

NAVER AI Lab

한국과학기술원 | 황의중*

1. 연구소 소개

본 기관탐방에서는 NAVER AI Lab을 소개한다. NAVER AI Lab은 인공지능 기술 서비스와 제품을 시장에서 선보이는 NAVER CLOVA 산하의 첨단 인공지능 연구에 중점을 둔 연구 기관이다. NAVER AI Lab의 미션은 크게 세 가지이다. 첫째, AI 커뮤니티에 기여하기 위해 영향력 있는 장기 AI 연구를 수행하는 것이다. 둘째, NAVER CLOVA의 글로벌 AI R&D 리더십에 기여하는 것이다. 셋째, NAVER를 포함 NAVER 플랫폼 생태계의 모든 사용자와 플레이어(중소기업, 파트너, 크리에이터)의 지속가능한 성장을 이끄는 것이다. 현재 주된 연구분야는 머신러닝/딥러닝, 컴퓨터 비전, 자연언어처리, 음성인식, 멀티모달 학습, 유제 모델링 및 추천 등이 있다. 최근에는 세계 최초 5G 기반 사람과 로봇이 공존하는 NAVER의 새로운 사옥인 1784 건물에 입주하였다.

2. 주요 연구 개발 내용

NAVER AI Lab의 역할을 NAVER CLOVA와 연계해서 설명한다(본 내용은 NAVER 발표 자료[1]를 참고하였다). NAVER CLOVA는 여러 AI 디바이스를 만들고 서비스를 제공한다. AI 디바이스로 CLOVA Friends, Clock, Lamp 등의 제품이 있다. 서비스로는 NAVER App, CLOVA, 및 Map이 대표적이다. 이러한 제품과 서비스를 가능케 하는 AI 기술로는 CLOVA ChatBot, Speech, Voice, OCR, Face, NSML, Video, NLP 등이 있다. 또한 AI 제품으로 CLOVA AI Contact Center, Document Insight, FaceSign, eKYC, Forecast, Logistics, CareCall, Studio 등이 있다. 마지막으로 B2C 서비스로 CLOVA Note와 Dubbing이 있다. 전체적인 AI 스택을 보면 가장 아래 레벨에는 Large-scale GPU Farm, MLOps, 및 데이터 파이프라인들이

있고 그 위에는 여러 가지 AI 기술들(StarGAN, CutMix, AdamP, ReXNet, PiT, PCME, CALM, ReBias 등)이 구현되어 있다. 이 기술들을 기반으로 Hyperscale AI (HyperCLOVA, Multimodal Backbone)과 Core AI Engine (ASR, TTS, OCR, 이미지/물체/비디오 인식, 시계열 예측)이 만들어지고 마지막으로 가장 상단에는 AI 어플리케이션과 솔루션(API, On-premis, Mobile 혹은 webapps, AI speaker 등)이 탑재된다. 여기서 NAVER AI Lab은 AI 연구, Hyperscale AI, Core AI Engine 레벨에서 핵심 연구 수행 및 실전에서 발생하는 도전적인 문제에 적용하는 역할을 맡고 있다.

NAVER AI Lab은 비전, 자연언어처리, 이미지 생성, 인간-컴퓨터 상호작용, 의료 인공지능 등의 분야에서 연구를 수행하고 20명 내외의 인공지능 연구자(인턴을 포함하면 30명 이상)들이 있다. 최근 몇 년간의 연구실적이 비약적으로 발전을 해왔다. 작년 2021년의 경우 NAVER CLOVA에서 인공지능 최고권위 학회 (NeurIPS, ICML, ICLR, CVPR, ICCV, ECCV, ACL, EMNLP, AAAI, ICASSP, Interspeech 등)에 66편의 논문을 게재하였고 이 중 59%인 39편이 NAVER AI Lab에서 나왔다. 전체 NAVER에서 나온 논문 중에 35%가 NAVER AI Lab 논문들이다. 이 논문들의 40%는 실제 인공지능 서비스에도 적용이 되었다. 올해에는 NAVER CLOVA에서 이보다 더 많은 88편의 논문이 게재되었고 그중 43편이 NAVER AI Lab에서 나왔다. 또한 48%의 논문 기법들이 실제 인공지능 서비스에 사용되었다. 논문뿐만 아니라 매년 여러 학회에서 활약하고 있다. 2021년에는 NeurIPS ImageNet PPF WS, NeurIPS Social: ML in Korea, ICLR Social: ML in Korea 등의 행사가 있었고 2022년에는 ICLR Social: ML in Korea, HyperscaleFAccT @ FAccT 2022, Shortcut learning Tutorial @ FAccT 2022 등이 있었다.

