

# 국외출장 결과보고서

기 간: 2022. 06. 17 ~ 2022. 06. 26

출장지: 미국(뉴올리언스)

출장자: 최경아, 임룡혁

# I. 출장개요

1. 출 장 지: 미국 루이지애나주 뉴올리언스(CVPR 2022 참석)
2. 출장기간: 2022. 06. 17 ~ 2022. 06. 26
3. 출 장 자

소속	직급	성명	비고
공간정보사회연구본부	부연구위원	최경아	연구책임자
공간정보사회연구본부	연구원	임룡혁	연구진

## 4. 출장목적

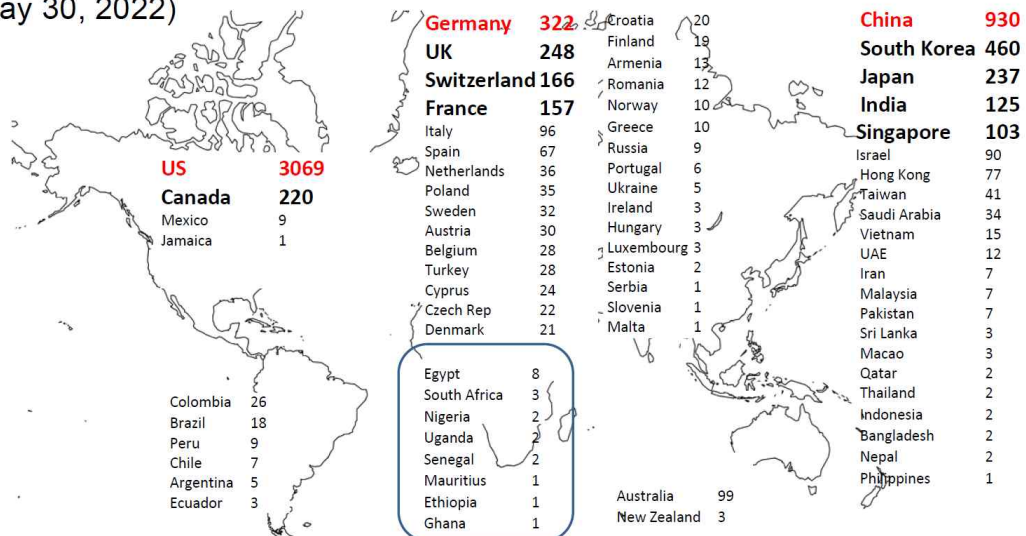
- 과제수행에 필요한 다중공간데이터 품질 검증 및 향상 알고리즘 개발을 위한 최신 연구 동향 파악 및 핵심 기술 습득
- 컴퓨터 비전분야의 공간정보 기술 적용 동향 파악 및 글로벌 네트워크 구축

## 5. CVPR 2022(Computer Vision and Pattern Recognition Conference 2022) 개요

- CVPR은 국제전기전자공학회(IEEE)와 국제컴퓨터비전재단(CVF)이 1983년부터 공동 주최하는 컴퓨터비전 분야의 세계 최고, 최대 규모의 학술대회로, 최신 AI 모델을 활용하여 영상을 처리하는 다양한 방법론과 응용에 관한 연구성과를 발표하며 활발한 토론 및 교류가 이루어지는 대표적인 학회임
- 올해에는 CVPR 2022에 투고된 8,161편의 논문 중 2,064편의 논문이 게재되었으며, 전 세계 곳곳에서 9,981명이 등록 및 참여하였음(22/06/20 기준)

### CVPR 2022: ~7,289 attendees from 74 countries/regions

(May 30, 2022)



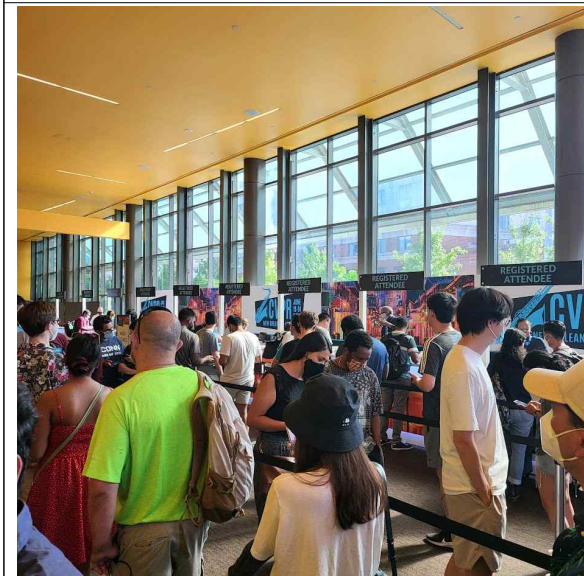
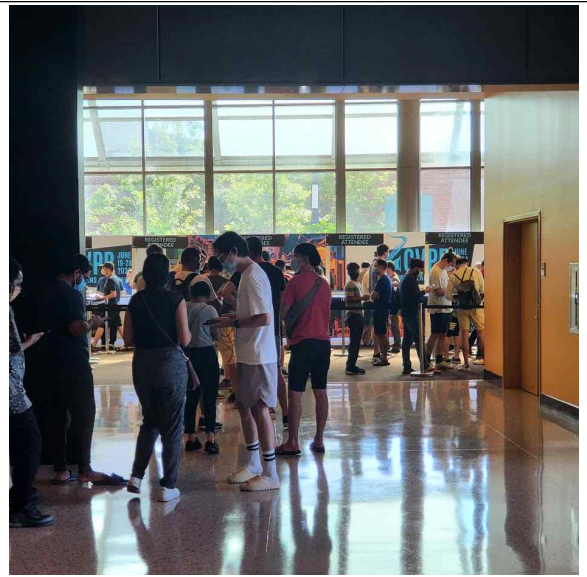
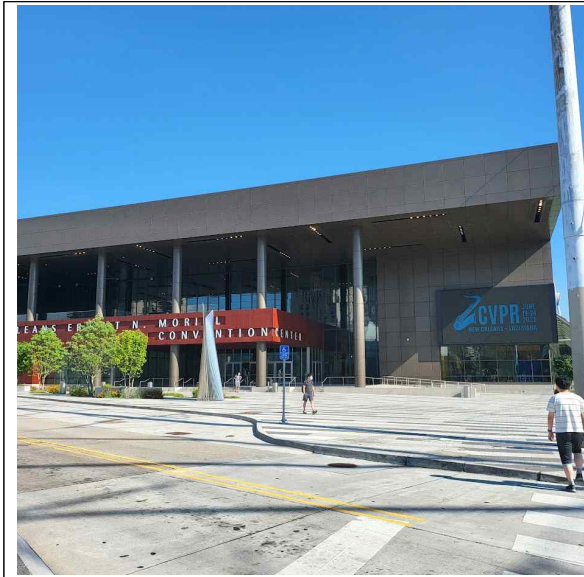
## II. 출장일정

일정 (요일)	출발지	도착지	업무수행내용	접촉 예정인물 (직책포함)
06/17(금)	인천	미국	인천->로스엔젤레스 6/17(금) 14:30 -> 6/17(금) 09:40 로스엔젤레스->뉴올리언스 6/17(금) 12:20 -> 6/17(금) 18:12	
06/18(토)			<ul style="list-style-type: none"> <li>연구진 회의: 학회 주요 일정 확인 및 자료수집 방향 논의</li> <li>학회 등록</li> </ul>	
06/19(일)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Workshop): "7th New Trends in Image Restoration and Enhancement Workshop and Challenges" "EarthVision: Large Scale Computer Vision for Remote Sensing Imagery"</li> </ul>	
06/20(월)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Workshop): "5th Efficient Deep Learning for Computer Vision" "The Fourth Workshop on Deep Learning for Geometric Computing"</li> </ul>	
06/21(화)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Main Conference): 자료 수집 및 학회 발표 관람</li> </ul>	
06/22(수)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Main Conference): 자료 수집 및 학회 발표 관람</li> </ul>	
06/23(목)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Main Conference): 자료 수집 및 학회 발표 관람</li> </ul>	
06/24(금)			<ul style="list-style-type: none"> <li>CVPR 2022 참석(Main Conference): 자료 수집 및 학회 발표 관람</li> </ul>	
06/25(토)	미국		뉴올리언스->애틀란타 6/25(토) 10:15 -> 6/25(토) 12:43	
06/26(일)		인천	애틀란타->인천 6/25(토) 15:00 -> 6/26(일) 19:10	

### III. 수행사항

#### 1. Registration

- 일시 및 장소: 2022. 06. 18(토) Ernest N. Morial Convention Center
- 목적: 학회 등록
- 주요 사진



## 2. 2022.06.19.(일) Workshops

□ 일시 및 장소: 2022. 06. 19(일) 08:30 Ernest N. Morial Convention Center

- (Workshop) EarthVision: Large Scale Computer Vision for Remote Sensing Imagery
- 지구 관측(EO, Earth Observation) 및 원격탐사(RS, Remote Sensing)는 컴퓨터 비전, 머신러닝, 신호 및 영상 처리 분야에서 지속적으로 성장하는 분야 중 하나임
  - 위치 기반 서비스, 온라인 지도 서비스, 3D 도시 모델링, 내비게이션 시스템, 자연 재해 예측 및 대응, 기후 변화 모니터링 등과 같은 수많은 분야에서 활용 수요가 있으며, AI 및 딥러닝에 필요한 엄청난 양의 데이터를 처리하고 분석하기 위해서는 자동화된 워크플로우가 필요
  - EO/RS의 데이터의 분석은 UN에서 정의한 지속 가능한 개발 목표의 40개 중 34개 지표(29개 목표 및 11개 목표)와 직접 연결되며, 또한 산불 및 기타 자연 재해 모니터링, 도시 성장, 삼림 벌채 및 기후 변화와 같은 다른 즉각적인 사회적 문제와 연결되는 중요한 분야임

