

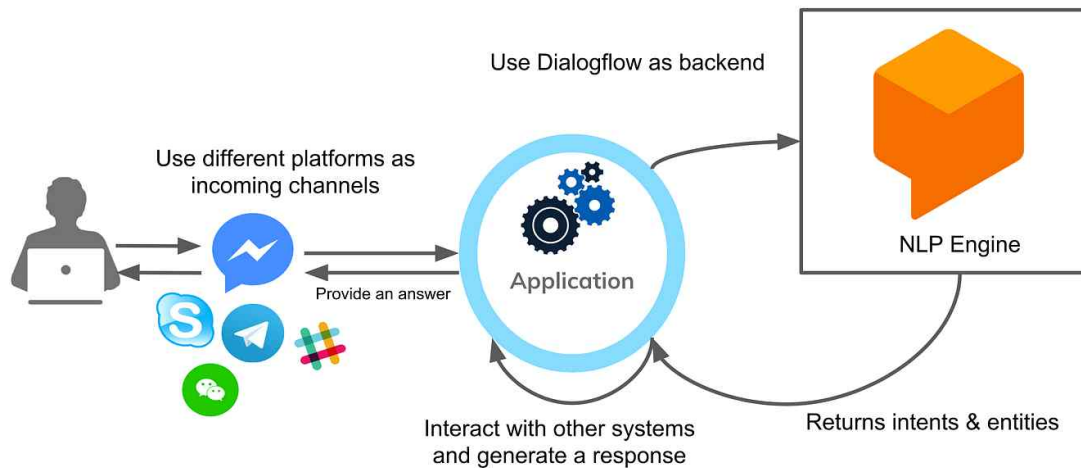
AICOSS 캡스톤디자인 산업체연계형 과제 주제 #1

1. 주제 (title)	멘토(Mentor)	고두현 코치
생성 모델(Generative AI Model)을 활용한 디지털휴먼 제작기술 고도화		
2. 목표 (Goal)		
(1차)		
- 대규모 자연어처리(NLP)모델을 활용하여 답변 추론의 신뢰도(Confidence) 성능지표 획득 - 선정된 모델을 API 서비스로 제공하기 위한 RESTful API 제작		
(2차)		
- Google Dialogflow V2 API를 활용한 디지털 휴먼 제작기술 파악 - 기존 상용 제작기술의 고도화 서비스 연구 (다국어 지원, 질의의 감정 분석, 답변 생성 등)		
3. 접근방법 (Approaches)		
- 주요 개발내용 · ChatGPT3, LLaMA, BERT 모델에 대한 자연어처리 테스트 성능 비교 · 우수한 자연어처리 모델을 활용하여 질의에 대한 답변을 학습시키는 모듈 제작 · 고도화 기능 선정 및 학습을 위한 전처리(토큰화, 정규화, 형태소 분석, 패딩) 작업 · 기존 Google ai 기반 인공지능 대화형 챗봇에 해당 기능 적용 및 사용성 테스트 - 성능목표 · 사전 학습된 대규모 자연어처리 모델에 대한 성능비교 연구 및 결론도출 (1차) · 새로운 대화를 학습시키기 위한 API 인터페이스 개발 (1차) · 기존 Google ai 기반 음성대화형 QA 챗봇 분석 및 연동(2차) · 고도화된 API를 적용한 디지털 휴먼 제작(2차) - 기능개발시 필요요소 · 생성형 인공지능, 음성인식 서비스, 음성출력(TTS) 서비스에 대한 이해(권장) · Python을 활용한 API(Interface) 제작역량(필수) · 앱, 웹, 게임엔진 등 Front-end 개발역량(권장)		
4. 테스트 방법 (Test)		
- 고도화 방향에 따른 학습모델 선정 및 성능지표 제시 - 생성형 모델을 학습하여 자연어 답변생성 기능이 정상적으로 구현되었는지 확인 - Google Dialogflow의 질문(의도)파악 모델에 대한 학습이 정상적으로 구현되었는지 확인 - 생성 모델과 기존 디지털휴먼 챗봇을 연동하여 고도화된 서비스를 제작하였는지 확인		
5. 주요 자원 (Resources)		
- 기존 상용 챗봇서비스 소개 및 개발가이드 제공 - 생성 모델을 활용한 디지털휴먼 제작기술 가이드 제시 - GCloud VM, 자연어 라벨링데이터(10만건)		
6. 기타사항		
※ AI Chatbot (Dialogflow)		
https://cloud.google.com/dialogflow/es/docs/reference/rest/v2-overview		
※ Generative AI Model (GPT3, LLaMA, BERT)		
https://openai.com/blog/introducing-chatgpt-and-whisper-apis		
https://ai.facebook.com/blog/large-language-model-llama-meta-ai/		
https://cloud.google.com/ai-platform/training/docs/algorithms/bert-start?hl=ko		

