germany	https://www.ie2025.fraunhofer.de/
06.23. - 06.25.	2025 Silicon Valley Cybersecurity Conference (SVCC)	San Francisco, California, USA	https://www.svcc-svcsi.org/
06.23. - 06.26.	2025 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom)	Chisinau, Moldova	https://blackseacom2025.ieee-blackseacom.org/
06.23. - 06.27.	2025 IEEE 11th International Conference on Network Softwarization (NetSoft)	Budapest, Hungary	https://netsoft2025.ieee-netsoft.org/
06.23. - 06.27.	2025 URSI International Symposium on Electromagnetic Theory (EMTS)	Bologna, Italy	https://www.emts2025.it/
06.24. - 06.26.	2025 International Conference on Clean Electrical Power (ICCEP)	Villasimius, Italy	https://www.iccep.net/
06.24. - 06.26.	2025 IEEE/ACM Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)	New York, New York, USA	https://conferences.computer.org/chase2025/
06.24.	2025 IEEE International Conference on AI and Data Analytics (ICAD)	Waltham, Massachusetts, USA	http://ieee-icad.org/
06.25. - 06.27.	2025 6th International Conference on Inventive Research in Computing Applications (ICIRCA)	Coimbatore, India	http://icircа18.com/
06.25. - 06.27.	2025 25th International Conference on Digital Signal Processing (DSP)	Pylos (Messinia, Southwest Peloponnese), Greece	https://2025.ic-dsp.org/
06.25. - 06.27.	2025 IEEE/ACIS 29th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD)	Busan, Korea (South)	https://acisinternational.org/conferences/snpd-2025-summeri/
06.25. - 06.26.	2025 Second International Conference on Cognitive Robotics and Intelligent Systems (ICC - ROBINS)	Coimbatore, India	https://kpriet.ac.in/conference/icc-robins
06.26. - 06.27.	2025 IEEE International Conference on LLM-Aided Design (ICLAD)	Stanford, California, USA	https://www.islad.org/
06.26. - 06.27.	2025 17th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)	Targoviste, Romania	https://ecai.ro/
06.26. - 06.27.	2025 International Compact Modeling Conference (ICMC)	San Francisco Bay Area, California, USA	https://2025.si2-icmc.org/
06.26. - 06.27.	2025 IEEE Cloud Summit	Washington, District of Columbia, USA	https://www.ieeecloudsummit.org/
06.26. - 06.27.	2025 Symposium on Maritime Informatics and Robotics (MARIS)	Syros, Greece	http://ieee-maris.aegean.gr/about.html
06.26. - 06.27.	2025 International Conference on Emerging Technologies in Computing and Communication (ETCC)	Bangalore, India	https://pesu-etcc2025.vercel.app/
06.27. - 06.29.	2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBE)	Penang, Malaysia	https://www.iccbe.asia/