□ Workshop Program

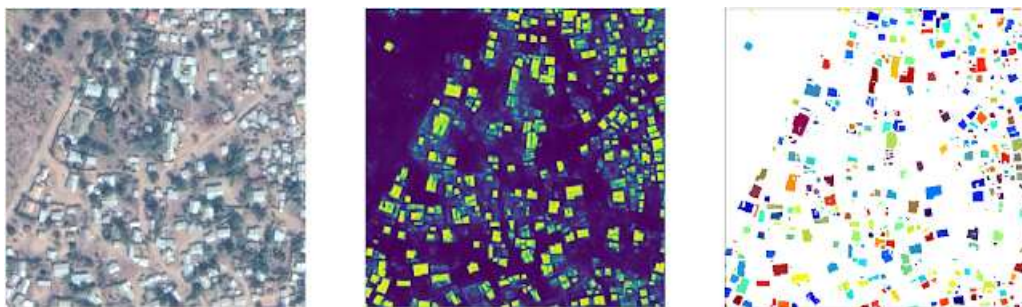
Time	Title	Author
08:30 - 08:45	Welcome	
08:45 - 09:15	Keynote 1 Building footprints for the African continent	John Quinn, Google
09:15 - 10:00	Oral Session 1 Semantic Analysis of Geospatial Data	
10:00 - 10:30	Coffee Break	
10:30 - 11:00	Keynote 2 Can Machines Learn to Predict the Weather? Using Deep Learning Instead of Physical Simulation for Weather Forecasting	Rich Caruana, Microsoft Research
11:00 - 11:45	Oral Session 2 Emerging Applications in Remote Sensing	
11:45 - 13:15	Lunch Break	
13:15 - 13:45	Keynote 3 Imageomics: images as the source of information about life	Tanya Berger-Wolf, Ohio State University
13:45 - 14:30	Oral Session 3 Time Series, Forecasting, Change	
14:30 - 16:00	Coffee Break / Poster Session	
16:00 - 16:30	SpaceNet 8 Challenge	
16:30 - 17:15	Oral Session 4 New Benchmarks for Machine Learning in Earth Observation	
17:15 - 17:30	Closing	

□ 주요 발표내용

- (Keynote 1) Building footprints for the African continent (John Quinn, Google)
  - 항공 및 위성영상에서의 건물 윤곽(building footprint)을 파악하는 것은 인구 추정, 도시계획, 인도주의적 활동, 환경과학 등 다양한 분야에서 매우 중요한 요소임
  - 위성영상 데이터는 아주 훌륭한 데이터 수집 방법이지만, 위성 영상은 지상에서 수백 킬로미터 떨어진 곳에서 지구를 촬영해야 하기 때문에 고해상도 (30~50cm/pixel)에서도 작은 건물은 식별이 쉽지 않으며, 건물과 쉽게 혼동될 수 있는 많은 유형의 지형지물, 인공구조물도 구별이 쉽지 않음
  - 특히 자연 소재로 지어진 건물이나 주변 환경과 시각적으로 조화를 이룰 수 있는 시골 지역 등에서는 분석 및 분류 작업이 훨씬 더 어려움
  - 위 연구에서는 아프리카의 사바나, 사막, 숲과 같은 다양한 지형에 걸쳐 건물을 감지하는 방법을 제시하였으며, 아프리카 대륙 대부분에 걸쳐있는 5억 1,600만 건물의 위치와 윤곽정보를 가지는 오픈 데이터 세트<sup>1)</sup>를 제작 및 제공
  - U-Net 모델을 기반으로 하여 아프리카 지역의 건물 윤곽을 추출하는 실험을 진행하였으며, 농촌 및 도시, 중간밀도를 갖는 지역 등 다양한 지역에서 모델의 성능을 평가하였으며
  - 정밀도와 재현율이 지역마다 다르게 나타났으며, 대륙 전체에서 일관된 성능을 달성하는 것은 지속적인 과제라고 정리하였음

<그림 1> 위성 영상에서의 건물 분할 예시

왼쪽: 원본 이미지, 중앙: 의미론적 분할(Semantic Segmentation), 오른쪽: 인스턴스 분할



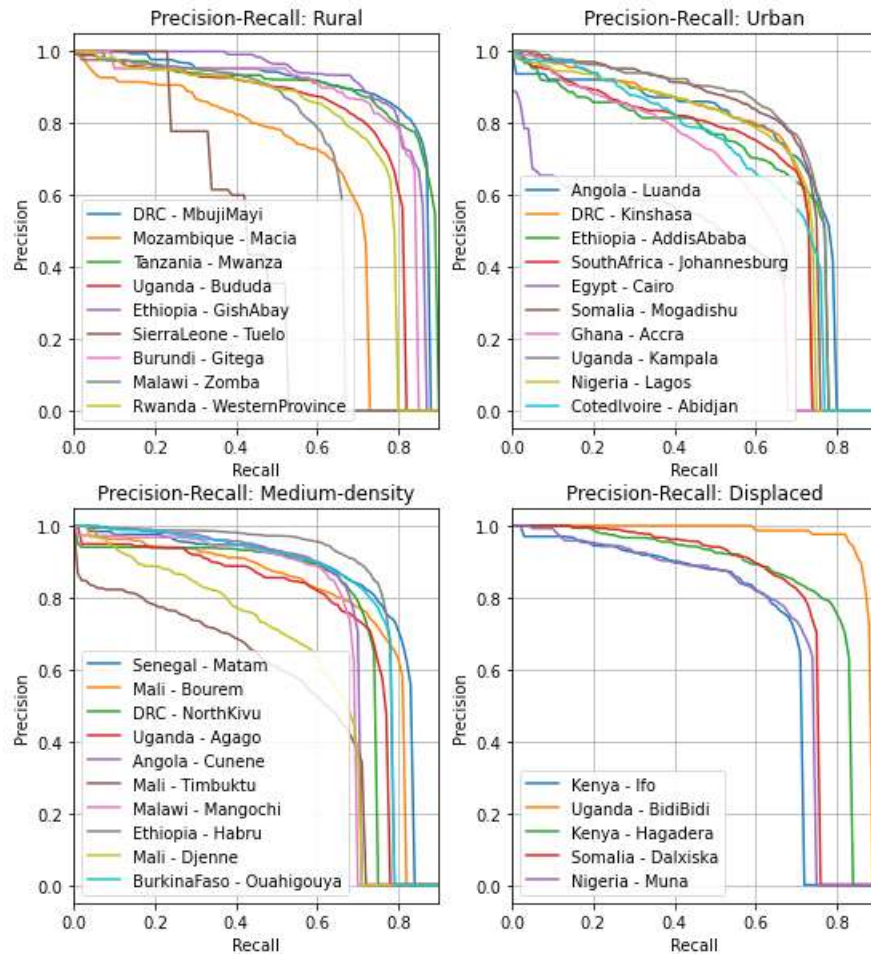
Source: Mapping Africa's Buildings with Satellite Imagery

(<https://ai.googleblog.com/2021/07/mapping-africas-buildings-with.html>, access July 1, 2022)

1) Open Buildings, <https://sites.research.google/open-buildings/>



<그림 2> 정밀도-호출 곡선(Precision-recall curves), Threshold 0.5



Source: Mapping Africa's Buildings with Satellite Imagery

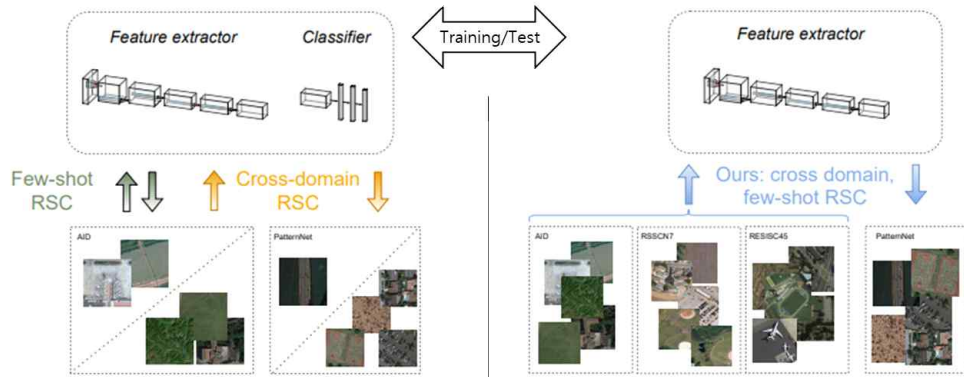
(<https://ai.googleblog.com/2021/07/mapping-africas-buildings-with.html>, access July 1, 2022)

◦ (Oral Session 1) Semantic Analysis of Geospatial Data

1) "Cross-dataset Learning for Generalizable Land Use Scene Classification", Dimitri Gominski (University of Copenhagen); Valérie Gouet-Brunet (LASTIG/IGN-UGE); Liming Chen (Ecole Centrale de Lyon);

