

CIDA Lab에 오신 것을 환영합니다

Computational Intelligence & Data Analytics Lab
신입생을 위한 연구실 소개

고상기 (Sang-Ki Ko)

2026

서울시립대학교 인공지능학과

**시작하며: 왜 이 이야기를
듣는가?**

신입생 여러분, 환영합니다

- 오늘 여러분은 **4년**을 함께 보낼 전공의 풍경을 볼 것입니다.
- 인공지능학은 단순히 “모델을 돌리는 학문”이 아닙니다.
- 이 시간 동안 저는 세 가지를 전하고 싶습니다:
 1. 인공지능학이 왜 **재미있는** 학문인지
 2. CIDA Lab에서는 어떤 연구를 하는지
 3. 학부생이 연구실에 어떻게 참여할 수 있는지
- 편하게 들으시고, 궁금한 것은 언제든지 질문해 주세요.

인공지능학은 어디서 왔는가?

- 인공지능학의 뿌리는 **컴퓨터과학(Computer Science)**입니다.
 - 튜링(Turing)의 “기계는 생각할 수 있는가?”(1950)로 시작
 - 알고리즘, 자료구조, 계산이론 — 모두 AI의 토대
- 그러나 인공지능은 **여러 학문의 영향**을 받으며 자라 왔습니다:
 - **수학** — 최적화, 확률론, 선형대수
 - **통계학** — 데이터로부터 배우는 원리
 - **뇌·신경과학** — 뉴럴네트워크의 영감
 - **언어학** — 자연어 처리의 기초
 - **철학·인지과학** — 지능이란 무엇인가
 - **경제학·게임이론** — 강화학습의 원리
 - **제어공학·로보틱스** — 의사결정과 행동
- 그래서 AI 연구자는 **여러 분야를 넘나드는 사고**를 할 수 있어야 합니다.

왜 지금 “연구”를 이야기하는가?

- 많은 분들이 “연구는 3 4학년 때나 할 일”이라고 생각합니다.
- 그러나 현실은:
 - AI 분야의 **일류 논문 저자**에 학부생이 자주 포함됩니다.
 - 대학원 진학, 해외 유학, 탐티어 인턴십 모두 **연구 경험**을 요구합니다.
 - 연구실은 수업에서는 경험할 수 없는 **실제 문제**를 다룹니다.
- 1학년부터 “연구란 무엇인가”를 이해해 두면, 남은 3년을 훨씬 더 전략적으로 설계할 수 있습니다.

이론 (Theory)

- 무엇이 계산 가능한가?
- 어떤 문제가 어려운가?
- 알고리즘의 한계는 어디인가?
- → 오토마타, 복잡도 이론, 논리

응용 (Applications)

- AI로 코드를 고칠 수 있을까?
- 축구 경기에서 “좋은 패스”란?
- 뉴럴넷을 믿을 수 있을까?
- → 딥러닝, 강화학습, LLM

CIDA Lab의 철학

우리는 이 두 세계를 **동시에** 연구합니다. “이론적 깊이”와 “현실 문제”가 만나는 지점에 진짜 흥미로운 문제가 있습니다.

인공지능학 전공의 매력

인공지능은 “보편 도구”이다

- 20세기의 보편 언어가 **수학**이었다면, 21세기는 **데이터와 AI**입니다.
- 생물학자는 유전자를 딥러닝으로 해독하고,
- 경제학자는 시장을 예측 모델로 분석하며,
- 축구 감독은 경기를 AI 지표로 평가합니다.
- ⇒ 인공지능학을 제대로 배우면, **어떤 분야에도** 진입할 수 있습니다.

지난 10년, AI에서 무슨 일이 있었나?

- 2012 — AlexNet, 딥러닝의 부활
- 2016 — AlphaGo, 이세돌 9단을 이기다
- 2017 — Transformer 등장 (“Attention Is All You Need”)
- 2020 — AlphaFold, 단백질 구조 예측의 혁명
- 2022 — ChatGPT, 언어 모델의 대중화
- 2023-현재 — 멀티모달, 에이전트, 코드 생성의 시대

여러분이 졸업할 즈음에는 또 어떤 세계가 펼쳐져 있을까요?

그러나 — “AI 시대”에 공부의 본질은?

- LLM이 코드를 써 주는 시대, 무엇을 공부해야 할까?
- 정답: **도구를 쓰는 사람**이 아닌, **도구를 설계하는 사람**이 되어야 합니다.
- 그러려면:
 - 알고리즘과 자료구조의 원리를 이해하고,
 - 수학적 사고로 문제를 추상화할 수 있으며,
 - 새로운 문제를 직접 정의할 수 있어야 합니다.
- 이것이 바로 “연구”의 훈련입니다.