[별첨]

1. 주제 개요

- 당사가 보유하고 있는 인공지능 기반 디지털휴먼의 제작기술은 음성을 인식하고 인식된 데이터의 의도어(Intent), 엔티티(Entity)를 판별하여 적절한 답변을 추출해내는 프레임워크 구조로 설계되어 있습니다. 저희는 올해 상용 챗봇제작기술을 고도화 하기위하여, 3가지 핵심 역량을 확보하고자 합니다.



<기존 상용 챗봇 플랫폼의 서비스 구조>

- 자연어처리(NLP) 분야에 사용되는 딥러닝 생성모델 중 전이학습(Transfor Learning)에 용이한 모델을 선정하고 새로운 성능비교를 해볼 수 있는 연구역량
- 사용자와 챗봇간의 대화 시나리오를 기반으로 새로운 답변을 생성하는 대화 데이터셋과 시뮬레이션 인터페이스 제작역량
- 대규모 데이터셋을 이용한 학습모델을 기반으로 새로운 데이터셋을 학습시켜 다양한 서비스 모델 연구

2. 연구방향 및 기대효과

- 인간의 언어에 대한 이해와 생성능력을 갖출 수 있습니다. 이를 통하여 당사의 기존 챗봇의 응답이 더 자연스러워지고 대화가 순조롭게 이어질 수 있습니다.
- 사용자의 질문에 대한 자세한 답변을 제공하거나, 감정인식/상황추론 등 추가 모델을 사용하여 사용자의 감정을 인식하고 그에 따라 대응할 수 있습니다.
- 당사의 기존 의도어에 대한 답변의 성능지표를 개선하고, 답변관리를 인공지능 모델을 통해 자동화 하는 효율성을 확보할 수 있습니다.