다음 절부터는 NAVER AI Lab의 주요 연구 분야들을 간단히 소개한다.

*swhang@kaist.ac.kr

2.1. 머신러닝/인공지능 연구

먼저 펀더멘탈한 머신러닝 연구 및 머신러닝 최적화 연구가 활발히 수행되고 있다. CutMix[2]는 훈련 이미지들 간에 중요한 부분들을 섞음으로써 데이터 증강 및 정규화 효과를 내는 많이 인용되는 기법이다. ReLabel[3]은 ImageNet 같은 데이터셋에서 이미지 당 여러 가지 클래스가 있음에도 불구하고 한가지 클래스로만 라벨링이 되었다는 문제점을 지적하고 외부 데이터에서 학습된 모델을 이용해서 여러 클래스 라벨링을 가능하게 하는 기법이다. AdamP[4]는 딥러닝에서 많이 사용되는 배치 정규화 기법이 Adam같은 기존 모델의 옵티마이저와 결합이 될 경우 수렴 속도가 지나치게 빠른 문제를 해결한다.

둘째, 비전의 백본에 대한 연구도 수행 중이다. ViDT[5]는 최근에 사용되는 비전과 디텍션 트랜스포머 기법들을 통합하여 효율적이고 효과적인 객체 탐지 기법이다. ReXNet[6]은 가벼운 비전 모델을 제한된 연산량으로 정확하게 학습시키는 채널 조절 기법이다.

셋째, 비전과 언어를 같이 학습하는 멀티모달 연구도 진행되고 있다. PCME[7]는 비전과 언어가 같이 다루어지는 멀티모달 세팅에서 임베딩을 생성할 때 확률적 임베딩을 제안함으로써 같은 이미지(캡션)가 여러 캡션(이미지)와 연계가 되는 것을 가능하게 한다. ViLT[8]는 기존 비전-언어 사전 훈련 모델에서 정확도의 저하 없이 콘볼루션 신경망을 제거함으로써 속도를 대폭 개선하는 기법이다.

마지막으로 신뢰 가능한 인공지능 연구가 진행되고 있다. ReBias[9]는 편향된 데이터에서의 모델 학습을 방지하기 위해 인위적인 편향성이 들어간 데이터와 다른 데이터에서 학습하는 기법이다. CGL[10]은 공정한 학습에서 성별 같은 그룹 정보가 부분적으로만 존재할 때 그것을 예측하는 모델을 따로 훈련시키고 신뢰도 값을 이용한 공정한 학습 기법을 제시한다. CALM[11]은 비전 모델을 설명할 때 이미지상의 위치까지 명확히 지목하도록 모델 구조를 개선해서 학습하는 기법이다.