- Few-shot & cross-domain 기법을 활용하여 토지이용분류 자동화 프레임워크를 제시하였으며, 제안된 방법은 기존 데이터셋에 분할되지 않은 새로운 데이터가 추가되었을 때, 토지이용분류에 대한 정확한 추정이 가능

<그림 3> Comparison of approaches for RSC

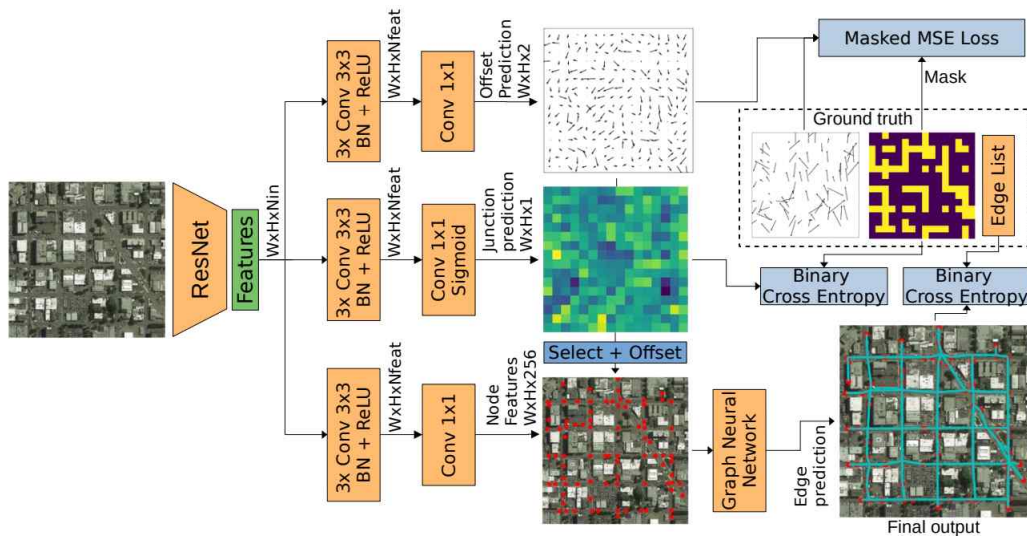


Source: Cross-dataset Learning for Generalizable Land Use Scene Classification(Dimitri Gominski et al., 2022)

2) "Single-Shot End-to-end Road Graph Extraction", Gaetan Bahl (INRIA); Mehdi Bahri (Imperial College London); Florent Lafarge (INRIA);

- 교차로, 막다른 골목, 회전과 같은 관심 지점(POD)을 찾는 완전 컨볼루션 네트워크 (Fully Convolutional Network)와 이들 지점 간의 연결을 예측하는 그래프 신경망을 결합 구성하여, 단방향 도로의 최종 도로 그래프를 직접 추론하는 방법을 제안

<그림 4> Single-shot road graph extraction neural network



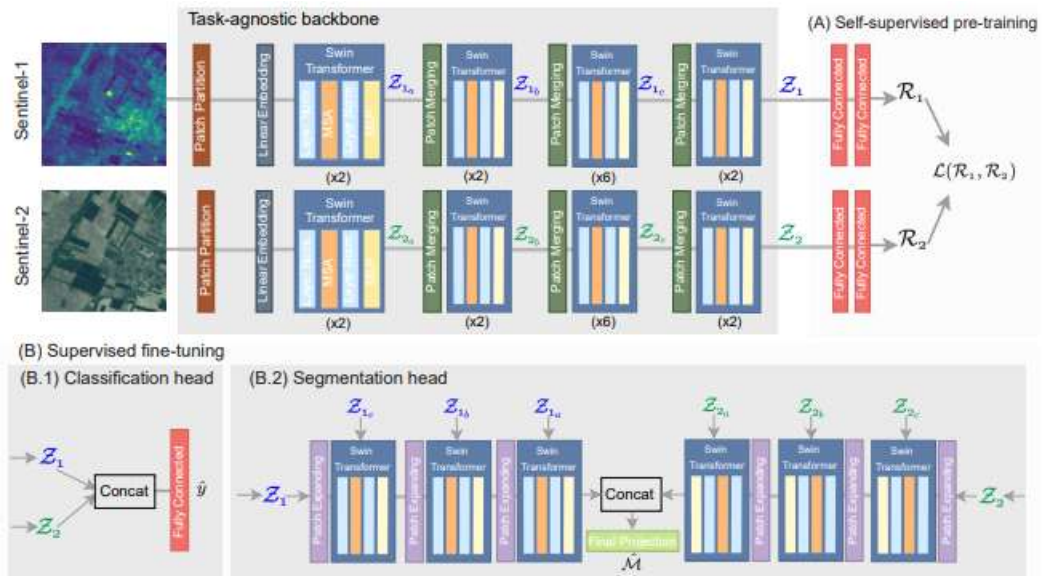
Source: Single-Shot End-to-end Road Graph Extraction(Gaetan Bahl et al., 2022)

3) "Self-supervised Vision Transformers for Land-cover Segmentation and Classification", Linus M. Scheibenreif (University of St. Gallen); Joëlle Hanna (University of St. Gallen); Michael Mommert (University of St. Gallen); Damian Borth (University of St. Gallen);



- Transformer는 내부적으로 RNN 레이어 없이, 어텐션(Attention)으로만 구현한 모델로, 기존에는 자연어 학습에서 속도와 성능을 모두 향상시키는 역할을 하였고, 최근에는 영상의 딥러닝 분석에 활용되고 있으며, 위 연구에서는 이러한 Transformer를 활용하여 토지피복분류를 위한 자체감독학습(Self-supervised) 방법을 제시하고 있음

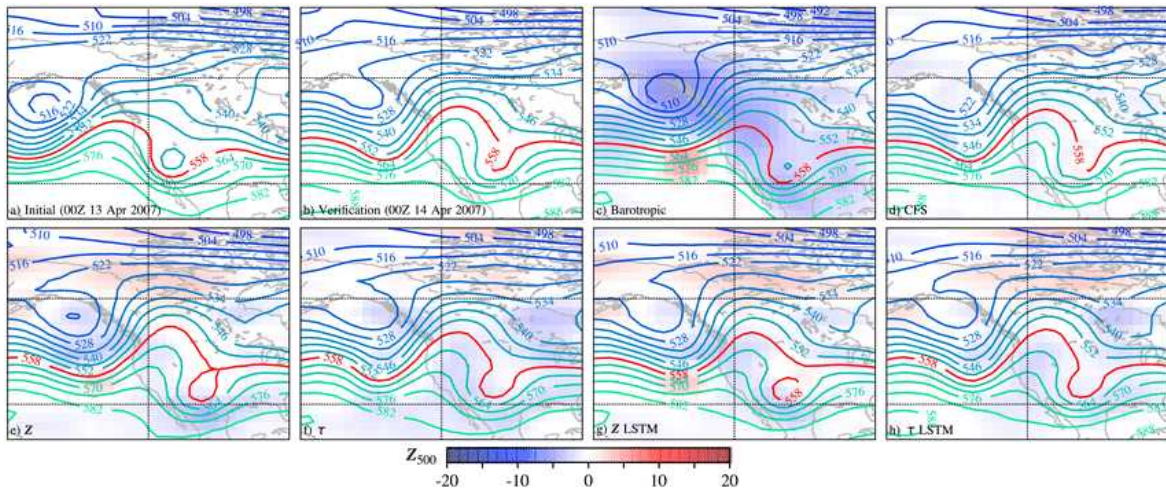
<그림 5> Network architecture



Source: Self-supervised Vision Transformers for Land-cover Segmentation and Classification(Linus M. Scheibenreif et al., 2022)

- o (Keynote 2) Can Machines Learn to Predict the Weather? Using Deep Learning Instead of Physical Simulation for Weather Forecasting (Rich Caruana, Microsoft Research)
  - 일기예보는 대부분 슈퍼컴퓨터에서 대규모 물리적 모델에 기반한 시뮬레이션을 통해 이루어짐
  - 위 발표에서는 시뮬레이션 대신 과거데이터를 심층신경망 훈련을 통해 학습하고, 날씨를 예측하는 방법에 대해 연구한 내용을 소개하였음
  - 유럽중기예보센터(ECMWF, European Centre for Medium-Range Weather Forecasts)에서 운용중인 모델과 비교하여 2주 이내의 날씨를 예측하는 부분에서는 다소 부족하며, 중기(2~6주)에 대한 일기예보에 대해서는 거의 정확한 예측 결과를 보여주고 있음
  - 일기예보 분야는 수십년에 걸친 모델의 개발과 발전이 필요했지만, 딥러닝 분야에서는 단 1주일만의 시간이 필요했다는 것과, 막대한 리소스를 요구로 하는 슈퍼컴퓨터에 비해, GPU를 이용한 일기예측은 일반 컴퓨터 수준에서도 가능하다는 것에 의미가 있다고 하였음

<그림 6> 일기예보모델(위)과 과거데이터 기반 딥러닝 추론결과(아래)



Source: Can Machines Learn to Predict Weather? Using DeepLearning to Predict Gridded 500-hPa Geopotential Height From Historical Weather Data (Jonathan A. Weyn et al., 2019)