일 자	학술대회명	개최장소	홈페이지/연락처
06.27. - 06.29.	2025 7th International Conference on Computer Communication and the Internet (ICCCI)	Tokushima, Japan	https://www.iccci.org/
06.27. - 06.29.	2025 IEEE International Symposium on Cyber-Physical Systems and Internet-of-Things: Design and Implementation (CPSIoTDI)	Bilaspur, Chhattisgarh, India	https://cpsi2k24.com/
06.27. - 06.28.	2025 International Conference on Computer Systems and Technologies (CompSysTech)	Ruse, Bulgaria	http://www.compsystech.org/
06.27. - 06.28.	2025 International Conference on Advancements in Power, Communication and Intelligent Systems (APCI)	Kannur, India	https://www.gcek.ac.in/APCI2025/
06.27. - 06.28.	2025 Annual International Conference on Data Science, Machine Learning and Blockchain Technology (AICDMB)	MYSURU, India	http://aicdmb2025.vvce.ac.in/
06.27. - 06.28.	2025 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)	Kuala Lumpur, Malaysia	https://sites.google.com/view/i2cacis/home
06.27. - 06.29.	2025 13th International Conference on Intelligent Computing and Wireless Optical Communications (ICWOC)	Chengdu, China	https://www.icwoc.org/
06.27. - 06.29.	2025 4th International Conference on Image Processing and Media Computing (ICIPMC)	Xi'an, China	https://www.icipmc.net/
06.28. - 06.30.	2025 IEEE 3rd International Conference on Image Processing and Computer Applications (ICIPCA)	Shenyang, China	http://www.icipca.org/
06.29. - 07.03.	2025 23rd International Conference on Solid-State Sensors, Actuators and Microsystems (Transducers)	Orlando, Florida, USA	https://www.transducers2025.org/
06.29. - 07.03.	2025 30th OptoElectronics and Communications Conference (OECC) and 2025 International Conference on Photonics in Switching and Computing (PSC)	Sapporo, Japan	https://www.oeccpsc2025.org/
06.29. - 07.03.	2025 IEEE Kiel PowerTech	Kiel, Germany	https://2025.ieee-powertech.org/
06.30. - 07.04.	2025 IEEE 10th European Symposium on Security and Privacy (EuroS&P)	Venice, Italy	https://eurosp2024.ieee-security.org/
06.30. - 07.05.	2025 International Joint Conference on Neural Networks (IJCNN)	Rome, Italy	https://2025.ijcnn.org/
06.30. - 07.02.	2025 22nd International Conference on Ubiquitous Robots (UR)	College Station, Texas, USA	https://2025.ubiquitousrobots.org/
06.30. - 07.03.	2025 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO)	Tyumen, Russia	http://media-publisher.ru/en/about-synchroinfo-2025/
06.30. - 07.03.	2025 IEEE 19th International Conference on Control & Automation (ICCA)	Tallinn, Estonia	http://www.mae.cuhk.edu.hk/~usr/icca2025/index.html
〉〉2025년 7월			
07.01. - 07.03.	2025 International Conference on Control, Automation and Diagnosis (ICCAD)	Barcelona, Spain	https://www.iccad-conf.com/
07.01. - 07.03.	2025 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0 & IoT)	Castelldefels, Spain	https://www.metroind40iot.org/
07.01. - 07.03.	2025 16th International Conference on Information and Communication Systems (ICICS)	Irbid, Jordan	https://www.just.edu.jo/icics/
07.02. - 07.05.	2025 IEEE Symposium on Computers and Communications (ISCC)	Bologna, Italy	https://ieee-iscc.computer.org/2025/
07.02. - 07.04.	2025 13th International Workshop on Advanced Ground Penetrating Radar (IWAGPR)	Thessaloniki, Greece	https://www.iwagpr2025.gr/
07.03. - 07.06.	2025 5th International Conference on Electrical, Computer and Energy Technologies (ICECET)	Paris, France	https://www.icecet.com/
07.03. - 07.05.	2025 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)	Bali, Indonesia	http://iaict.org/
07.03. - 07.05.	2025 11th International Conference on Smart Computing and Communications (ICSCC)	Kochi, India	http://icscc.net/
07.04. - 07.06.	2025 IEEE North-East India International Energy Conversion Conference and Exhibition (NE-IECCE)	Silchar, India	https://ne-iecce2025.org/
07.04. - 07.05.	2025 International Conference on Computing Technologies & Data Communication (ICCTDC)	HASSAN, India	https://icctdc.com/
07.04. - 07.06.	2025 International Conference on Smart & Sustainable Technology (INCSST)	Chikodi, India	https://incst.org/