3. 인공지능 연구용 언어모델 추천

- GPT-3 davinci, Curie, Babbage, Lama LaBSE, MBERT/KBERT

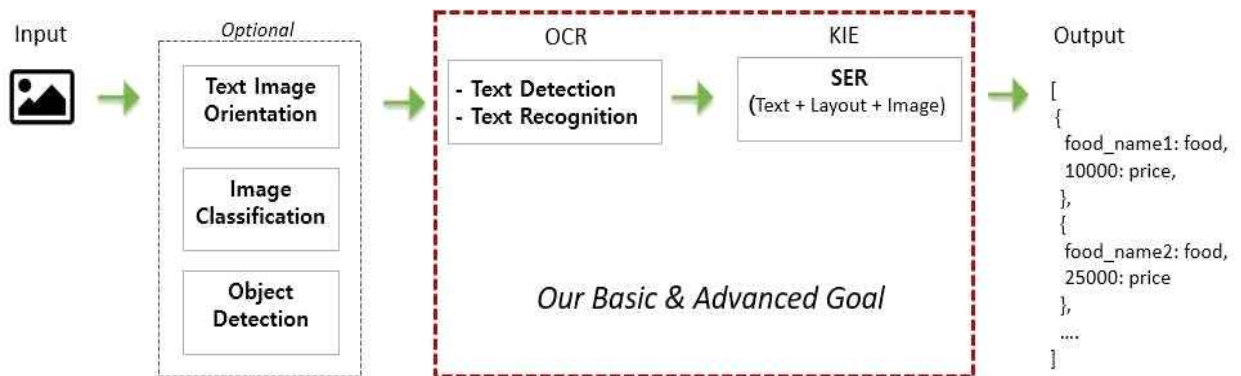
4. 성능지표 기준

- Perplexity 65 미만대 형성
- Dialog Confidence 0.7 이상

AICOSS 캡스톤디자인 산업체연계형 과제 주제 #2

1. 주제 (title)	멘토(Mentor)	길가영 코치
AI-OCR: 메뉴판 인식 및 개인화 음식 추천 서비스		
2. 목표 (Goal)		
(1차)		
- 사진 촬영 이미지 OCR 서비스 구현 (Text Detection + Text Recognition)		
(2차)		
- KIE(Key Information Extraction) 구현 및 API호출 통해서 음식 칼로리 정보 추출/추천		
3. 접근방법 (Approaches)		
- 음식 메뉴판 사진에 대해 OCR 적용하여 텍스트를 추출하고, 추출된 정보중에서 유의미한 정보(예: 음식명)를 추출하는 기술 (<i>추출 음식명에 대해서 음식 DB 연동한 칼로리 정보, 개인화된 음식 추천 서비스 등으로 확장 가능</i>)		
- 서비스 구현 예시: <i>별첨 참조</i>		
■ OCR algorithms (Two-stage approach)		
1) Text Detection: 이미지로부터 텍스트의 위치 인식(Bounding Box)		
- DB, DB++, EAST, SAST, PSE, ...		
2) Text Recognition: 인식된 텍스트 박스의 내용 인식		
- CRNN, SRN, SVTR, ABINet, VisionLAN, RobustScanner, ...		
■ SER(Semantic Entity Recognition): OCR통해 추출된 텍스트의 엔티티를 분류하는 기술		
- SDMGR, InfoXLM, XLM-RoBERTa, LayoutXLM, Vi-LayoutXLM, ..		
4. 테스트 방법 (Test)		
■ 부가적인 서비스 구현(사진촬영, 칼로리 정보 추출)보다는 OCR 및 KIE 모델 구현에 중점을 두어서 테스트를 진행 합니다. 각 모델의 테스트는 아래의 평가 매트릭스로 진행합니다.		
- Text Detection: hmean score (<i>baseline: 0.5</i>)		
- Text Recognition: Avg accuracy (<i>baseline: 0.5</i>)		
- SER: F1 score (<i>baseline: 0.5</i>)		
5. 주요 자원 (Resources)		
- 학습용 데이터: 관광 분야 인공지능 학습용 지식베이스-식당 메뉴판 이미지/어노테이션		
- 부가 툴: 추가 학습 데이터 생성용(Det., Rec. KIE) labelMaker		
- 학습용 GPU Server/Workstation		
6. 기타사항		
■ Open Source & Papers		
https://paperswithcode.com/task/scene-text-detecti		
on https://github.com/open-mmlab/mimocr		
https://github.com/PaddlePaddle/PaddleOCR		
https://github.com/microsoft/unilm : Multimodal(X + Language) 섹션 참조		

[별첨]