○ (Oral Session 4) New Benchmarks for Machine Learning in Earth Observation

1) "OpenSentinelMap: A Large-Scale Land Use Dataset using OpenStreetMap and Sentinel-2 Imagery", Noah Johnson (Vision Systems, Inc.); Wayne Treible (Vision Systems, Inc.); Daniel E Crispell (Vision Systems, Inc.);

- OpenStreetMap 데이터에서 파생된 도로, 건물, 물 그리고 12종류의 토지이용분류에 대한 정보를 라벨로 하여, Sentinel-2 위성 영상을 토지이용 데이터셋으로 구축한 사례

2) "Hephaestus: A large scale multitask dataset towards InSAR understanding", Nikolaos I Bountos (National Observatory of Athens); IOANNIS PAPOUTSIS (National Observatory of Athens); Dimitrios Michail (Harokopio University of Athens); Andreas Karavias (Harokopio University of Athens); Panagiotis Elias (National Observatory of Athens); Isaak Parcharidis (Harokopio University of Athens);

- InSAR(Interferometric SAR)는 지질학, 인공구조물의 지질공학적 특성에 대한 고유한 정보를 제공해주는 중요한 데이터이지만, 대규모 데이터셋을 만들기 어려움
- 데이터를 생성하는 과정에 드는 비용이 높으며, 데이터 해석이 가능한 전문가 팀이 필요한 등 제약조건이 많음
- 위 연구에서는 전 세계의 서로다른 44개의 화산에서 얻어진 216,106개의 분할 InSAR 영상에서 19,919개의 주석을 포함하는 데이터셋을 생성하였음

### 3. 2022.06.20.(월) Tutorials

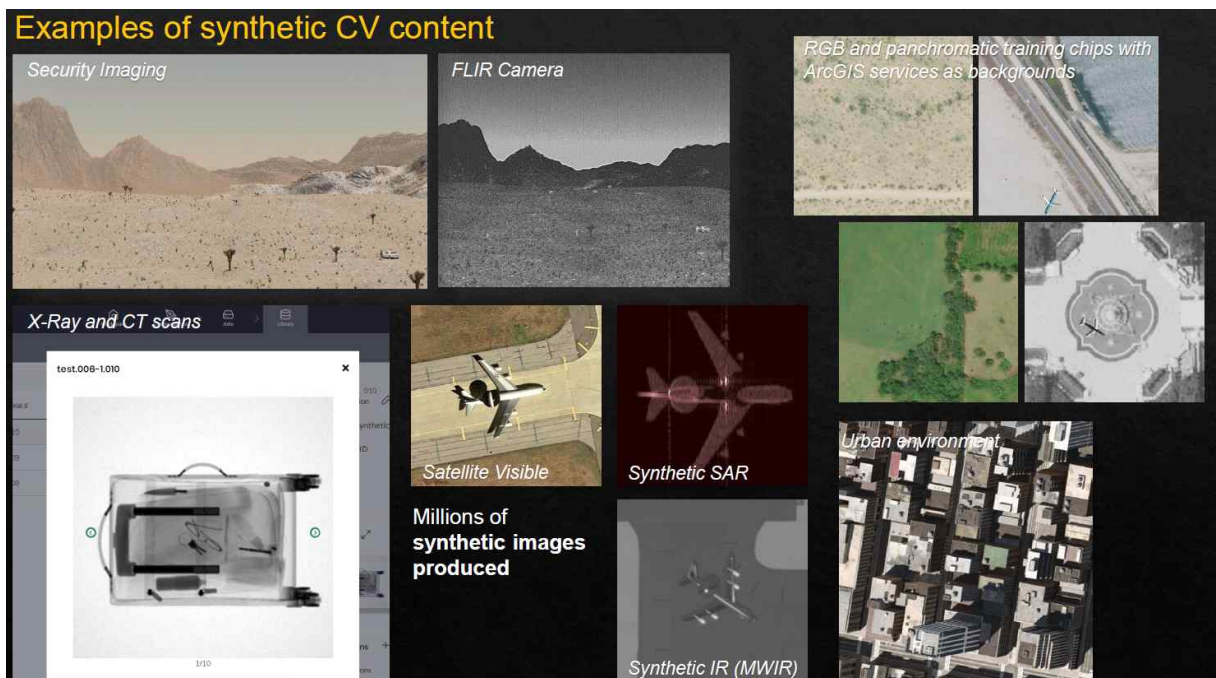
□ 일시 및 장소: 2022. 06. 20(월) 09:00 Ernest N. Morial Convention Center

□ (Tutorial) Creating and Using Synthetic Data for Computer Vision Applications - Rendered.ai

□ 주요 발표내용

#### ○ Synthetic Data

- 데이터 분석과 연구는 실제 물리적 센서 및 센서 플랫폼을 사용하여 취득한 실제 데이터를 활용하는 경우가 많지만, 개인 정보, 보안 및 안전 문제 등 다양한 이유로 실제 데이터를 취득하기 어려운 경우가 많음
- 합성 데이터(Synthetic Data)는 실제 데이터를 취득하기 어려운 경우, 또는 설계/개발 중인 사양에 맞는 센서데이터를 포함하여 특정 요구사항에 맞게 데이터 세트를 설계하고 생성해주는 일종의 가상데이터라고 할 수 있음
- Rendered.ai 에서는 이러한 합성데이터를 생성해볼 수 있는 플랫폼 서비스를 제공하고 있으며, RGB 광학영상 및 적외선, SAR 등 다양한 센서데이터의 생성이 가능

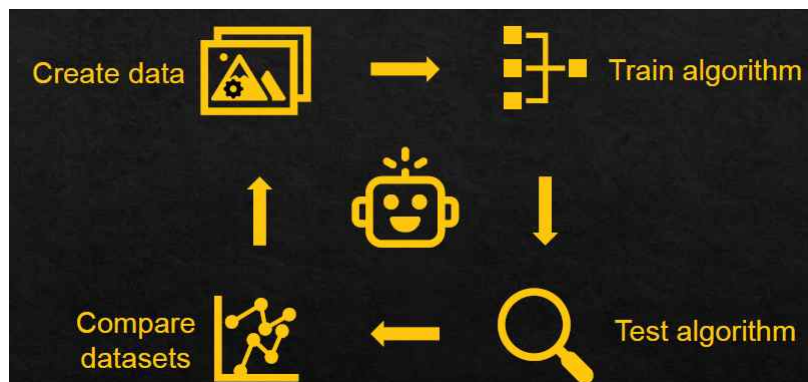




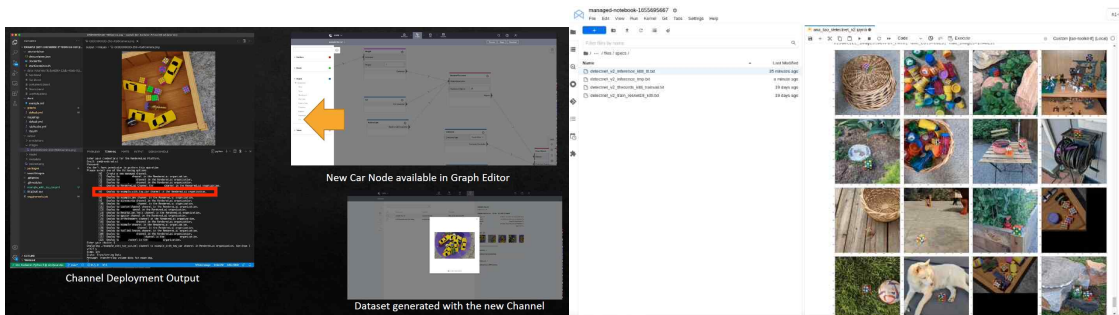
- 기존의 인공지능 워크플로우는 데이터를 습득하거나 찾아서 알고리즘을 훈련시키고 테스트하는 방식이었다면,



- 합성 데이터는 데이터를 만들고 테스트하여 데이터셋을 최적화 하는 방향으로 발전할 수 있도록 도움이 된다고 함



- 튜토리얼 진행을 통해, 실제로 합성데이터를 생성해볼 수 있었음



#### 4. 2022.06.21.(화) Main Conference & Exhibition

□ 일시 및 장소: 2022. 06. 21(화) 09:00 Ernest N. Morial Convention Center

□ (Exhibition) 주요 전시 내용

- 전시공간에는 Tesla, Apple, Microsoft, Meta 등 AI & 딥러닝 분야의 가장 대표적인 기업들을 포함하는 약 100여개의 기업이 참여하였으며, 국내기업에서도 네이버, FuriosaAI 등이 참여
- Tesla, Zoox, Motional, Cruise 에서는 주로 차량의 자율주행에 활용되고 있는 기술들에 대해 소개하고 있었으며, Amazon에서는 소형 배달 로봇에 적용중인 자율주행기술을 소개
- 국내기업은 FuriosaAI에서는 자체개발한 NPU(Neural Processing Unit)를 전시 및 소개하고 있었으며, 네이버에서는 자사의 음성인식기술, 웹툰관련 AI 기술을 소개하였음
- LG에서는 운전중인 운전자의 얼굴 및 시선을 인식하여, 졸음운전 등을 방지해주는 운전자 모니터링 시스템을 소개
- SuperbAI에서는 AI 데이터셋의 자동 라벨링, 데이터 어노테이션 기술 등을 소개하였으며, 이 솔루션을 이용해 데이터 중복제거, 균등화된 랜덤샘플링, 데이터 추천에 대한 데모 시현
- Meta에서는 Project Aria Glasses를 소개하며, 사용자 참여 기반의 AR 데이터 수집 프로젝트에 대해 설명, 장치의 센서는 착용자의 비디오 및 오디오는 물론 시선 추적 및 위치 정보를 캡처하여 암호화된 서버로 전송하고, Meta는 이 데이터를 연구 및 활용
- 이 외에도 학습 모델 최적화 및 경량화, 영상 필터 및 안면인식, 주변환경 인식과 자동 데이터 수집 및 라벨링 기술 등 AI/딥러닝 분야의 최신 기술을 견식할 수 있었음
- 주요 사진