일 자	학술대회명	개최장소	홈페이지/연락처
07.04. - 07.05.	2025 International Conference on Information, Implementation, and Innovation in Technology (I2ITCON)	Pune, India	https://commcon.in/
07.04. - 07.06.	2025 4th International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC)	Chengdu, China	https://www.raiic.org/
07.04. - 07.07.	2025 7th International Conference on Power and Energy Technology (ICPET)	Shanghai, China	https://www.icpet.org/
07.04. - 07.05.	2025 International Conference on Engineering Innovations and Technologies (ICoEIT)	Bhopal, India	http://icoeit.org/
07.06. - 07.10.	2025 IEEE International Conference on Fuzzy Systems (FUZZ)	Reims, France	https://fuzzieee2025.conf.lip6.fr/
07.06. - 07.09.	2025 IEEE Computer Society Annual Symposium on VLSI (ISVLSI)	Kalamata, Greece	https://www.ieee-isvlsi.org/ISVLSI_2025_Website/index.html
07.07. - 07.10.	2025 American Control Conference (ACC)	Denver, Colorado, USA	https://acc2025.a2c2.org/
07.07. - 07.09.	2025 IEEE Region 10 Symposium (TENSYMP)	Christchurch, New Zealand	https://attend.ieee.org/tensymp25/
07.07. - 07.12.	2025 IEEE World Congress on Services (SERVICES)	Helsinki, Finland	https://services.conferences.computer.org/2024/
07.07. - 07.09.	2025 IEEE 31st International Symposium on On-Line Testing and Robust System Design (IOLTS)	Ischia, Italy	https://orion.polito.it/iolts/
07.07. - 07.13.	2025 17th International Conference on Advanced Computational Intelligence (ICACI)	Bath, United Kingdom	http://conference.cs.cityu.edu.hk/icaci/
07.07. - 07.11.	2025 Joint Telematics Group/IEEE Information Theory Society Summer School in Information Theory, Signal Processing, Telecommunication, and Networking	Mumbai, India	http://ee.iitb.ac.in/~jtg2025/
07.07. - 07.10.	2025 21st International Conference on Synthesis, Modeling, Analysis and Simulation Methods, and Applications to Circuits Design (SMACD)	Istanbul, Turkiye	https://smacd-conference.org/
07.07. - 07.11.	2025 28th International Conference on Information Fusion (FUSION)	Rio de Janeiro, Brazil	https://fusion2025.org/
07.08. - 07.10.	2025 Signal Processing Symposium (SPSympo)	Warsaw, Poland	https://spsympo.ise.pw.edu.pl/en
07.08. - 07.11.	2025 IEEE World Haptics Conference (WHC)	Suwon, Korea (South)	https://2025.worldhaptics.org/
07.08. - 07.10.	2025 Fifth International Symposium on 3D Power Electronics Integration and Manufacturing (3D-PEIM)	Denver, Colorado, USA	https://www.3d-peim.org/
07.08. - 07.11.	2025 IEEE 49th Annual Computers, Software, and Applications Conference (COMPSAC)	Toronto, Ontario, Canada	https://ieeecompsac.computer.org/2025/
07.08. - 07.10.	2025 38th International Vacuum Nanoelectronics Conference (IVNC)	Reykjavik, Iceland	https://www.vacuumnanoelectronics.org/
07.08. - 07.11.	2025 Sixteenth International Conference on Ubiquitous and Future Networks (ICUFN)	Lisbon, Portugal	https://icufn.org/
07.09. - 07.12.	2025 IEEE 5th International Conference on Sustainable Energy and Future Electric Transportation (SEFET)	Jaipur, India	https://www.sefet.in/
07.10. - 07.18.	2025 IEEE Nuclear and Space Radiation Effects Conference (NSREC)	Nashville, Tennessee, USA	https://www.nsrec.com/
07.10. - 07.13.	2025 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT)	Bengaluru, India	https://ieee-conecct.org/
07.11. - 07.13.	2025 IEEE 26th China Conference on System Simulation Technology and its Applications (CCSSTA)	Shenzhen, China	http://www.ccssta.org.cn/index
07.12. - 07.13.	2025 Intelligent Methods, Systems, and Applications (IMSA)	Giza, Egypt	http://imsa.msa.edu.eg/
07.12. - 07.14.	2025 10th International Conference on Signal and Image Processing (ICSIP)	Wuxi, China	https://www.icsip.org/index.html
07.12. - 07.15.	2025 International Conference on Machine Learning and Cybernetics (ICMLC)	Bali, Indonesia	https://www.icmlc.com/
07.12. - 07.15.	2025 International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR)	Bali, Indonesia	https://www.icwapr.org/
07.13. - 07.18.	2025 IEEE International Symposium on Applications of Ferroelectrics (ISAF)	Graz, Austria	https://2025.ieee-isaf.org/
07.13. - 07.18.	2025 IEEE International Symposium on Antennas and Propagation and North American Radio Science Meeting (AP-S/CNC-USNC-URSI)	Ottawa, Ontario, Canada	https://sites.google.com/view/ieeeap-sursi2025