AICOSS 캡스톤디자인 산업체연계형 과제 주제 #3

1. 주제 (title)	멘토(Mentor)	김진혁 코치
Re-Identification(재인식)을 활용한 추적기 성능 개선		
2. 목표 (Goal)		
1차 목표 : 새로운 추적 ID가 생성될 때 Re-ID로 기존의 ID에 존재하던 객체인지 판별		
2차 목표 : Re-ID를 활용한 추적 ID 스위칭 문제를 해결		
3. 접근방법 (Approaches)		
기본적인 프로세스는 다음과 같습니다.		
1. 입력 영상이 들어온다.		
2. 객체(사람) 인식을 한다.		
3. 추적기가 객체마다 ID를 부여한다.		
4. ID가 부여된 객체를 Re-ID 모델에 넣어 기존에 존재하던 ID인지 판별한다. (1차 목표)		
5. 만약 기존에 존재하는 ID와 같다고 판별됐는데 ID가 서로 다른 경우 기존에 있던 ID로 바꿔준다. (2차 목표)		
Re-ID를 적용했을 때와 안했을 때 추적 ID 개수 차이를 비교하여 성능 향상이 이뤄졌다는 것을 확인할 수 있어야 합니다. 입력 영상에 객체별로 추적된 박스 영역 혹은 Instance Segment 영역을 그리고 추적 ID를 표시하여야 합니다. 마지막 추적 ID를 영상 한곳에 계속 카운트하여 표시해주셔야 합니다.		
1. [별첨]에서 제시하는 모델들은 제안하는 모델들이며 더 좋다고 생각하는 성능의 모델이 있다면 사용해도 무방합니다.		
2. [별첨]에서 제시한 성능 목표 달성 방법 2가지 기준 중 하나라도 충족되어야 합니다.		
4. 테스트 방법 (Test)		
모델별로 성능 목표를 달성했는지 테스트 영상을 통해 다음과 같은 내용을 확인합니다.		
1. 객체인식, 추적기, Re-ID에서 각 모델별로 제시된 지표들을 만족하는가		
2. Re-ID를 통해 추적기 ID 스위칭 문제가 개선 되었는가		
시연에 필요한 방식에 제한은 두지 않겠습니다. jupyter-lab을 사용하셔도 되고 웹사이트 형태로 구성하셔도 좋습니다. 입력 영상을 넣으면 모델별 성능지표 결과와 결과 영상을 확인할 수 있는 구조라면 상관없습니다.		
5. 주요 자원 (Resources)		
- 제공하는 Dataset : coco2017, MOT17, Market-1501		
- 과제에 필요한 딥러닝 자료 및 개발 가이드		
- 개발 환경이 셋팅된 Docker Image 제공		
- 추가 학습을 위한 데이터셋 제작이 필요한 경우 라벨링 툴 제공		
6. 기타사항		
Object Detection 정리 : https://techblog-history-younghunjo1.tistory.com/178		
Tracking 정리 : https://gaussian37.github.io/vision-concept-vot_mot/		
Re-ID 정리 : https://junha1125.github.io/blog/artificial-intelligence/2021-02-17-reID_Survey/		
yolov5 : https://github.com/ultralytics/yolov5		
yolov7 : https://github.com/WongKinYiu/yolov7		
Mask-RCNN, MAsk2Former, RTMDet : https://github.com/open-mmlab/mmdetection		
Byte-Track : https://github.com/open-mmlab/mmottracking/tree/master/configs/mot/bytetrack		
Re-ID : https://github.com/DengpanFu/LUPerson/tree/main/fast-reid		
필요한 프로그래밍 스킬 : Pytorch, Python 프로그래밍 기술, Deep-learning 개념		

[별첨]

과제를 수행하는데 추천하는 모델 목록

Object-Detection, Tracking, Re-Identification 측면으로 나눠서 진행

객체 인식 (Object-Detection) : YOLOv5, YOLOv7, Mask-RCNN, Mask2Former, RTMDet 등의 모델을 활용

추적기 (Tracking) : ByteTrack 활용

재인식 (Re-Identification) : Fast-reid 활용

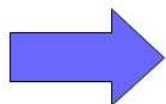
성능 목표 달성 방법

각 모델별로 해당 지표들의 수치를 만족해야 합니다.

- Object-Detection (coco2017 Dataset Test 기준) : mAP : 0.5 이상 or boxAP : 45 이상
- Tracking (MOT17 Dataset Test 기준) : MOTA : 70 이상 or HOTA : 60 이상
- Re-Identification (Market-1501 Dataset Test 기준) : mAP : 95 이상 or Rank 1 : 95 이상

Re-ID 예시

처음 아이디를 찾았을 때는 545개의 아이디가 나왔으나 Re-ID 후 중복을 제거해 40개로 합쳐짐.

