

[인공지능혁신융합대학사업단]

2025년 1학기 캡스톤디자인 산업체연계형 주제 목록

번호	프로젝트 주제명	멘토
1	AI 아트워크 자동 생성 & 스타일 전환 시스템 개발	고두현
2	온디바이스 LLM기반 모바일 작업로봇의 직관적 연동 인터페이스 개발	류지형
3	다수 카메라를 활용한 사용자의 위치 및 동작 추정	
4	생성형 AI를 적용한 디지털 액자	박병운
5	딸기 재배시 발행하는 다양한 작업에 대한 알림 고지	주재영
6	토마토 농약 살포를 위한 영상처리 알고리즘 개발	
7	범용카메라를 활용한 안면감지(Face Recognition)와 시선추적(Eye Tracking)이 가능한 온디바이스 AI 모델 개발	황재준

캡스톤디자인 산업체연계형 주제 #1

1. 주제 (title)
AI 아트워크 자동 생성 & 스타일 전환 시스템 개발
2. 목표 (Goal)
(1차) - Kohya_ss, ComfyUI 활용하여 LoRA(소규모 모델 추가 학습) 모델을 학습하고, 최적화된 모델 생성 - 학습된 LoRA 모델을 포함한 워크플로우(Workflow) 설계 및 이미지(비디오) 생성 테스트 - REST API를 활용한 서비스 구축, 다양한 AI 이미지 생성 및 변환 서비스 개발(자유주제)
(2차) - 사용자가 직접 LoRA 모델을 학습할 수 있는 프로세스 설계 및 실행 - ComfyUI 기반으로 LoRA 모델을 적용한 이미지 생성/변환 자동화 서비스 구축 및 모델 경량화 - REST API 연동을 통한 외부 애플리케이션 AI 생성서비스 구현
3. 접근방법 (Approaches)
- Diffusion Transformer model GUI Framework(Kohya_ss, ComfyUI 등)를 활용한 LoRA 학습 · SDXL 및 기타 이미지 생성 모델을 기반으로 특정 스타일을 학습 · 학습된 LoRA 모델을 다양한 스타일 변환 및 이미지 생성에 활용 · 저작권이 해소된 데이터를 기반으로 AI 학습 진행 (예: 고희, 모네, 일본 전통화 등) - ComfyUI 기반 AI 모델 서비스화 · ComfyUI의 노드 기반 워크플로우 설계 · 워크플로우를 JSON으로 추출하여 다양한 서비스에서 활용 가능하도록 개발 · REST API를 활용한 모델 호출 및 결과 반환 시스템 구축 - REST API를 통한 서비스 개발 · 학습된 LoRA 모델을 호출하여 이미지 생성 및 변환 API 제공 · API 호출을 통해 웹사이트, 모바일 앱, 외부 애플리케이션과 연동 가능하도록 지원
4. 테스트 방법 (Test)
- LoRA 학습 성능 평가 (사용성 평가, Inception 및 FID Score) · 동일한 원본 이미지에 대해 학습 전/후 비교 및 품질 테스트 · 특정 스타일(예: 고희 스타일) 적용 정확도 비교(인물의 경우: Inception Score 적용) - ComfyUI 기반 이미지 생성 테스트 · JSON 기반 워크플로우로 변환된 AI 모델을 실행하여 결과물 비교 · API 호출을 통한 이미지 생성 속도 및 성능 테스트 - REST API 연동 테스트 · FastAPI 서버를 통해 AI 모델을 호출하고, JSON 데이터로 응답을 반환 · API 응답 속도 및 이미지 품질 테스트
5. 주요 자원 (Resources)
- 이미지(비디오) 생성모델 구축 및 학습 가이드라인 제공 - 저작권 해소된 학습용 데이터셋 제공 (반고흐, 고갱, 모네, 뭉크, 드가 등) - ComfyUI Workflow 데이터를 활용한 Rest API 제작 튜토리얼 제공
6. 기타사항
- SDXL Document : https://platform.stability.ai/docs/api-reference#tag/SDXL-and-SD1.6 - ComfyUI Git : https://github.com/comfyanonymous/ComfyUI - Stable Diffusion WebUI : https://github.com/AUTOMATIC1111/stable-diffusion-webui