일 자	학술대회명	개최장소	홈페이지/연락처
07.13. - 07.16.	2025 IEEE 25th International Conference on Nanotechnology (NANO)	Washington, District of Columbia, USA	https://2025.ieeenano.org/
07.13. - 07.18.	2025 IEEE CNC-USNC-URSI North American Radio Science Meeting (Joint with AP-S Symposium)	Ottawa, Ontario, Canada	https://sites.google.com/view/ieeeap-sursi2025
07.14. - 07.18.	2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)	Copenhagen, Denmark	https://embc.embs.org/2025/
07.14. - 07.17.	2025 IEEE International Conference on Advanced Learning Technologies (ICALT)	Changhua, Taiwan	https://tc.computer.org/tclt/icalt-2025/
07.14. - 07.18.	2025 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)	Hangzhou, China	https://www.aim2025.org/
07.15. - 07.18.	2025 IEEE International Conference on Environment and Electrical Engineering and 2025 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe)	Chania, Crete, Greece	https://www.eeeic.net/
07.15. - 07.22	2025 12th International Conference on Machine Intelligence Theory and Applications (MiTA)	Frankfurt, Germany	http://www.mita2025.com/
07.15. - 07.17.	2025 IEEE 8th International Conference on Signal Processing and Machine Learning (SPML)	Hohhot, China	https://www.spml.net/
07.15. - 07.18.	2025 16th International Conference on Mechanical and Aerospace Engineering (ICMAE)	Rome, Italy	https://www.icmae.org/
07.15. - 07.18.	2025 17th Biomedical Engineering International Conference (BMEiCON)	Chiang Mai, Thailand	http://www.bmeicon.org/bmeicon2025/
07.16. - 07.18.	2025 6th International Conference of the Portuguese Society for Engineering Education (CISPEE)	Setúbal, Portugal	https://www.cispee2025.ips.pt/
07.16. - 07.20.	2025 25th International Conference on Software Quality, Reliability and Security (QRS)	Hangzhou, China	https://qrs25.techconf.org/
07.16. - 07.20.	2025 25th International Conference on Software Quality, Reliability, and Security Companion (QRS-C)	Hangzhou, China	https://qrs25.techconf.org/
07.16. - 07.18.	2025 International Conference on Computer, Information and Telecommunication Systems (CITS)	Colmar, France	https://cits.udg.edu/
07.16. - 07.18.	2025 10th International Conference on Image, Vision and Computing (ICIVC)	Chengdu, China	https://www.icivc.org/index.html
07.17. - 07.19.	2025 IEEE International Conference on Advanced Robotics and its Social Impacts (ARSO)	Osaka, Japan	https://ieee-arso.org/
07.17. - 07.18.	2025 International Symposium on Signals, Circuits and Systems (ISSCS)	Iasi, Romania	http://scs.etti.tuiasi.ro/isscs2025/
07.18. - 07.21.	2025 International Conference of Clean Energy and Electrical Engineering (ICCEEE)	Changchun, China	http://www.icceee.com/expire.php
07.18. - 07.19.	2025 International Conference on Computing, Intelligence, and Application (CIACON)	Durgapur, India	https://www.ciacon.in/
07.18. - 07.19.	2025 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)	Ernakulam, India	http://ic7.fisat.ac.in/
07.18. - 07.20.	2025 4th International Conference on Power System and Energy Technology (ICPSET)	Chengdu, China	https://www.icpset.org/
07.19. - 07.21.	2025 International Conference on Advanced Machine Learning and Data Science (AMLDS)	Tokyo, Japan	http://amlds.site/
07.20. - 07.23.	2025 IEEE 20th International Symposium on Antenna Technology and Applied Electromagnetics (ANTEM)	St. John's, Newfoundland and Labrador, Canada	https://antem2025.ieee.ca/
07.20. - 07.23.	2025 IEEE International Professional Communication Conference (ProComm)	Sønderborg, Denmark	https://attend.ieee.org/procomm-2025/
07.20. - 07.22.	2025 IEEE 9th International Test Conference India (ITC India)	Bangalore, India	https://itctestweekindia.org/
07.21. - 07.24.	2025 IEEE 15th International Conference on Power Electronics and Drive Systems (PEDS)	Penang, Malaysia	http://ieee-peds.org/
07.21. - 07.25.	2025 IEEE Symposium on Radiation Measurements and Applications (SORMA West)	Berkeley, California, USA	http://sormawest.org/
07.21. - 07.23.	2025 IEEE 4th German Education Conference (GECon)	Hamburg, Germany	https://attend.ieee.org/gecon-2025/
07.21. - 07.23.	2025 IEEE International Conference on Computational Photography (ICCP)	Toronto, Ontario, Canada	https://iccp2025.iccp-conference.org/