“좋은 개발자”와 “연구자”의 차이

개발자

- “어떻게 만들지?”
- 명확한 요구사항
- 기존 기술의 숙련된 적용
- 제품의 완성도

연구자

- “이것이 과연 가능한가?”
- 문제 자체가 불분명
- 기존에 없는 방법 제안
- 새로운 지식의 생산

핵심

연구자의 훈련을 받은 사람은 “좋은 개발자”가 되기 쉽지만, 반대는 어렵습니다.

CIDA Lab 소개

Computational Intelligence & Data Analytics Lab

- 서울시립대학교 인공지능학과 소속 연구실
- 이론 전산 + 최첨단 인공지능 응용의 교차점
- 주요 연구 4개 축:
 1. **Theory of Computation** — 계산이론
 2. **Formal Verification of Neural Networks** — 뉴럴넷 검증
 3. **Programming Language Understanding & Generation** — 프로그래밍 언어 분석
 4. **Sports Data Analytics** — 스포츠 데이터 분석

- 연세대학교 컴퓨터과학과 Ph.D. (이론 전산학)
- 한국전자기술연구원(KETI) 선임연구원
- 강원대학교 컴퓨터공학과 조교수
- 현재: 서울시립대 인공지능학과 부교수
- **한국프로축구연맹 K리그 TSG 위원**
(Technical Study Group)
- 관심사: 오토마타, 형식언어, 뉴럴넷 검증, 코드 LLM, 축구 데이터 분석

한마디로

“이론 전산학 전공자가 어쩌다 딥러닝으로 축구를 연구하게 된 이야기”

- 우리는 **특정 기술**을 연구하지 않습니다.
- 우리는 **어려운 문제**를 연구합니다.
- 그 문제가 마침 — 형식언어일 수도, 추구일 수도, 코드 LLM일 수도, SNN 검증일 수도 있습니다.
- **문제가 재미있다면 분야는 상관없다.** 이것이 CIDA Lab의 원칙입니다.

결과

KDD, IJCAI, EMNLP, CIKM, MIT Sloan SAC 등 다양한 **탐티어 학회**에 꾸준히 논문 발표 중.

연구 분야

질문: “이 문제, 기계가 풀 수 있을까? 풀 수 있다면 얼마나 빨리?”

- 튜링(Turing), 처치(Church), 괴델(Gödel)로 거슬러 올라가는 **CS의 뿌리**
- 핵심 도구: 오토마타, 정규표현식, 형식언어
- 겉보기에는 순수 수학이지만 — 컴파일러, 검색엔진, 유전자 분석, XML 파서 등 **모든** 곳에 들어 있습니다.

- **Edit-distance of formal languages**

두 언어 사이의 “거리”를 계산할 수 있을까?

- **Descriptive complexity**

같은 언어를 표현하는 가장 작은 오토마타/정규식은 얼마나 작을까?

- **Regular language inference**

긍정/부정 예시만 보고 정규 언어를 역추론할 수 있을까?

- **Simon's congruence**

두 문자열이 같은 “부분 수열” 패턴을 가지는지 판정하기.

재미있는 응용 예: Regex 문제 자동 채점

- 정규표현식 수업의 과제를 교수가 일일이 채점해야 하는가?
- **아이디어:** 학생 답과 정답 정규식의 언어 동치성을 자동 검사.
- 단순 “일치 테스트”와 달리, 진짜 동등한지를 **수학적으로** 판단합니다.
- ⇒ **ESOP 2023** (프로그래밍 언어 분야 탐티어) 발표.

학부생에게

이산수학, 오토마타 이론, 알고리즘 수업을 좋아했다면 — 당신에게 딱 맞는 분야입니다.

- **Information and Computation 2025** — Alternating Finite Automata의 polynomial width 결정가능성
- **Theoretical Computer Science 2024** — Simon's Congruence Pattern Matching
- **ESOP 2023** — Automated Grading of Regular Expressions
- **JALC 2025** — SplitRegex: 신경망 기반 정규식 합성

왜 “뉴럴넷을 검증”해야 하는가?


- 의료 영상 진단, 자율주행, 금융 거래 — 모두 **안전 중요(safety-critical)** 시스템입니다.
- 그런데 뉴럴넷은 “블랙박스”: 왜 그런 출력을 냈는지 설명이 어렵습니다.
- **질문:** “이 모델은 모든 입력에 대해 안전한가?”
- 테스트로는 불충분 — 모든 입력을 시험할 수 없습니다.
- **해답:** 수학적으로 증명하기 — **Formal Verification.**

- 계산이론에서 오래 쓰여 온 기법:
 - Model checking (모델 체킹)
 - Abstract interpretation (추상 해석)
 - SMT solving (충족 가능성 검사)
- 이것을 **뉴럴넷**에 적용하면?
- ⇒ 특정 입력 범위에서 출력이 안전한지 자동 증명 가능.
- 특히 **Spiking Neural Network (SNN)** — 뇌를 모사한 차세대 뉴럴넷 — 의 검증이 우리의 특화 분야.