캡스톤디자인 산업체연계형 주제 #2

1. 주제 (title)
온디바이스 LLM기반 모바일 작업로봇의 직관적 연동 인터페이스 개발
2. 목표 (Goal)
(1차) - LLM(Large Language Model)의 온디바이스 이식 - 모바일 작업로봇의 명령어와 LLM 사이 REST API활용 연동 튜닝
(2차) - 실제 로봇 또는 가상환경의 로봇과 온디바이스 LLM 연동 개발 - 다양한 환경에서 모바일 작업로봇 Test
3. 접근방법 (Approaches)
○ 주요 개발내용 - Nvidia Xavier/Orin 보드에 적합한 온디바이스 LLM 선정(Mini-GPT, LLAMA2 등) - Nvidia 보드에 LLM 적용 및 질의 응답 Test - LLM과 제작할 모바일 작업로봇간 API 설계 및 연동 - 모바일 작업로봇 제작(이동형 로봇 + 매니퓰레이터) 또는 가상환경 세팅 - 모바일 작업로봇의 Test 또는 가상환경에서의 모바일 작업로봇 제작 - 실제 로봇 또는 가상환경의 로봇과 온디바이스 LLM 연동 개발 - 다양한 환경에서 모바일 작업로봇 Test 및 성능 측정
4. 테스트 방법 (Test)
- 실제환경/가상환경에서 모바일 작업로봇에게 명령을 내리고 의도한 대로 동작하는지 분석 - 분석한 데이터를 토대로 LLM활용 재학습 또는 연동 디버깅 - 사용자의 다양한 요청사항에 대해 신뢰도를 측정
5. 주요 자원 (Resources)
- 작업용 모바일 로봇(이동 + 매니퓰레이터) - 온 디바이스 LLM 구현을 위한 Jetson Xavier NX보드
6. 기타사항
- https://say-can.github.io/ - https://sites.research.google/palm-saycan - https://llama.meta.com/ - https://emanual.robotis.com/docs/kr/platform/turtlebot3/overview/ - J. Wei et al., "Chain-of-thought prompting elicits reasoning in large language models," in Advances in Neural Information Processing Systems, vol. 35, 2022, pp. 24824-24837. - A. Brohan et al., "Rt-2: Vision-language-action models transfer web knowledge to robotic control," arXiv preprint, CoRR, 2023, arXiv: 2307.15818.



캡스톤디자인 산업체연계형 주제 #3

1. 주제 (title)
다수 카메라를 활용한 사용자의 위치 및 동작 추정
2. 목표 (Goal)
(1차) - 다수의 카메라 연동을 위한 설치 및 캘리브레이션 - 다수의 카메라에서 객체(사람)인식 및 사람의 위치 추정
(2차) - 다수의 카메라 연동 사람의 포즈 추정 - Web기반 사용자 인터페이스 개발
3. 접근방법 (Approaches)
○ 주요 개발내용 - 3대 이상의 다수의 카메라를 비정형 장소에 설치 및 캘리브레이션 - 3대의 카메라 각각에서 객체(사람)인식 - 인식된 사람의 바운딩 박스 활용 전역좌표계상 사람의 위치 추정 - 다수의 카메라에서 여러사람들의 포즈 추정 - 위치 및 포즈에 따라 관찰자에게 알람을 줄 수 있는 Web기반 사용자 인터페이스 개발
4. 테스트 방법 (Test)
- 다수의 카메라에서 객체(사람)의 인식 결과 분석 - 다수의 카메라에서 사람의 포즈 추정 결과 분석 - 사용자 인터페이스(Web)상에서의 결과값 확인
5. 주요 자원 (Resources)
- 위치 및 포즈 추정을 위한 3대의 스테레오 카메라 - Jetson Xavier NX 보드
6. 기타사항
- Montero, D., Aranjuelo, N., Leskovsky, P. et al. "Multi-camera BEV video-surveillance system for efficient monitoring of social distancing.", Multimed Tools Appl 82, 34995–35019 (2023). https://doi.org/10.1007/s11042-023-14416-y - Human Pose Estimation with Deep Learning – Ultimate Overview in 2024, https://viso.ai/deep-learning/pose-estimation-ultimate-overview/ - https://www.intelrealsense.com/ - https://github.com/IntelRealSense/librealsense