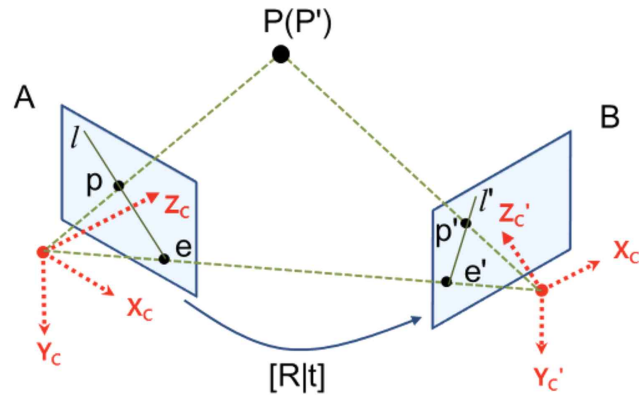


AICOSS 캡스톤디자인 산업체연계형 과제 주제 #4

1. 주제 (title)	멘토(Mentor)	이정훈 코치
실내 공간에서의 자율주행 로봇을 위한 다중 카메라 이용 특정 인물의 실제 위치		
2. 목표 (Goal)		
- 다중 카메라를 이용한 실내 공간 지면 좌표화		
- 인상착의 기반 특정 인물 인식 및 공간에서의 위치값 추출		
3. 접근방법 (Approaches)		
궁극적 목표는 무인 매장 내의 CCTV 카메라를 활용해 특정 인물만을 탐지하고 객체의 위치를 실제 공간의 좌표값으로 자율주행 로봇에 전달하기 위함임.		
※ 목표를 달성하기 위한 절차 또는 주요 개발 내용		
1-① camera calibration을 통해 카메라 내/외부 파라미터 구하기		
1-② 2개의 카메라를 epipolar geometry의 triangulation을 이용하여 공간의 지면을 좌표화		
2-① 인물 객체 탐지		
2-② 입고 있는 의상정보를 중심으로 인상착의를 인식할 수 있는 분류		
2-③ 공간의 2D 좌표계 내에서 탐지된 인물의 위치 좌표 구하기		
4. 테스트 방법 (Test)		
※ 목표 달성 결과에서 제시할 항목과 테스트 방법 등		
1. 객체 탐지 성능 측정		
2. 의상 분류 성능 측정		
3. 인물의 위치 좌표와 실제 위치의 차이 측정 (비즈니스적 기준)		
5. 주요 자원 (Resources)		
※ 제공되는 데이터와 규모		
DeepFashion2 Dataset (클래스당 8000여개의 데이터)		
※ 개발에서 필요한 컴퓨팅 환경 등		
6. 기타사항		
※ 참여 팀들이 문제를 이해하는데 필요한 정보 또는 사이트 등		
영상처리 관련 정보 블로그		
https://darkpgmr.tistory.com/category/%EC%98%81%EC%83%81%EC%B2%98%EB%A6%AC		
의상 분류 데이터셋 깃허브		
https://github.com/switchablenorms/DeepFashion2		

[별첨]

① 다중 카메라를 이용한 실내 공간 지면 좌표화



<그림 1>

- CCTV 영상과 실제 공간 차원이 다른 두 좌표계 간의 변환 관계를 만족하는 행렬을 구하고 카메라 파라미터를 포함한 두 이미지의 실제 픽셀 좌표 사이의 기하학적 관계를 구하여 스테레오 비전에서 3D 공간좌표 P를 결정할 수 있다.

② 인상착의 기반 특정 인물 인식 및 공간에서의 위치값 추출



- 영상 내 특정한 옷을 입고 있는 사람만 탐지하고 지면 좌표와 매핑하여 어느 좌표에 위치하는지 실제 위치를 구한다.

※ 이미지 출처

- <https://darkpgmr.tistory.com/153>
- <https://trepo.tuni.fi/bitstream/handle/123456789/27490/granat.pdf?sequence=4&isAllowed=y>

AICOSS 캡스톤디자인 산업체연계형 과제 주제 #5

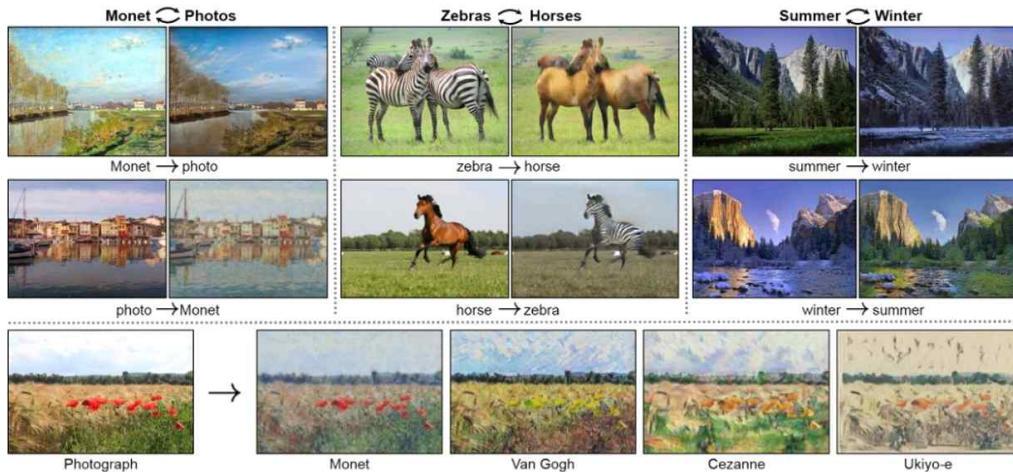
1. 주제 (title)	멘토(Mentor)	주소현 코치
CycleGAN을 이용한 데이터 증강		
2. 목표 (Goal)		
- CycleGAN을 이용하여 희귀한 케이스의 데이터 증강		
- 희귀한 케이스 데이터를 Domain Adaptive Deeplearning으로 모델 성능 향상		
3. 접근방법 (Approaches)		
- CycleGAN을 이용한 백돼지 사진을 흑돼지(외국 돼지 품종 등) 사진으로 변경 - 마찬가지로 한우와 젓소 transfer - stable diffusion 기법을 활용한 데이터 증강 시도 - 일반 훈련 모델과 DA 모델 성능 비교		
4. 테스트 방법 (Test)		
- 모델 훈련 결과 mAP (testset : 흑돼지 가정) - 1. 일반 COCO모델(pretrain) 2. COCO + transfer learning(백돼지 - 라벨 O) 3. COCO + transfer learning(생성된 흑돼지 - 라벨 O) 4. COCO + DA(생성된 흑돼지 - 라벨 O)		
5. 주요 자원 (Resources)		
- cuda11 + 3090 + docker + linux - 제공되는 데이터는 회사와 협의		
6. 기타사항		
https://junyanz.github.io/CycleGAN/ https://paperswithcode.com/task/domain-adaptation https://paperswithcode.com/task/few-shot-learning https://github.com/WongKinYiu/yolov7		