AI가 코드를 이해한다는 것

- ChatGPT, Copilot, Claude Code — 이미 우리는 AI로 코딩합니다.
- 그러나 “정말로 잘하는가?”라는 질문이 남습니다:
 - AI가 올바른 코드를 생성하는가?
 - AI가 코드의 복잡도를 예측할 수 있는가?
 - AI가 좋은 테스트 케이스를 만들 수 있는가?
- 우리는 이 모든 질문을 **이론+딥러닝**으로 동시에 공략합니다.

AI 기반 소스코드 분석 및 수정 기술 연구

- 한국연구재단 **기초연구실 (BRL)** 지정 과제
- 연세대학교, 성균관대학교, 서울시립대학교 공동
- 3년간 다수의 탐티어 논문 배출 중
-  프로젝트 웹사이트

연구 1: 코드 복잡도 예측 (CodeComplex)

- 문제: 주어진 Python/Java 코드의 최악시간복잡도 $O(\cdot)$ 를 자동 예측하기
- 기존 방법: 정적 분석은 실제 코드에 적용 어려움
- 우리 방법: 딥러닝 + 오토마타 기반 구조 인코딩
- ⇒ **EMNLP 2025** (자연어처리 탐티어) 발표.

연구 2: 자동 테스트 케이스 생성 (LogiCase)

- 문제: 경쟁 프로그래밍 문제를 자연어 설명만 보고 테스트 케이스를 생성하기
- 핵심 아이디어: 자연어 제약을 **논리적 표현**으로 번역
- 그 다음 — 형식문법 기반으로 실제 입력을 합성
- ⇒ **IJCAI 2025** 발표 (인공지능 탐티어).

“AI 연구자가 왜 축구를?”

- 축구는 **가장 어려운 멀티에이전트 시스템** 중 하나:
 - 22명 + 공 = 23개 에이전트
 - 확률적, 비정형, 목표가 명확하지 않은 행동
 - 학습 데이터는 매우 희소
- 자율주행, 로봇틱스, 군집 지능 — 모두 같은 문제 구조를 가집니다.
- 축구는 이 모든 문제의 **훌륭한 실험대**입니다.
- 동시에 — 실제 프로 구단과의 협업 기회가 열려 있습니다.

Valuing La Pausa

- “속도보다 타이밍이다” — 최적 패스 타이밍의 정량화
- **200편 이상 중 Top 7 Finalist!**
- MIT Sloan Sports Analytics Conference는 스포츠 데이터 분야 **세계 최고 권위 학회**

2년 연속 MIT Sloan 입상 (2025 exPress → 2026 La Pausa)

- **Player valuation** — 선수 한 명의 행동이 팀 승률에 얼마나 기여했는가?
- **Trajectory imputation** — 일부 선수 위치 데이터만 있을 때 나머지를 복원
- **Formation analysis** — 경기 중 포메이션 변화를 자동 탐지 (KDD 2022)
- **Ball trajectory inference** — 공의 궤적을 선수 움직임에서 추정 (KDD 2023)
- **Playing style embedding** — 팀의 “스타일”을 벡터로 표현
- **Strategy optimization** — 강화학습으로 전술 의사결정 최적화

- 프로축구 구단 및 스포츠 데이터 기업과 **실제 경기 데이터**를 기반으로 공동 연구
- 학생들은 논문과 실제 산업 문제를 동시에 다룰 수 있습니다.
- 스포츠 팬이 아니어도 OK — “멀티에이전트 시스템”으로 접근하면 됩니다.
- 현재 4명의 대학원생 + 학부생 2명이 이 주제에 참여 중.

연구실 생활

- 박사과정 1명 — Edge AI, AGI
- 석사과정 8명 — 추구 분석, 강화학습, 형식언어, SNN 검증, 디지털 트윈 등
- 학부 연구원 6명 — 다양한 주제에 참여
- 졸업생 — KETI 연구원, KAIST 인턴, 하나은행 데이터 분석가 등으로 진출

- 자율적인 출퇴근 — 정해진 시간이 아닌, **결과** 중심 문화
- 주간 미팅 — 교수-학생 1:1 진행상황 공유
- 주제별 세미나 — 축구팀, 코드팀, 이론팀 등 소그룹 활동
- 논문 읽기 모임 — 최신 탐티어 논문 함께 공부
- 학회 참가 — 매년 여러 명이 해외 학회 발표