캡스톤디자인 산업체연계형 주제 #4

1. 주제 (title)

생성형 AI를 적용한 디지털 액자

2. 목표 (Goal)

(1차) 자연어처리(인공지능) 프로젝트를 시작하고, 운용하고, 관리하는 방법 학습

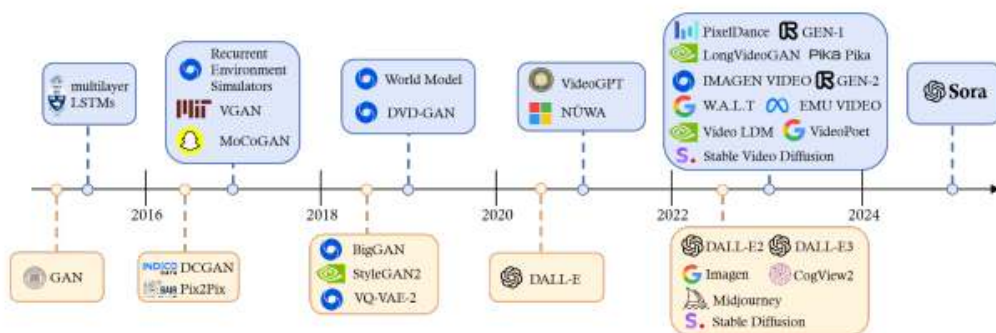
- attention is all you need
- pytorch-lightning과 wandb로 깔끔하게 구현해보는 트랜스포머
- Ankur A. Patel 및 Ajay Uppili Arasanipalai의 O'Reilly Media_Applied Natural Language Processing in the Enterprise

(2차) LLM 모델을 활용한 알고리즘 개발

- ChatGPT와 같은 LLM에 대한 심층 분석
- GPT-2(124M)를 구현 해 보기
- 전시장 부스에서 활용할 수 있는 범용 콘텐츠 서비스 개발

3. 접근방법 (Approaches)

- 랭체인(Langchain)으로 Naive RAG 구현하기
- 프롬프트 엔지니어링 쉽게 접근해 보기
- 고객 서비스를 위한 디지털 휴먼 NVIDIA AI Blueprint는 ACE 기술을 기반으로 하는 워크플로인 NVIDIA Tokkio로 구동되어 3D 애니메이션 디지털 휴먼 인터페이스로 엔터프라이즈 애플리케이션을 생생하게 구현합니다. 친근하고 인간적인 상호 작용을 통해 고객 서비스 애플리케이션은 기존 고객 서비스 옵션에 비해 더욱 매력적인 사용자 경험을 제공할 수 있습니다. 이 워크플로는 RAG를 사용하여 구축된 기존 생성 AI 애플리케이션에 통합되도록 설계되었습니다. 이 워크플로를 사용하여 데이터 센터, 클라우드 또는 에지에서 실행되는 애플리케이션을 진화시켜 완전한 디지털 휴먼 인터페이스를 포함시킬 수 있습니다.



(<https://arxiv.org/pdf/2402.17177>)

4. 테스트 방법 (Test)

- 디지털 액자 아키텍트 구성 디자인 및 설계
- (선택1)디스플레이 장치 + 라즈베리 파이 + 우분투 + 생성형 AI
- (선택2)클라우드 플랫폼(매직인포+PostgreSQL or Mariadb or SQL Server 2022) + 스마트 사이니지 디스플레이
- (선택3)아이패드

5. 주요 자원 (Resources)

- IDE(VS Code, Pycharm등)에서 Copilot에 의한 코드 자동 생성 기술 지원
- PyTorch, Tensorflow 딥러닝 프레임워크 예제, 튜토리얼 교육
- AWS, GCP 클라우드 서비스 데이터센터 GPU 활용 교육
- 깃허브를 통한 개인 및 팀 과제를 통한 협력 체계 구축 방안 제시
- 노션을 통한 기술 보고서 공유 활성화 제안
- 생성형 AI, LLM 자연어 처리 교육

마이크로소프트와 AWS에서 상당한 크레딧을 제공받고 있기 때문에 좋은 GPU 서버를 사용
NVIDIA의 관련 있는 많은 인공지능 제품과 소프트웨어

GitHub 마켓플레이스 모델

GitHub 에듀케이션

코랩 Colaboratory

6. 기타사항

- nvidia/Build a Digital Human

산업 전반의 고객 서비스를 위한 지능적이고 대화형 아바타를 만듭니다.