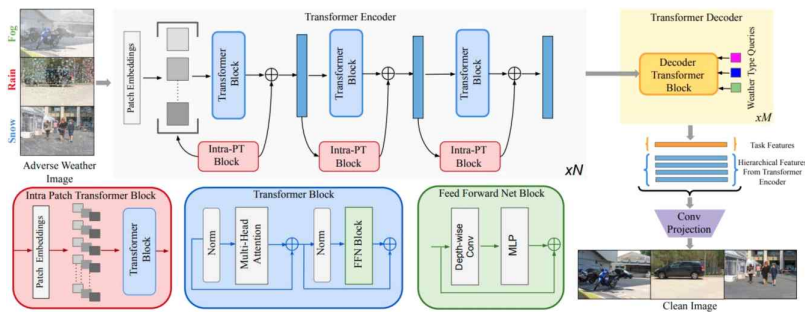




□ (Main Conference) Poster Session

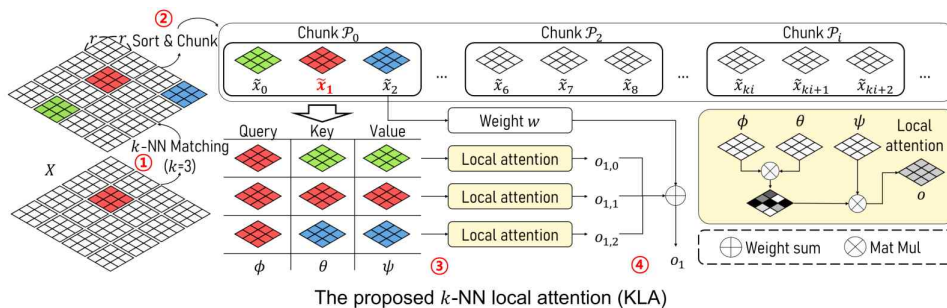
○ TransWeather: Transformer-Based Restoration of Images Degraded by Adverse Weather Conditions

- 눈, 비, 안개 등 기상에 따라 발생하는 영상의 노이즈를 제거하기 위해 최근에는 모든 기상 조건을 한 번에 제거하기 위해 신경 아키텍처 검색(All-in-One)을 사용하는 CNN 기반 방법이 제안되어왔음
- 위 연구에서는 모든 악천후 제거 문제에 대한 효율적인 방법을 개발하기 위해 인코더와 디코더가 하나뿐인 트랜스포머 기반의 end-to-end 모델을 제안하고, 이를 활용하여 기상조건에 의해 열화된 영상을 복원하는 방법을 제시 하였음



○ KNN Local Attention for Image Restoration

- k-means 클러스터링 아이디어를 CNN 트랜스포머에 적용하여 영상의 노이즈 제거, 디블러링 등을 수행하였음
- 위 연구에서는 벤치마크 테스트를 통해 다른 복원방식에 비해 더 뛰어난 성능을 가진다고 하였음



- 디지털 국토정보 구축 효율화를 위한 다차원/다시점 공간 데이터 기반 국토정보 변화인식 및 자동갱신 기술개발 과제에서 개발하게 될 클라우드 소싱 데이터의 품질 향상 기술에 상기 논문에서 제안한 접근법의 적용 가능성을 검토할 예정

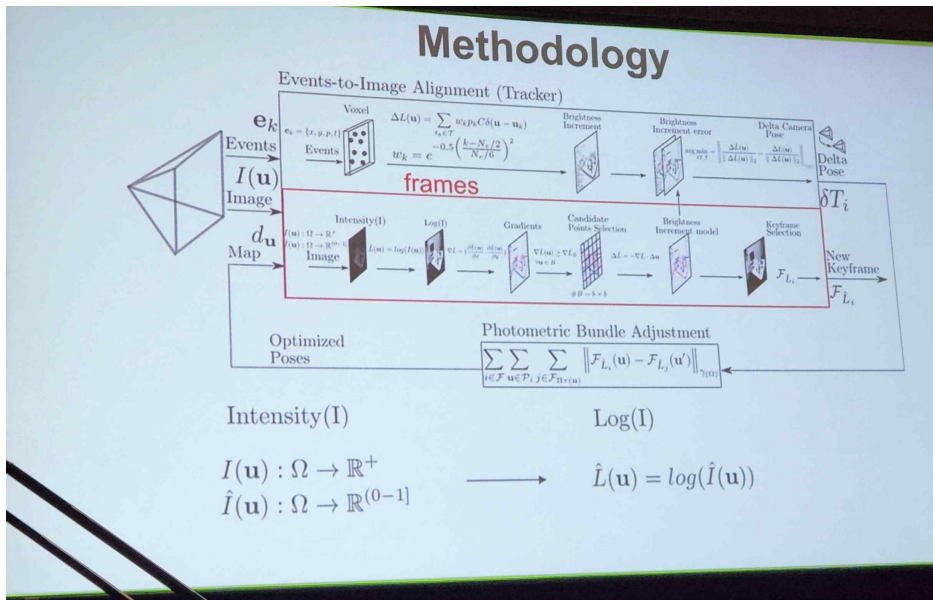
5. 2022.06.22.(수) Main Conference & Expo

□ 일시 및 장소: 2022. 06. 22(수) 09:00 Ernest N. Morial Convention Center

□ 주요 발표내용

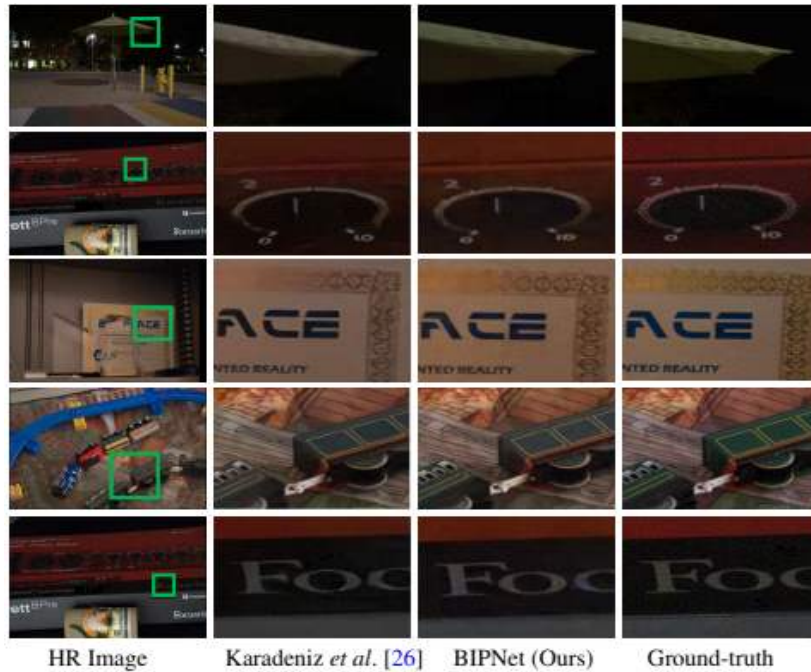
○ (Session 1) Low-Level Vision

- Minsu Cho(POSTECH), Kavita Bala(cornell University) 좌장으로 18개 발표가 진행됨
- 영상에 존재하는 그림자 제거, 조도가 낮을 때 촬영된 어두운 영상의 밝기 향상, 해상도가 낮은 영상을 초고해상도 영상으로 변환, 영상 스티칭, 영상과 라이다 데이터 융합, 열화상과 광학 영상의 융합, 이벤트 신호를 이용한 카메라 모션 추정 등 영상의 품질 향상과 위치 추정과 관련된 low-level 컴퓨터 비전 연구들이 발표됨
- “Event-Aided Direct Sparse Odometry”는 이벤트 카메라를 이용하여 연속적으로 취득된 영상 프레임 간의 블라인드 타임에 밝기값의 변화를 이용하여 카메라 모션을 추적하는 방법론을 제안하였고, 이는 기존 방법과 다르게 on-demand 상황에서 저전력으로 동작할 수 있어 활용성이 높을 것으로 판단됨



- 영상 데이터의 위치 및 자세를 결정할 수 있으므로 디지털 국토정보 구축 효율화를 위한 다차원/다시점 공간 데이터 기반 국토정보 변화인식 및 자동갱신 기술개발 과제에서 크라우드 소싱 데이터 확보 수단으로 이벤트 카메라를 적용할 수 있음
- 특히, “Burst Image Restoration and Enhancement”에서 소개된 burst super-resolution, burst low-light image enhancement, burst denoising으로 이루어진 영상 복구 및 향상