우리가 원하는 것

- 호기심
- 끈기
- 수학에 대한 거부감이 없는 태도
- 영어 논문을 읽는 것을 두려워하지
않음
- 자기주도적 학습

우리가 제공하는 것

- 탐티어 논문 지도
- GPU 서버 및 연구 장비
- 학회 참가 지원
- 산학 협력 기회
- 글로벌 네트워크

논문 성과: 지난 3년간

- IJCAI 2025 — LogiCase (테스트 케이스 생성)
- EMNLP 2025 — CodeComplex (복잡도 예측)
- CIKM 2025 — 축구 궤적 복원
- ECML PKDD 2025 — 궤적 imputation
- MIT Sloan 2025 & 2026 — 2년 연속
- ESOP 2023 — 정규식 자동 채점
- KDD 2022, 2023 — 축구 포메이션, 공 궤적

BK21 최상위 학회(IF: 3, 4) 다수. — 학부생도 공저자로 참여한 논문 다수.

학부생을 위한 가이드

1. 학부 연구생 (Undergraduate Researcher)

- 정식으로 연구실에 소속되어 주제를 갖고 연구
- 보통 3 4학년, 때로는 2학년부터 가능

2. 졸업논문/캡스톤 프로젝트

- 4학년 졸업 작품의 형태
- 결과가 좋으면 학회 논문으로 확장

3. 관심 표명 + 세미나 참여

- 1 2학년이라면 먼저 세미나부터 참여 가능
- 메일로 편하게 문의 주세요.

1 2학년 때 무엇을 준비하면 좋을까?

- **필수 과목을 제대로 공부**
 - 이산수학 / 자료구조 / 알고리즘 / 확률과 통계
 - 선형대수 — 딥러닝의 언어입니다.
- **프로그래밍 기본기**
 - Python, C++ 중 하나는 “편안하게” 쓸 수 있도록
 - Git, 리눅스 환경 익숙해지기
- **영어 논문 읽기에 도전**
 - 처음엔 다 모르겠지만 괜찮습니다. 3편만 읽어 보세요.
- **수학에 대한 긍정적 태도**

“저는 아직 아무것도 몰라요” — 괜찮습니다

- 여러분이 모르는 게 **정상**입니다.
- 중요한 건 “얼마나 아는가”가 아니라 “얼마나 배우고 싶은가”.
- 지금 연구실의 박사/석사생들도 3년 전에는 같은 자리에 있었습니다.
- 한 가지만 약속해 주세요: **지속성**.
 - 3개월 하고 그만두는 것이 아닌, 1년, 2년을 함께 성장하기.

대학원 진학은 어떤 선택인가?

석사 (M.S.)

- 2년 과정
- 1 2편의 학회 논문
- 기업 R&D 부서 진출
- 박사 진학의 디딤돌

박사 (Ph.D.)

- 4 5년 과정
- 국제 탐티어 다수 발표
- 교수/연구원 커리어
- 해외 유학도 선택지

CIDA Lab의 방침

학부 때 먼저 연구실에 와서 해 보고, 잘 맞으면 대학원 진학. 무작정 입학부터 하지 않습니다.

- 대기업 R&D — 삼성, LG, 네이버, 카카오 등
- 국책 연구소 — KETI, ETRI, KISTI 등
- 해외 유학 — 박사 과정
- AI 스타트업 — 연구 기반 창업
- 학계 — 교수/연구원
- 공통점: “연구해 본 사람”으로서의 차별화.

마무리




오늘 기억했으면 하는 3가지

1. 인공지능학은 **이론**과 **응용**이 만나는 학문이다 — 둘 다 공부하세요.
2. 연구는 “3학년부터”가 아니다 — 1학년부터 **의식적으로** 준비할 수 있다.
3. CIDA Lab은 **재미있는 어려운 문제**를 환영합니다 — 분야는 상관없습니다.

우리가 환영하는 “당신”

- 오토마타의 아름다움에 매료된 이론 애호가
- LLM으로 뭔가 근본적인 걸 바꾸고 싶은 시스템 빌더
- 축구 경기를 수학으로 설명하고 싶은 스포츠 덕후
- 뉴럴넷을 “믿을 수 있는” 소프트웨어로 만들고 싶은 안전 중시자
- 아직 자신이 뭘 좋아하는지 모르는 신입생

— 모두 환영합니다.

-  narame7.github.io — 연구실 홈페이지
-  narame7@uos.ac.kr — 교수 이메일
-  Google Scholar — 논문 전체 리스트
- **다음 세미나 일정 등 연구실 게시판 확인**

오늘 이후 가능한 액션

- 메일로 “학부생 연구 참여 문의” 보내기
- 연구실 세미나 방청 요청
- 관심 있는 논문 1편 읽고 질문거리 만들어 오기

감사합니다

질문을 받게됩니다.