2.2. 생성 연구

생성 연구로는 이미지 생성 모델, 멀티모달 생성, 콘텐츠 생성 등의 연구가 수행 중이다. StarGAN[12]은 여러 도메인 간 이미지 변환을 가능하게 하는 잘 알려진 적대적 생성 신경망이다. StarGAN v2[13]는 기존 StarGAN을 발전시켜서 여러 가지 스타일로의 이미지 변환을 가능하게 한다. StyleMapGAN[14]은

실제 이미지를 잠재 공간에 투사할 때 공간 차원까지 가지게 함으로써 기존보다 더욱 정확한 임베딩을 생성하는 기법이다. C3GAN[15]은 세밀한 클러스터링을 위해 클러스터의 경계가 명확한 것을 장려하는 GAN 학습 기법이다. FSMR[16]은 GAN을 이용한 이미지 생성 시 질감이나 색상 같은 스타일에 편향되지 않게 하는 정규화 기법을 제안한다. GGDR[17]은 조건 없는 생성모델 훈련 시 판별자가 생성자의 이미지를 가짜 이미지인지 판별하는 것과 함께 이미지의 의미론적인 특징까지도 알아내는 세밀한 훈련 기법이다.

2.3. 언어 연구

자연언어처리 분야에서도 활발한 연구가 수행 중이고 하이퍼스케일 인공지능, 언어 표현, 인공지능 윤리 등이 연구되고 있다. SOM-DST[18]는 대화의 상황을 추적할 때 이전의 대화를 메모리에 저장한 상태에서 선택적으로 삭제하는 기법이다. SG-OPT[19]는 BERT를 기반으로 문장 임베딩을 정확하게 생성할 수 있는 대조적 학습 기법이다. HyperCLOVA[20]는 국내 최초의 초대규모 한국어 모델로 AI 노코드 플랫폼인 클로바 스튜디오에 튜닝 기능에 활용이 되는 등 사용자 목적에 최적화가 가능하다.

2.4. 인간-컴퓨터 상호작용 연구

먼저 큰 언어모델을 활용하는 유저 인터페이스 연구가 수행 중이다. TaleBrush[21]는 GPT-3같은 언어모델을 이용해서 글을 쓸 때 단순히 다음 문장을 생성하는 것뿐만 아니라 전체적인 흐름을 선으로 표현함으로써 원하는 스토리를 유도할 수 있게 한다. 또한 머신러닝을 이용해서 사람의 행동을 이해하고 보완하는 연구도 수행 중인데 Policy Modulation[22]에서는 여러 인터페이스에 대해서 사용자의 행동을 시뮬레이션하고 평가할 때 파라미터가 변해도 업데이트가 빠른 행동 정책을 이용함으로써 사용자 모델링을 빠르게 할 수 있다.

3. NAVER 신사옥 건물 (1784)

이번 탐방에서는 성남에 있는 NAVER의 신사옥 건물인 1784도 견학하였다. 이 건물은 새로 지은 세계 최초 5G 기반 사람과 로봇이 공존하는 스마트 시티 건물이다(그림 1a). 1784라는 숫자는 최초의 산업혁명이 시작된 해이기도 하고 건물 자체가 큰 테스트베드의 컨셉을 가진다고 한다. 바로 앞에 있는 그린 팩토리 건물에서 이 건물로 최근 여러 팀이 옮겨왔다고 한다. 때마침 방문한 날은 NAVER 지식인 20주년 기

넘이라 무료 아이스크림을 나누어주고 있었고 많은 직원들로 활기가 넘쳤다. 이 건물에서 가장 인상적인 것은 NAVER AI 기술을 적용할 수 있는 데는 모두 적합한 건물이라는 점이다. 1~2층은 일반인에게 공개가 되어 있는데 AI 로봇들이 사람처럼 펜을 들고 그림을 그리고 있었고(그림 1b) 곳곳에 배달 로봇들이 지나다녔는데 음식, 음료수 등을 건물 어느 곳이든 배달하고 있었다(그림 1c). 건물을 디자인할 때부터 로봇이 잘 지나다닐 수 있게 하였다고 한다. 미래에 사람과 로봇이 공존하는 모습을 미리 보는 느낌이었다.