프레임워크는 낮은 품질의 영상에 대하여 식별력을 증대시키는 것에 있어서, 우수한 성능을 보여주었음



- 디지털 국토정보 구축 효율화를 위한 다차원/다시점 공간 데이터 기반 국토정보 변화인식 및 자동갱신 기술개발 과제에서 개발하게 될 클라우드 소싱 데이터의 품질 향상 기술에 상기 논문에서 제안한 접근법의 적용 가능성을 검토할 예정

○(Session 2) Transfer / Low-Shot / Long-Tail Learning

- Lorenzo Torresnani(Dartmouth College), Mei Chen(Microsoft), Ehsan Adeli(Stanford University) 좌장으로 15개 발표가 진행됨
- 사전 학습된 대형 모델을 새로운 대상 문제에 적용할 때, 추가적인 학습 데이터 없이(zero shot) 또는 아주 작은 데이터만을 학습시키고(few shot)도 추론의 정확도와 일관성을 확보하기 위한 연구와 대상 도메인이 바뀔 때 기존 도메인에 학습시킨 결과를 잇는 롱테일(long-tail) 효과를 없애기 위한 연구들이 발표됨
- “Balanced MSE for Imbalanced Visual Regression”에서는 연령 추정이나 포즈 추정 등의 시각적 회귀모델에서 데이터의 불균형은 모델의 일반화 및 성능에 영향을 끼치게 되는데 학습 데이터의 불균형 분포를 수용할 수 있는 새로운 손실함수로서 Balanced MSE를 제안함
- “Robust fine-tuning of zero-shot models”에서는 사전 훈련된 대규모 모델을 제로샷 추론에 적용하면 데이터 분포 범위에 일관된 정확도를 얻을 수 있지만, 미세조정을 하면 주어진

문제에 정확도를 향상시키지만 분포 이동에 대한 견고성이 떨어진다는 단점이 존재하여 제로샷과 미세 조정 모델의 앙상블을 제안하여 대상 문제의 정확도와 분포 이동에 대한 견고성을 모두 성취할 수 있도록 함

- “Learning to learn and remember super long multi-domain task sequence”에서는 사전 학습된 모델이 적용되는 도메인이 바뀌면, 이전 도메인에서 학습시킨 결과를 잊게 되는 롱테일 현상 해결하기 위해 메타 옵티마이저를 제안하였고, 10개의 서로 다른 도메인에 대해 포괄적으로 실험한 결과 도메인에 구애받지 않고 메타 학습이 가능한 것을 보여줌

University at Buffalo  
JD EXPLORE ACADEMY

Learning to Learn and Remember Super Long Multi-Domain Task Sequence  
Zhenyi Wang, Li Shen, Tiegang Duan, Donglin Zhan, Le Fang, Mingchen Gao

#### Introduction

Most existing meta-learning focuses on learning a single domain, but sequential domain shift frequently happens in real-world applications.

How to avoid forgetting?

Meta-learning model

Sequential domain meta-learning (SDML): the meta-learning model is required to learn a sequence of domains without forgetting the knowledge in previous domains.

#### Problem Definition

**Definition 1. non-stationary heterogeneous domains.** A sequence of domains,  $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_J$ , arrive sequentially. Each domain  $\mathcal{D}_i$  is represented as a labeled dataset  $\{(x^i, y^i)\}_{i=1}^J$  with  $I_i$  labeled datapoints; where  $x^i$  are the datapoints and  $y^i$  are the labels. All the domains do not share class labels.  $\mathcal{D}_1, \mathcal{D}_2, \dots, \mathcal{D}_J$  are called non-stationary heterogeneous domains.

**Domain-aware SDML:** Domain identities and boundaries are known during training.

**Domain-agnostic SDML:** Domain identities and boundaries are unknown during training.

**Goal:** Learn new domains without forgetting previous knowledge.

#### Methodology

**Intuition:**  
More important parameters should be updated slower to avoid forgetting.  
Less important parameters could be updated faster for efficient learning of the current domain.

**Solution:**  
Maintain a small memory buffer to store a small number of training tasks from previous domains.  
Propose a meta optimizer to adaptively and separately adjust the learning rates for each meta parameter.

#### Training Algorithm

```

for  $t = 1$  to  $N_T$  do
  update parameters  $\theta$  by meta training on  $\mathcal{T}_t$ 
   $\theta_{t+1} = \theta_t - \lambda_t \frac{\partial \mathcal{L}(\theta; \mathcal{T}_t)}{\partial \theta}$ 
   $\lambda_{t+1} = \lambda_t - \eta \frac{\partial^2 \mathcal{L}(\theta; \mathcal{T}_t)}{\partial \theta^2}$ 
  Reservoir sampling to update task memory  $\mathcal{M} \leftarrow \mathcal{M} \cup \mathcal{T}_t$ 
  if decided to store the task  $\mathcal{T}_t$ 
end for
    
```

#### Domain Shift Detection

In domain agnostic SDML, domain boundaries are unknown. We need to detect when domain shift happens.

We design a low-dimensional latent space to capture latent domain change.

We apply Bayesian online changepoint detection (BOCPD) to detect domain shifts in this latent space.

#### Experiments

Construct a large-scale dataset consisting of a sequence of 10 domains (datasets) with sequential domain shift.

Randomly sample 10K tasks from each dataset and construct a super long 100K tasks.

Equip the meta optimizer with both MAML(ANIL) and Prototypical Network (PNNet).

Compared to SOTA continual-learning methods.

The following table shows the results for 5-way 5-shot learning (integrating the meta optimizer with PNNet and ANIL).

Algorithm	5-way 5-shot	
	ACC	BWT
PNNet-Sequential	46.83 ± 0.10	-22.95 ± 0.12
PNNet-EWC	49.88 ± 0.15	-14.51 ± 0.14
PNNet-HAT	50.25 ± 0.26	-16.32 ± 0.28
PNNet-UCB	49.06 ± 0.22	-15.83 ± 0.20
PNNet-A-GEM	49.21 ± 0.31	-20.01 ± 0.39
PNNet-RS	49.56 ± 0.18	-18.87 ± 0.19
PNNet-MER	50.38 ± 0.24	-15.10 ± 0.24
PNNet-DEGCL	50.79 ± 0.37	-13.82 ± 0.45
PNNet-GPM	49.73 ± 0.51	-14.91 ± 0.58
Ours	<b>55.28 ± 0.19</b>	<b>-11.15 ± 0.27</b>
Joint-training	66.32 ± 0.18	N/A

Algorithm	5-way 5-shot	
	ACC	BWT
ANIL-Sequential	45.85 ± 0.46	-23.47 ± 0.43
ANIL-EWC	45.45 ± 0.29	-21.99 ± 0.34
ANIL-HAT	40.58 ± 0.19	-28.89 ± 0.24
ANIL-UCB	47.21 ± 0.28	-20.18 ± 0.22
ANIL-A-GEM	48.08 ± 0.33	-20.30 ± 0.35
ANIL-RS	46.97 ± 0.27	-21.37 ± 0.33
ANIL-MER	47.96 ± 0.52	-19.25 ± 0.50
ANIL-DEGCL	47.91 ± 0.45	-18.57 ± 0.53
ANIL-GPM	47.73 ± 0.53	-19.76 ± 0.46
Ours	<b>51.56 ± 0.21</b>	<b>-16.07 ± 0.20</b>
Joint-training	68.16 ± 0.11	N/A

- 학습데이터와 테스트데이터의 도메인이 상당히 다르더라도 잘 예측할 수 있도록 하는 것이 domain adaptation이고, 다양한 시공간에서 취득된 데이터를 이용한 공간정보분야에서의 딥러닝 연구는 domain adaptation에 대해서 꼭 고려해야 함
- “DINE: Domain Adaptation from Single and Multiple Black-box Predictors”에서는 대상 데이터 구조를 고려하여 소스 예측자의 지식을 맞춤형 대상 모델로 추출한 다음 대상 도메인에 더 적합하도록 추출된 모델을 미세 조정하는 DIstill 및 fine-tuNE(DINE)라는 2단계 지식 적용 프레임워크를 제안함
- 제안된 방법은 도메인 전체에서 신경망이 동일할 필요가 없으며 리소스가 적은 장치에서도 효과적인 적응이 가능하여 이전 및 관련 레이블이 지정된 데이터 세트(소스)의 지식을 레이블이 없는 새로운 데이터 세트(대상)로 이전하는 것을 목표로 하는 UDA(Unsupervised Domain Adaptation)에 경쟁력 있는 것으로 보임



## Proposed Method (Distill and Fine-tune, DINE)

### A. Adaptive Self-Knowledge Distillation

$$AdaLS(p, r) = \begin{cases} p_i, & i \in T_p^r \\ \frac{1 - \sum_{j \in T_p^r} p_j}{K - r}, & \text{otherwise.} \end{cases}$$

Initialization

$$P^T(x_t) \leftarrow \gamma P^T(x_t) + (1 - \gamma) f_t(x_t) \quad \longrightarrow \quad \mathcal{D}_{kl}(P^T(x_t) || f_t(x_t))$$

$T_p^r$  denotes the index set of top-r classes in source predictions, and  $K$  denotes the size of  $p$ .