[별첨]

○ 목표 : 돼지 품종, 일령, 스타일을 모두 커버할 수 있는 모델 개발

- ▶ 다양한 방법으로 모델을 훈련하는 프레임워크 개발(DA, FS, AL 등)
- ▶ 이미지 생성 프레임워크 개발(CycleGAN)

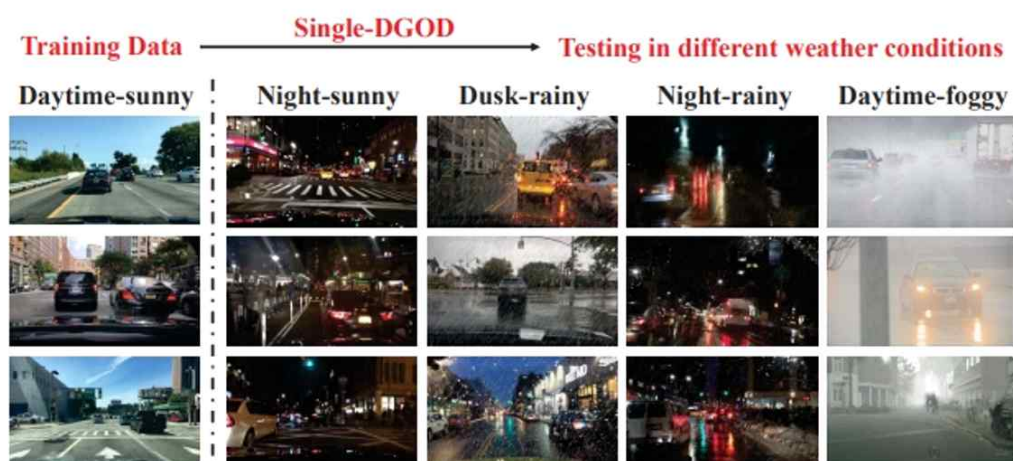
○ CycleGAN이란?



링크: <https://junyanz.github.io/CycleGAN/>

- ▶ 위의 예시로 사진을 어떤 화가의 화풍으로 변경, 일반 말을 얼룩말로, 여름 사진을 겨울 사진으로 변경하는 등 하나의 도메인을 다른 도메인으로 스타일을 변경하는 것
- ▶ 훈련 시 라벨링 데이터가 필요하지 않음(비용 문제X, 직접 라벨할 필요 없음)**
- ▶ 저희가 제공할 데이터에는 백돼지 도메인 라벨이 제공되어 흑돼지로 transfer진행 후 해당 라벨을 재사용하여 O.D. 가능함