3층부터는 직원들을 위한 구역이라 NAVER CLOVA 기술로 만든 안면 인식 장치를 통과하였다(그림 1d). 또한 엘리베이터도 AI 알고리즘으로 작동하고 있었고 화면에 원하는 층을 누르면 이동이 가장 빠른 엘리베이터를 추천하고 있었다(그림 1e). 직원들은 5G 기술로 카드나 현금이 전혀 필요 없이 핸드폰만 있으면 모든 활동이 가능했다. 예를 들어 핸드폰으로 건물 식

당의 음식이나 음료를 주문하고 창구에 가서 가져오는 식이었다. 또한 건물 내 주차된 차의 위치도 자동으로 알려주는 서비스가 있었다. 건물 식당에 여러 음식 업체들이 있었는데 그중에는 AI 로봇으로 음식을 만드는 NAVER 내부의 스타트업도 있는 것이 흥미로웠다(그림 1f). 로봇들의 충전소도 있었는데 로봇들이 필요하면 스스로 충전을 하는 모습을 볼 수 있었다(그림 1g). 다른 공간에서는 NAVER 내의 여러 스타트업들을 위한 공간도 있었고(그림 1h) 의료시설, 헬스장, 카페, 매점, 꽃집도 있는 등 직원들의 복지가 인상적이었다(그림 1i).

NAVER-KAIST 초창의적 AI 연구센터 및 NAVER AI Lab도 방문하였다. 초창의적 AI 연구센터는 지난해 5월에 AI 분야 연구 발전과 글로벌 AI 리더십 확보를 위해서 NAVER와 KAIST가 설립을 하였고 KAIST 교수 10여명과 KAIST의 AI 연구원 100명이 참여하는 대규모 산학협력이 진행되고 있었다(그림 1j). NAVER