일 자	학술대회명	개최장소	홈페이지/연락처
07.21. - 07.24.	2025 IEEE International Conference on Decentralized Applications and Infrastructures (DAPPS)	Tucson, Arizona, USA	https://ieeedapps.com/
07.21. - 07.24.	2025 IEEE International Congress on Intelligent and Service-Oriented Systems Engineering (CISOSE)	Tucson, Arizona, USA	https://conf.researchr.org/home/cisose-2025
07.21. - 07.23.	2025 IEEE Photonics Society Summer Topicals Meeting Series (SUM)	Berlin, Germany	https://www.ieee-sum.org/
07.21. - 07.24.	2025 IEEE 11th International Conference on Big Data Computing Service and Machine Learning Applications (BigDataService)	Tucson, Arizona, USA	https://conf.researchr.org/track/cisose-2025/bigdataservice-2025
07.21. - 07.23.	2025 IEEE Space, Aerospace and Defence Conference (SPACE)	Bangalore, India	http://www.ieeespace.org/
07.21. - 07.23.	2025 1st International Symposium on E-CARGO and Applications (E-CARGO)	Guangzhou, China	http://e-cargoschool.com/2025#/
07.21. - 07.24.	2025 IEEE Intelligent Mobile Computing (MobileCloud)	Tucson, Arizona, USA	https://conf.researchr.org/track/cisose-2025/imc-2025
07.21. - 07.24.	2025 IEEE International Conference on Artificial Intelligence Testing (AITest)	Tucson, Arizona, USA	https://ieeeaitest.com/
07.22. - 07.25.	2025 International Symposium on Educational Technology (ISET)	Bangkok, Thailand	https://hksmic.org.hk/iset/2025/
07.23. - 07.25.	2025 Conference on Information Communications Technology and Society (ICTAS)	Durban, South Africa	https://www.ictas.org/
07.23. - 07.25.	2025 International Seminar on Intelligent Technology and Its Applications (ISITIA)	Surabaya, Indonesia	https://elib.its.ac.id/conf/isitia/main/
07.25. - 07.26.	2025 International Conference on Innovations in Intelligent Systems: Advancements in Computing, Communication, and Cybersecurity (ISAC3)	Bhubaneswar, India	https://isac3.in/
07.25. - 07.27.	2025 3rd World Conference on Communication & Computing (WCONF)	Raipur, India	https://wconf.in/
07.25. - 07.27.	2025 International Conference on Computational Intelligence and Robotics (CIR)	Guangzhou, China	https://www.ic-cir.com/
07.25. - 07.27.	2025 IEEE 7th International Conference on Artificial Intelligence, Computer Science, and Information Processing (AICSIP)	Hangzhou, China	https://www.aicsconf.cn/
07.25. - 07.26.	2025 3rd International Conference on Data Science and Network Security (ICDSNS)	Tiptur, India	https://icdsns.co.in/
07.25. - 07.27.	2025 6th International Conference on Energy Power and Automation Engineering (ICEPAE)	Zhengzhou, China	https://www.icepae.org/
07.25. - 07.27.	2025 8th Asia Conference on Energy and Electrical Engineering (ACEEE)	Qingdao, China	https://www.aceee.net/
07.26. - 07.27.	2025 IEEE 4th World Conference on Applied Intelligence and Computing (AIC)	GB Nagar, Gwalior, India	http://scrs.in/conference/aic2025
07.27. - 07.31.	2025 IEEE Power & Energy Society General Meeting (PESGM)	Austin, Texas, USA	https://pes-gm.org/
07.27. - 07.30.	2025 16th International Conference on Reliability, Maintainability and Safety (ICRMS)	Shanghai, China	http://www.icrms.cn/#/
07.28. - 08.01.	2025 International Conference on Sampling Theory and Applications (SampTA)	Vienna, Austria	https://sampta25.univie.ac.at/
07.28. - 07.30.	2025 44th Chinese Control Conference (CCC)	Chongqing, China	https://ccc2025.cqu.edu.cn/
07.28. - 08.01.	2025 IEEE International Conference on Manipulation, Manufacturing and Measurement on the Nanoscale (3M-NANO)	Changchun, China	http://www.3m-nano.org/2025/main/index.asp
07.28. - 07.31.	NAECON 2025 - IEEE National Aerospace and Electronics Conference	Dayton, Ohio, USA	https://attend.ieee.org/naecon-2025/
07.28. - 08.01.	2025 IEEE Space Computing Conference (SCC)	Los Angeles, California, USA	https://2025.smcit-scc.space/
07.28. - 08.01.	2025 IEEE 11th International Conference on Space Mission Challenges for Information Technology (SMC-IT)	Los Angeles, California, USA	https://2025.smcit-scc.space/
07.29. - 08.01.	2025 IEEE MTT-S International Conference on Numerical Electromagnetic and Multiphysics Modeling and Optimization (NEMO)	Tianjin, China	http://www.em-conf.com/nemo2025/
07.29. - 07.31.	2025 8th International Conference on Computing Methodologies and Communication (ICCMC)	Erode, India	https://icocmc.com/ICCMC-25/
07.30. - 08.01.	2025 10th International STEM Education Conference (iSTEM-Ed)	Pattaya, Thailand	http://www.istem-ed.com/2025/