<https://build.nvidia.com/nvidia/digital-humans-for-customer-service/blueprintcard>

- Intro to Large Language Models

<https://github.com/karpathy/build-nanogpt>

- State of GPT

<https://youtu.be/bZQun8Y4L2A?si=ob6kSmSCDgs24v6Z>

- 메타의 인공지능 리서치

<https://ai.meta.com/>

- AWS

<https://aws.amazon.com/ko/>

- 구글의 인공지능 리서치

<https://cloud.google.com/ai/generative-ai>

- 답시크 '량원평 매직' 비결은?

<http://www.hellodd.com/news/articleView.html?idxno=106873>

캡스톤디자인 산업체연계형 주제 #5

1. 주제 (title)

딸기 재배시 발행하는 다양한 작업에 대한 알림 고지

2. 목표 (Goal)

(1차) 딸기 재배시 다양한 수작업이 필요한데 이를 고지 하기 위한 프로그램 개발

- 딸기의 재배시 발생하는 고품광을 인식하여 고지함
- 레일 타입의 카메라 영상 취득을 통한 자동 동영상 촬영과 알림

(2차) 딸기 생육에 필요한 다양한 모니터링 시스템 데이터를 취득하여, 최적의 생육조건을 확보

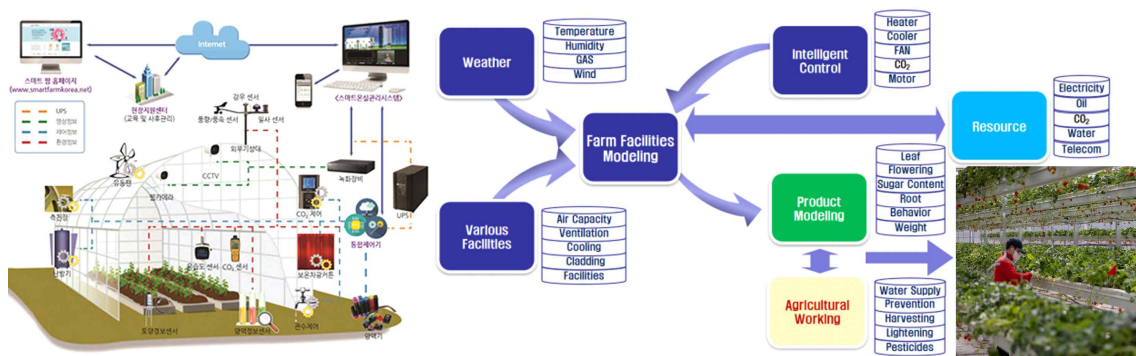
- 딸기의 생육에 따른 화방에 개수를 파악하고, 수정이 된 딸기를 파악하여 생산량을 도출
- 양액 및 조도 등의 데이터를 확보하여 최적의 생육조건을 확보

3. 접근방법 (Approaches)

- 딸기의 재배 최적화를 위한 다양한 문헌 자료 검토
- 영상 기반의 딸기 재배 및 고품광이 파악 분류
- 실내 위치데이터 센서를 통한 문제의 딸기재배 재역의 도출 및 농민 고지
- 다양한 영상처리 기법을 통한 딸기 생육시 발생하는 문제에 대한 영상으로 학습데이터 구축
- 딸기 생산 현장에서의 직접 데이터 수집 및 분류
- 딸기 재배 현장에서 발행하는 다양한 문제들을 농가로부터 배워해 해당 사진들을 취득하고, 이를 학습하고 동일한 영상이 나올 때 고품광이 및 잎따기는 다양한 작업을 지지할 수 있도록 분류함
- 분류된 내용을 토대 사람의 작업이 필요한 영상을 추출하여 알림을 제공함

4. 테스트 방법 (Test)