### B. Distillation with Structural Regularizations

$$\mathcal{L}_{mix} = l_{ce}(Mix_\lambda(f_t^i(x_i^t), f_t^j(x_j^t)), f_t(Mix_\lambda(x_i^t, x_j^t)))$$

$$Mix_\lambda(a, b) = \lambda a + (1 - \lambda)b$$

$$\mathcal{L}_{im} = \mathcal{H}(\mathbb{E}_{x_t \in X_t} f_t(x_t)) - \mathbb{E}_{x_t \in X_t} \mathcal{H}(f_t(x_t))$$

$$\mathcal{H}(p) = -\sum_i p_i \log p_i$$

← pairwise structural distillation

← global (batch-level) structural distillation

○ (기조 강연) Toward Integrative AI with Computer Vision

- 발표자: Microsoft Azure AI의 Xuedong Huang
- 빅 데이터, 컴퓨팅 성능 및 최신 네트워크 아키텍처 덕분에 지난 10년 동안 AI는 빠르게 발전하고, 사람들의 일상 생활에도 침투하고 있으며, 앞으로 잠재력이 무궁무진함
- 현재 AI는 음성 인식, SuperGLUE 및 이미지 캡션과 같은 몇 가지 잘 정의되고 좁은 범위의 연구에서는 인간과 유사한 성능을 보여주지만, 개방적이고 복잡한 것을 이해하는 데 있어서는 아직 한계가 존재함
- AI가 인간의 지능을 모방하기 위해서는 세계 각 AI 작업이 역할을 나눠서 이루어져야 하며, 그 결과들이 공유되어야 함
- 분야마다 대량의 데이터를 각각 구축하고 있어, 지식을 공유하고 이전하는 과정을 통해 시너지를 기대하기 어려운 상황임
- 마이크로소프트에서는 음성, 언어 및 시각의 다양한 작업을 통합하기 위해 총체적이고 의미론적인 표현을 사용하여 이 문제를 해결하는 다국어, 다중모드 접근이 가능한 통합 AI를 추진
- 최근에는 통합 AI를 컴퓨터 비전에 적용하기 위해 대규모 이미지와 언어 사전 훈련을 통해 의미 계층 개념을 도입한 플로렌스를 개발하였고, 시각적 지식과 추론을 이미지 및 텍스트 변환기로 추출하여 인식, 감지, 분할 및 캡션과 같은 일반적인 컴퓨터 비전 작업을 위한 제로샷 및 퓨어샷 기능을 가능하게 함
- 마이크로소프트와 같은 거대 기업에서는 플로렌스와 같은 통합된 고도화된 범용 AI API를 제공하고, 중소기업, 연구진 등에서는 이를 이용하여 특정 분야의 특정 문제를 해결하지는 의견을 제시하였으나, 합리적이지는 하지만 너무 대기업에 종속되는 것은 아닌지 우려됨





## 6. 2022.06.23.(목) Main Conference & Expo

□ 일시 및 장소: 2022. 06. 23(목) 09:00 Ernest N. Morial Convention Center

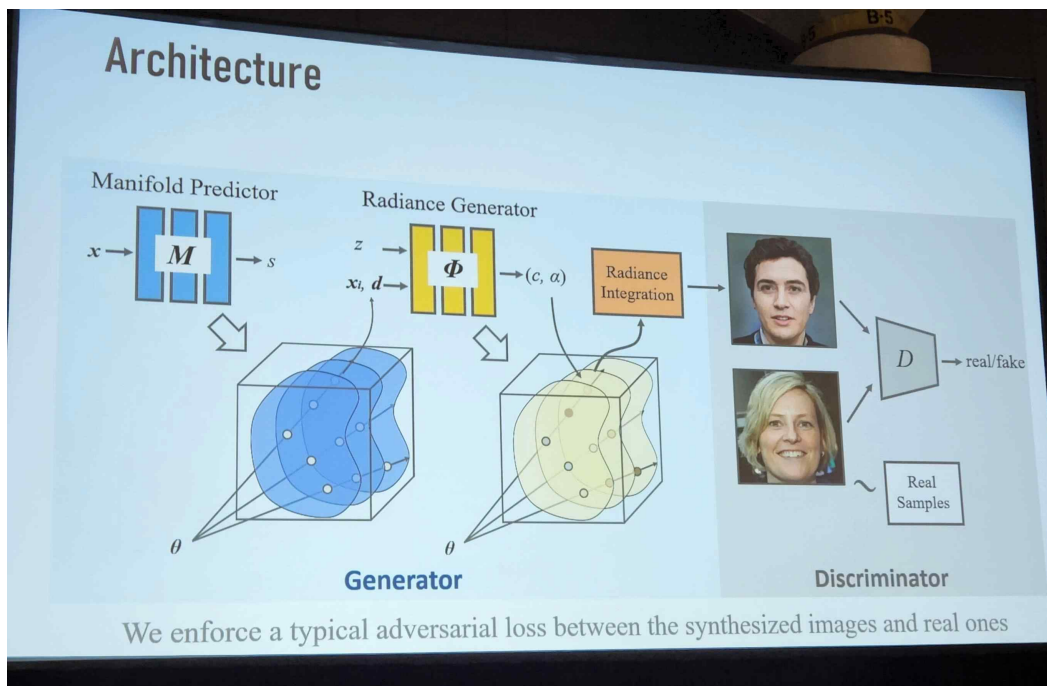
□ 주요 발표내용

### ○(세션 3) Image & Video Synthesis and Generation(I)

- Sharon Xiaolei Huang (Pennsylvania State University), Shaodi You (University of Amsterdam) 좌장으로 15개 발표가 진행됨
- 영상 인코딩 & 디코딩, GAN을 이용한 영상 합성 및 생성, 네트워크의 품질 및 다양성 제어, 3D 인식 이미지 생성, 고해상도 이미지 합성, 텍스트-이미지 합성, 주야간 이미지 합성, 영상 마스킹, 양질의 데이터셋 확보를 위한 distillation, continual learning 등 주어진 영상으로부터 예측을 통해 새로운 영상을 만드는 것과 관련된 연구들이 발표됨
- “Marginal Contrastive Correspondence for Guided Image Generation”에서는 영상 분할에 대한 조건과 예시 스타일을 입력으로 주면 두 영상 사이의 조밀한 대응을 설정하여 조건에 따라 예시 스타일을 적용하여 새로운 영상을 생성하여 주는 대조 학습 네트워크(MCL-Net)를 제안함
- 건축 설계나 리모델링 등에 소비자가 원하는 스타일을 가져오면, 스타일에 맞춰 현재 건물에 대한 리모델링 후의 모습, 또는 건축 후의 모습에 대한 청사진 생성에 유용할 것으로 기대됨

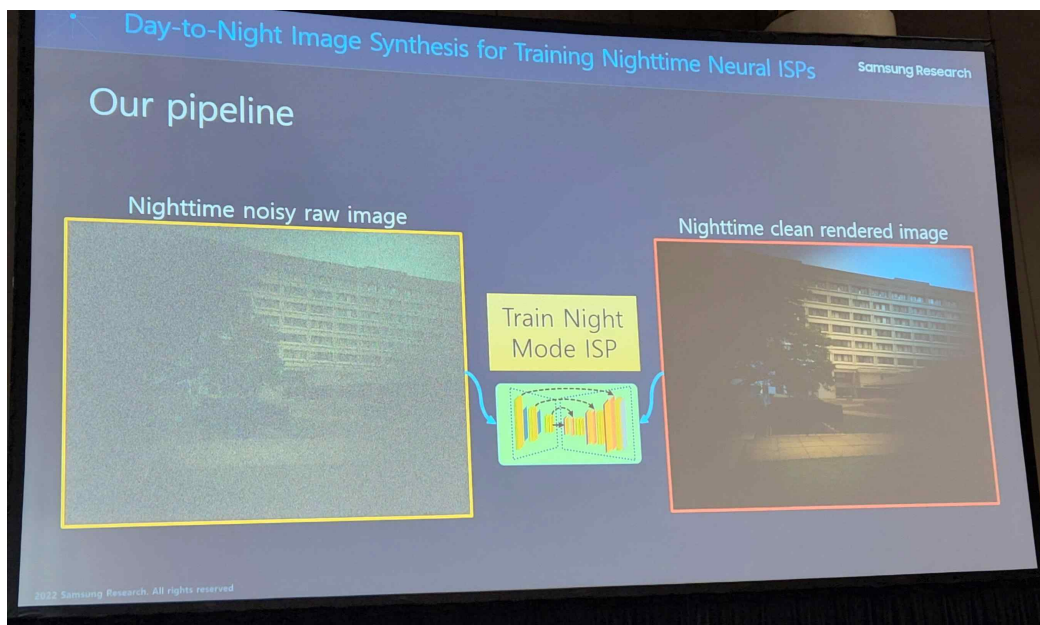


- 특정 객체에 대하여 다양한 방향에서 취득될 수 있는 이미지를 생성하는 3D-Aware 이미지 모델링은 체적 표현에 많은 연산과 메모리가 필요하기 때문에 radiance 구현을 위한 샘플링 수를 제한하여 사실적인 이미지 생성에 한계가 존재함
- “GRAM: Generative Radiance Manifolds for 3D-Aware Image Generation”에서는 3D 볼륨에서 학습된 암시적 표면 세트로부터 구현된 2D 매니폴드에서 포인트 샘플링 및 radiance field learning을 조절하는 새로운 접근 방식을 제안하여, 각 샘플 광선에 대해 광선-표면 교차를 계산하고 네트워크에서 생성된 복사 휘도를 누적하고 이러한 광도 매니폴드를 훈련하고 렌더링함으로써 사실적인 디테일을 가진 고품질 이미지를 생성할 수 있음