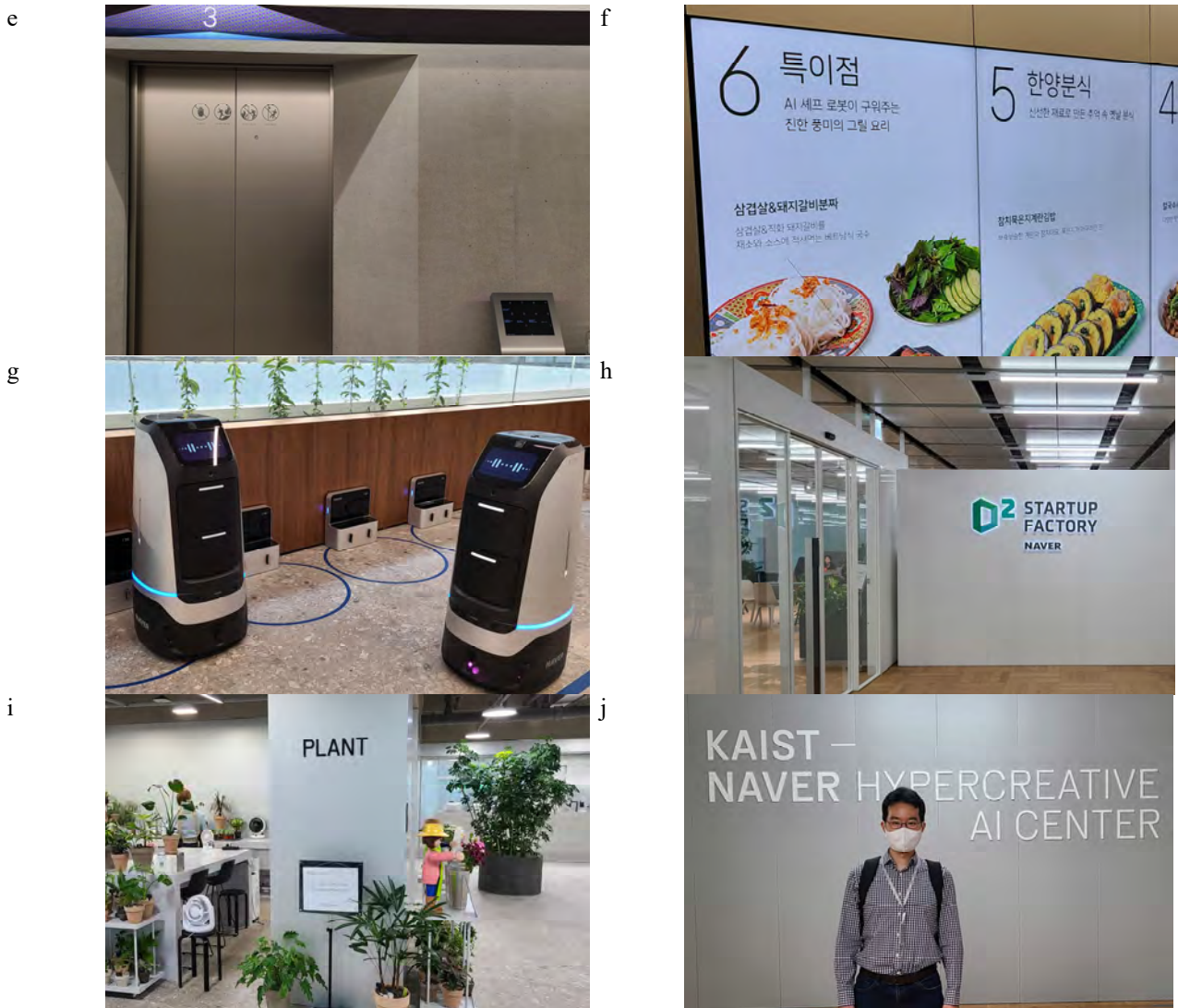


그림 1 NAVER 신사옥 1784 건물의 내부

AI Lab에는 연구원들을 위한 쾌적한 공간과 함께 로봇들이 어김없이 지나가고 있었다. 연구원의 생활에 대해서도 들을 수 있었다. NAVER AI Lab은 한국 뿐만 아니라 전세계적으로도 보기 드문 연구소로서 연구뿐만 아니라 실질적인 AI 제품도 염두에 둔다는 것이 중요한 특징이다. 따라서 연구를 수행할 때 논문을 폭넓게 읽게 되고 서로 협력하기 위해 같이 스터디하는 것이 활성화 되어 있다. 우수한 인턴들과의 연구 협력도 활발하다.

4. 탐방을 마치며

NAVER AI Lab은 한국을 대표하는 소프트웨어 기업인 NAVER의 핵심 AI 연구소이다. 세계적으로 뛰어난 연구실적을 내고 있고 AI 제품과 서비스를 통해

실제 임팩트도 미치고 있다. 이번 방문을 통해서 NAVER AI Lab의 연구가 어떻게 수행되는지 더욱 잘 이해할 수 있었고 특히 1784 건물에서 AI 기술들이 실현되는 모습들이 인상적이었다. 투어를 허락해 주시고 많은 시간을 내주신 송환준 박사님, 이화란 박사님, 그리고 하정우 소장님께 깊은 감사의 말씀을 드린다.

참고문헌

- [1] Sangdoo Yun, NAVER AI Lab @NAVER CLOVA 발표자료, 2022.
- [2] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe, Youngjoon Yoo, “CutMix: Regularization Strategy to Train Strong Classifiers With Localizable Features”, ICCV 2019.