- 개발된 모델을 다른 딸기 재배 하우스에 적용하여 성능 테스트
- 양액, 일조량, 내부온도, 풍량 등에 대한 데이터를 취득하고 해당 내용을 토대로 일반 하우스형 딸기 재배농가에서 보다 높은 생산성을 얻기 위한 개선방안 마련



(그림) 딸기 성장 최적화를 위한 인공지능 모델 개발 프로세스

5. 주요 자원 (Resources)

- 딸기 하우스 재배지를 방문해서 환경데이터를 수집하기 위한 센서를 부착하고, 해당 데이터를 불러들이고 매주 수확량 데이터를 농가로부터 받아서 최적화 조건을 도출함
- 각 작업로 사이에 카메라를 설치하여 이동하면서 주기적으로 영상을 촬영함

6. 기타사항

- <https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE08761977>

캡스톤디자인 산업체연계형 주제 #6

1. 주제 (title)

토마토 농약 살포를 위한 영상처리 알고리즘 개발

2. 목표 (Goal)

(1차) 토마토 농약 살포를 위한 장애물 파악

- 토마토 재배시 농약을 살포하는데 다양한 장애물이 있어, 장애물의 거리 및 위치를 인식시켜 농약 살포시 해당 장애물을 피해서 살포하도록 함
- 앞의 뒷면(기공면) 살포를 위한 앞의 위치 파악 및 살포 시간 조절

(2차) 적절 분사를 위한 가지 및 잎의 위치 확인

- 단안 카메라에서 오는 데이터로 장애물과 잎의 거리를 측정하는 알고리즘 개발
- 토마토 채취를 위한 수확 가능 여부 파악 알고리즘 개발

3. 접근방법 (Approaches)

- 토마토 재배환경에서의 다양한 영상 자료 취득
- 영상의 분류를 통한 가지 및 나무의 거리 데이터 분석 및 도출
- 잎의 수량에 따른 농약 분무 타이밍의 최적화
- 다양한 영상처리 기법을 통한 토마토 재배 데이터 학습
- 토마토 생산 현장에서의 직접 데이터 수집 및 분류
- AI Data 허브에 있는 토마토 생육에 대한 데이터를 활용하여 학습

4. 테스트 방법 (Test)

- 개발된 모델을 다른 토마토 재배 하우스의 생육 사진들을 직접 취득하여 개발된 모델의 근거리 잎의 밀집도를 직접 파악한 것과 비교하여 성능 테스트
- 추출된 잎 데이터들의 거리값 도출 결과를 실제와 비교하여 그 성능을 테스트함

5. 주요 자원 (Resources)

- 토마토재배지에서의 영상 데이터 취득 가능, 필요시 카메라 등을 설치하여 데이터를 원격지에서 획득 가능
- 데이터량이 많을 경우 원격 조정으로 고성능 연산 컴퓨터 사용 가능

6. 기타사항

- <https://www.aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&dataSetSn=534>



캡스톤디자인 산업체연계형 주제 #7

1. 주제 (title)

범용카메라를 활용한 안면감지(Face Recognition)와 시선추적(Eye Tracking)이 가능한 온디바이스 AI 모델 개발

2. 목표 (Goal)

■ 1차 목표 (Basic Goal)

- 안면구분을 위한 적합한 인공지능 모델 선정
- 카메라 위치에 따른 정상이미지와 비정상이미지의 분류
- 설치형 카메라로 촬영한 얼굴이미지를 분석하여 연령별(5 Classes) 성별(Gender) 구분
- 얼굴위치 인식을 통한 메뉴높이 조정(상하)
- 얼굴인식서비스를 통한 연령별 맞춤형 키오스크 UI/UX 제공 (보통글씨, 큰글씨)
- 키오스크(PC) 혹은 모바일에 적용가능한 형태의 Application 제작.