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
롯데렌탈(주)	최진환	경기도 안양시 동안구 전파로88 (신원비전타워 8층)	02-3453-8970	https://www.lotterental.com
리얼텍코리아 주식회사	창텅치	서울시 서초구 사임당로 18, 석오빌딩 5층	070-4120-7966	www.realtek.com
비젼테크	이원복	대전 유성구 테크노2로 187, 미건테크노월드2차 1층 118호	042-934-0236	http://www.visiontechkorea.com
머플 주식회사	안수남	경기 성남시 분당구 판교로 289번길 20, 2동		http://www.murple.ai/
삼성전자㈜	한종희	서울시 서초구 서초2동 1320-10 삼성전자빌딩	02-1588-3366	https://www.samsung.com
스카이칩스	이강윤	수원시 장안구 서부로 2066, 산학협력센터 85511호	031-299-6848	http://www.skaichips.co.kr
스테코㈜	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

회원사	대표자	주 소	전 화	홈페이지
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
㈜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	홍세경	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〉 제52권 제2호(통권 제489호) The Magazine of the IEIE

2025년 2월 20일 인쇄
2025년 2월 25일 발행

발행및 편집인 **(사) 대한전자공학회** **회장 백 광 현**
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

2025년도 회비납부 안내

1. 회비의 납부 및 유효기간

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

◆ 2025년도 회원 연회비는 다음과 같습니다.
- 정 회 원 : 70,000원 (입회비 : 10,000원)
- 학생회원 : 30,000원 (입회비 면제)
- 평생회원 : 700,000원
  - 평생회비 할인 제도 : 학회 홈페이지 안내 참조
  - 평생회비 분납 제도(1년 한) : 평생회비 분할 납부를 원하시는 회원께서는 회원 담당에게 요청하여 주시기 바랍니다.
  - 7월 1일부터 연회비 50% 할인 적용

2. 논문지(eBook) 제공

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

3. 회비의 납부방법

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

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

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

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

5. 문의처

◆ 대한전자공학회 사무국 곽새롬 서기(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