-
- [3] Sangdoo Yun, Seong Joon Oh, Byeongho Heo, Dongyoon Han, Junsuk Choe, Sanghyuk Chun, “Re-Labeling ImageNet: From Single to Multi-Labels, From Global to Localized Labels”, CVPR 2021.
 - [4] Byeongho Heo, Sanghyuk Chun, Seong Joon Oh, Dongyoon Han, Sangdoo Yun, Gyuwan Kim, Youngjung Uh, Jung-Woo Ha, “AdamP: Slowing Down the Slowdown for Momentum Optimizers on Scale-invariant Weights”, ICLR 2021.
 - [5] Hwanjun Song, Deqing Sun, Sanghyuk Chun, Varun Jampani, Dongyoon Han, Byeongho Heo, Wonjae Kim, Ming-Hsuan Yang, “ViDT: An Efficient and Effective Fully Transformer-based Object Detector”, ICLR 2022.
 - [6] Dongyoon Han, Sangdoo Yun, Byeongho Heo, YoungJoon Yoo, “Rethinking Channel Dimensions for Efficient Model Design”, CVPR 2021.
 - [7] Sanghyuk Chun, Seong Joon Oh, Rafael Sampaio de Rezende, Yannis Kalantidis, Diane Larlus, “Probabilistic Embeddings for Cross-Modal Retrieval”, CVPR 2021.
 - [8] Wonjae Kim, Bokyung Son, Ildoo Kim, “ViLT: Vision-and-Language Transformer Without Convolution or Region Supervision”, ICLR 2021.
 - [9] Hyojin Bahng, Sanghyuk Chun, Sangdoo Yun, Jaegul Choo, Seong Joon Oh, “Learning De-biased Representations with Biased Representations”, ICML 2020.
 - [10] Sangwon Jung, Sanghyuk Chun, Taesup Moon, “Learning Fair Classifiers With Partially Annotated Group Labels”, CVPR 2022.
 - [11] Jae Myung Kim, Junsuk Choe, Zeynep Akata, Seong Joon Oh, “Keep CALM and Improve Visual Feature Attribution”, ICCV 2021.
 - [12] Yunjey Choi, Minje Choi, Munyoung Kim, Jung-Woo Ha, Sunghun Kim, Jaegul Choo, “StarGAN: Unified Generative Adversarial Networks for Multi-Domain Image-to-Image Translation”, CVPR 2018.
 - [13] Yunjey Choi, Youngjung Uh, Jaejun Yoo, Jung-Woo Ha, “StarGAN v2: Diverse Image Synthesis for Multiple Domains”, CVPR 2020.
 - [14] Hyunsu Kim, Yunjey Choi, Junho Kim, Sungjoo Yoo, Youngjung Uh, “Exploiting Spatial Dimensions of Latent in GAN for Real-Time Image Editing”, CVPR 2021.
 - [15] Yunji Kim, Jung-Woo Ha, “Contrastive Fine-grained Class Clustering via Generative Adversarial Networks”, ICLR 2022.
 - [16] Junho Kim, Yunjey Choi, Youngjung Uh, “Feature Statistics Mixing Regularization for Generative Adversarial Networks”, CVPR 2022.
 - [17] Gayoung Lee, Hyunsu Kim, Junho Kim, Seonghyeon Kim, Jung-Woo Ha, Yunjey Choi, “Generator Knows What Discriminator Should Learn in Unconditional GANs”, ECCV 2022.
 - [18] Sungdong Kim, Sohee Yang, Gyuwan Kim, Sang-Woo Lee, “Efficient Dialogue State Tracking by Selectively Overwriting Memory”, ACL 2020.
 - [19] Taeuk Kim, Kang Min Yoo, Sang-goo Lee, “Self-Guided Contrastive Learning for BERT Sentence Representations”, ACL 2021.
 - [20] Boseop Kim et al., “What Changes Can Large-scale Language Models Bring? Intensive Study on HyperCLOVA: Billions-scale Korean Generative Pretrained Transformers”, EMNLP 2021.
 - [21] John Joon Young Chung, Wooseok Kim, Kang Min Yoo, Hwaran Lee, Eytan Adar, Minsuk Chang, “TaleBrush: Sketching Stories with Generative Pretrained Language Models”, CHI 2022.
 - [22] Hee-Seung Moon, Seungwon Do, Wonjae Kim, Jiwon Seo, Minsuk Chang, Byungjoo Lee, “Speeding up Inference with User Simulators through Policy Modulation”, CHI 2022.
-