■ 2차 목표 (Advanced Goal)

- 고가의 전용 카메라가 아닌 저가의 설치형 카메라를 활용하여 동공(눈동자)을 인식
- 사용자가 안내 커서를 따라 동공을 움직이며 Calibration 진행

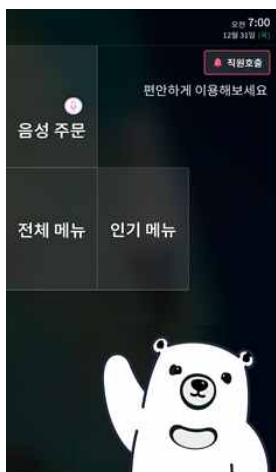
=> Calibration 시간 단축

- 눈동자의 움직임을 따라 마우스 커서가 움직임
- 키오스크(PC) 혹은 모바일에 적용가능한 형태의 Application 제작.

○ 오픈소스를 활용한 시선추적 알고리즘 구현

○ 시선추적을 통한 Easy Menu Click 기능 개발

- 사용자의 상황과 몰입도, 거리, 머문 시간 등의 정보를 수집하여 통계치를 기반으로 상황인지형 인터랙티브 콘텐츠를 제공함으로써 실시간으로 사용자 맞춤형 서비스를 제공



Eye Tracking UI 예시

○ 온디바이스 적용을 위한 모델 경량화 (Option)

1. 시선의 움직임 감지

2. 세밀한 터치가 불가하므로 공간을

분할하여 특정 영역에 일정시간 동안

시선이 머물면 메뉴 동작

※ 영역내 머무른 시간, UI 상의 분할

기준 등에 대해서는 인식 기술 개발

검증과 병행하여 수립

3. 접근방법 (Approaches)

[인공지능 역량기술]

- 인공지능 이론 전반에 대한 이해,

- 다양한 딥러닝 기반 모델 이해,
- PyTorch, TensorFlow 프레임워크, Python, C언어 사용 가능,
- 컴퓨터 비전 관련 인공지능 모델 구현 경험
- PC 프로그래밍 or Mobile 프로그래밍 (개발 Framework은 무관)

[핵심역량 및 해결방안]

- AI허브 안면이미지 영상을 통한 데이터 구축
- 기존의 안면인식 연구에서 다루었던 인공지능 모델을 탐색하고 프로젝트를 위한 모델 제시
- 유사한 사용자의 나이와 성별 구분.
- 사용자의 나이와 성별을 구분하여 메뉴 구조 모델 개발: 나이(Age), 성별(Gender)
- 동공의 움직임에 따라 메뉴전환
- 학습 데이터모델 작성 및 입력
- PC Demo 또는 Mobile App을 통한 얼굴인식과 Eye Tracking 성능테스트

[성능목표]

- 얼굴 위치에 따른 메뉴 높낮이 조정 (1차)
- 나이별, 성별 구분 정확도 분석 (1차)
- 나이 및 성별에 따른 UI/UX 제공(1차)
- 동공을 인식하여 Calibration 진행
- 동공의 움직임과 멈춤에 따라 메뉴(UI/UX) 구현 (2차)
- 키오스크 적용 가능한 Solution 제공 (2차)
- => 키오스크의 넓은 화면에서 시선추적을 위한 방안 제시(2차)

4. 테스트 방법 (Test)

- 실시간 안면인식 확인 (정량적)
 - 정확도 (성별, 나이별), 처리시간, 에러 표출 여부
- 실시간 시선추적 확인 (정량적)
- UI/UX 가이드에 맞는 메뉴 Process 설계/구현 (정성적)

5. 주요 자원 (Resources)

- 안면인식을 위한 세그멘테이션 및 레이블 처리된 데이터
(안면 인식 에이징(aging) 이미지 데이터)
- 시선추적을 위한 기본 model 제공(Source 포함)
- 성능 평가를 위한 테스트 데이터
- 애플리케이션 기능 적용을 위한 카메라
- 키오스크 UI/UX 가이드 제공

6. 기타사항

- 얼굴인식과 시선추적에 대한 응용기술 확보 및 다양한 키오스크에 활용 가능
- 협력기업에서 진행 중인 키오스크 UI 플랫폼 가이드 개발사업에 활용
- 비대면 시대에 주목받고 있는 자동주문시스템 전문기업들과 전략적인 기술제휴를 통해 AI 스마트 키오스크 기술을 적용한 비대면 비접촉 주문시스템을 개발하여 사업의 확대에 적극적으로 활용 예정