○ Domain Adaptive Learning

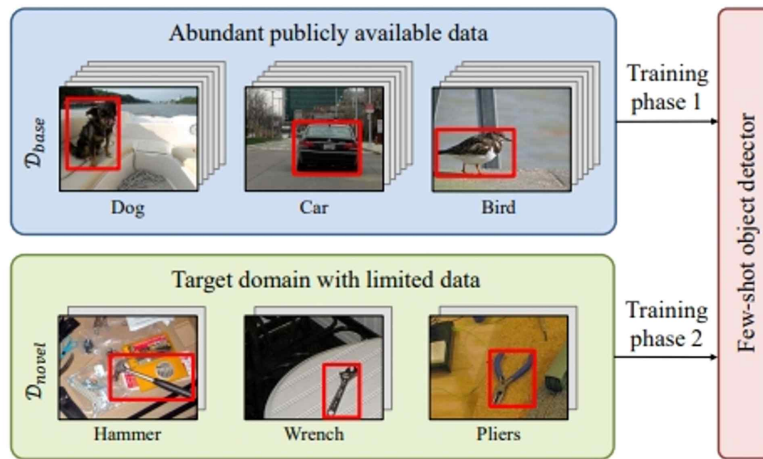


링크 : Single-Domain Generalized Object Detection in Urban Scene via
Cyclic-Disentangled Self-Distillation

- ▶ 이미 훈련된 모델을 다른 도메인에서도 성능이 나을 수 있도록 훈련하는 기법
- ▶ 예를 들어 낮 도메인에서 모은 데이터를 가지고 훈련모델을 만들었다면, 저녁, 안개 등에서는 대처

하기 힘들다. 하지만 저녁, 안개의 일부 데이터를 가지고 transfer learning을 진행한다면 낮 도메인에서 성능이 나빠질 것이다. 이를 해결할 수 있는 방법론

○ Few Shot Learning



출처 : <https://arxiv.org/pdf/2112.11699.pdf>

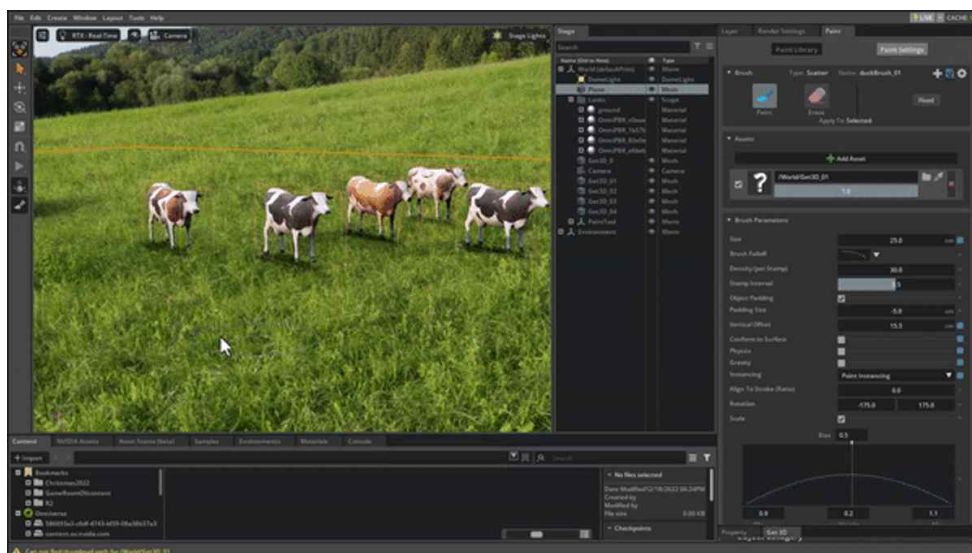
- ▶ 학습된 모델에 새로운 클래스를 추가하고 싶을 때 사용하는 방법론
- ▶ 혹은 학습된 모델에서 잘 감지하지 못하는 특정이미지 소수에 대해서 재학습시켜주는 용도로 사용

○ 구상중인 아이디어

Stable Diffusion

- ▶ 텍스트에서 이미지 만들기 or 이미지에서 이미지 만들기 등으로 희귀 케이스 이미지를 늘리는데 사용될 것으로 판단되나, 모델의 크기가 너무 크고 내용을 이해하기 어려울 수 있음
- ▶ "몽쳐있는 돼지 사진" → 이미지 출력되는 모델

Nvidia Metaverse



- ▶ 마찬가지로 희귀 케이스 이미지를 생성하는 데 사용할 것으로 판단됨