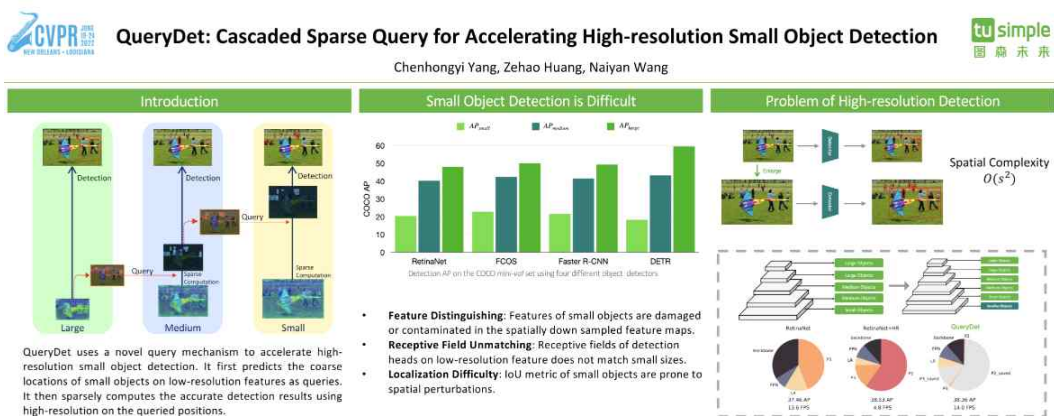
- 딥러닝 기술은 대량의 학습데이터가 필요하다는 한계점이 있으며, 훈련된 모델이 전체 데이터 세트에 대해서 훈련된 모델의 테스트 정확도와 일치하기 어렵기 때문에 이를 극복하기 위해 작은 훈련 데이터 세트를 합성하는 데이터 세트 증류를 수행함
- “Dataset Distillation by Matching Training Trajectories”에서는 많은 훈련 단계에서 실제

- 데이터로 훈련된 것과 유사한 상태로 네트워크를 안내하기 위해 증류된 데이터를 최적화하는 방법을 제안함. 네트워크가 주어지면 증류된 데이터에 대해 여러 번 반복하여 훈련하고 훈련된 매개변수와 실제 데이터에서 훈련된 매개변수 사이의 거리와 관련하여 증류된 데이터를 최적화하게 됨. 대규모 데이터 세트에 대한 초기 및 대상 네트워크 매개변수를 효율적으로 얻기 위해 실제 데이터 세트에서 훈련된 전문가 네트워크의 훈련 궤적을 미리 계산하고 저장하여 사용함
- “Mask-Aware Transformer for Large Hole Image Inpainting”에서는 고해상도 이미지를 효율적으로 처리하기 위해 트랜스포머와 컨볼루션의 장점을 통합한 대형 홀 인페인팅을 위한 새로운 트랜스포머 기반 모델을 제안하였고, 이는 공간정보분야에서 다루는 고해상도의 훼손된 영상 복원에 적용할 수 있을 것으로 판단됨
  - “Day-to-Night Image Synthesis for Training Nighttime Neural ISPs”에서는 스마트폰으로부터 야간에 취득된 이미지에 잡음, 모션 블러 등이 발생하기 쉬어, 이를 해결하기 위해 노이즈가 적은 주간 이미지와 야간 이미지를 합성하여 야간 이미지를 개선하는 방법을 제안함
  - 디지털 국토정보 구축 효율화를 위한 다차원/다시점 공간 데이터 기반 국토정보 변화인식 및 자동갱신 기술개발 과제에서 입력 영상으로 고려하는 5종의 데이터에 스마트폰으로 취득된 영상도 포함되며, 일반 사용자가 임의의 시간에 촬영한 영상일 수 있으며, 이에 상기 방법을 적용 가능할 것으로 판단됨



○ (세션 4) Reognition, Learning for Vision and Robot Vision

- Kris Kitani (Carnegie Mellon Univ.), R. Venkatesh Babu (Indian Inst. of Science), Noha Radwan (Google) 좌장으로 15개 발표가 진행됨
- 딥러닝을 적용한 객체 및 장소 인식 방법론을 중심으로 인식 정확도 향상, 인식 모델 훈련의 가속화, 소형 물체의 인식 가속화, 모르는 물체 감지 등 성능 향상을 위한 연구와 쌍별 제약 조건으로 신경망 보정, 신경망 해석을 위한 인과 관계 기반 잠재 변수 모델링, 로봇 조작을 위한 학습, 실시간 맵-뷰 시맨틱 분할을 위한 연구들이 발표됨
- 객체 인식에 있어서 딥러닝 적용은 상당히 일반화되었지만, 작은 크기의 객체 인식에 있어서는 아직 고해상도 이미지나 피쳐맵을 사용하여 극복하고 있으며, 이 경우 이미지 및 피쳐맵의 사이즈가 증가함에 따라 계산량이 제곱으로 증가함
- “QueryDet: Cascaded Sparse Query for Accelerating High- Resolution Small Object Detection”에서는 고해상도 이미지의 피쳐맵을 최대한 효율적으로 활용하기 위해 피쳐 피라미드를 구성하여 이를 기반으로 객체 인식 속도를 가속화하는 QueryDet를 제안함
- QueryDet은 먼저 저해상도 피쳐에서 작은 물체의 대략적인 위치를 예측한 다음 이러한 대략적인 위치에 의해 안내되는 고해상도 피쳐를 사용하여 탐지를 수행함으로써, 고해상도 피쳐맵의 이점을 얻으면서 불필요한 계산을 줄이게 됨
- 저고도에서 촬영된 드론 영상, 스마트폰 영상, 지상 차량에서 획득된 영상 등 작은 객체를 많이 포함하는 근접 촬영된 영상 데이터 셋에 적용함으로써 효과적인 객체 인식을 수행할 수 있을 것으로 예상됨









○(기조 강연 2) Understanding Visual Appearance from Micron to Global Scale

- 발표자: Cornell University의 Kavita Bala 교수
- 증강 현실/혼합 현실(AR/MR)은 컴퓨터 비전과 컴퓨터 그래픽을 결합하여 매력적이고 몰입도 높은 사용자 경험을 만들고 있음
- 사실적인 시각적 모습과 렌더링을 위한 그래픽 모델, 모양과 재료의 재구성, 장면 이해를 위한 시각적 검색 및 인식과 같은 보완적인 영역에 대해 수행한 연구를 소개함
- 특히, 옷감의 텍스처와 같은 마이크로한 스케일에서부터 행성과 같은 글로벌 스케일까지 영상을 기반으로 시각적 모습을 이해하는 사례에 대한 소개는 매우 흥미로움
- 특히, 핵심 빌딩 블록으로 인식을 사용하여 전 세계적으로 이미지의 시각적 발견이 지리와 시간에 걸쳐 시각적 패턴과 경향을 발견할 수 있는 방법을 소개함
- 예를 들어, 전세계 스트리트 영상을 수집하여, 옷차림을 살펴 클러스터링을 하여 살펴보면, 아웃라이어가 발생하는 시점에 특별한 옷차림을 하는 이벤트가 있거나, 특정 아이템이 유행하는 현상 등을 발견할 수 있음
- 우리가 사는 세상에 대해 마이크로 스케일부터 글로벌 스케일까지의 시각적 이해를 실현하기 위해 섬유 과학자, 원단 디자이너, 패션 디자이너, 산업 디자이너, 문화 인류학자, 농업 과학자 등 아주 다양한 전문가와 해당 전문 지식을 융합되어야 함을 느낌



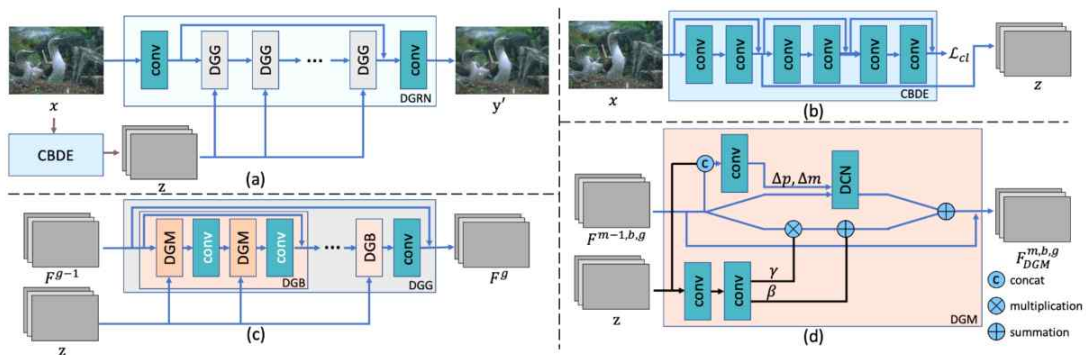
## 7. 2022.06.24.(금) Main Conference

□ 일시 및 장소: 2022. 06. 24(금) 09:00 Ernest N. Morial Convention Center

□ 주요 발표내용

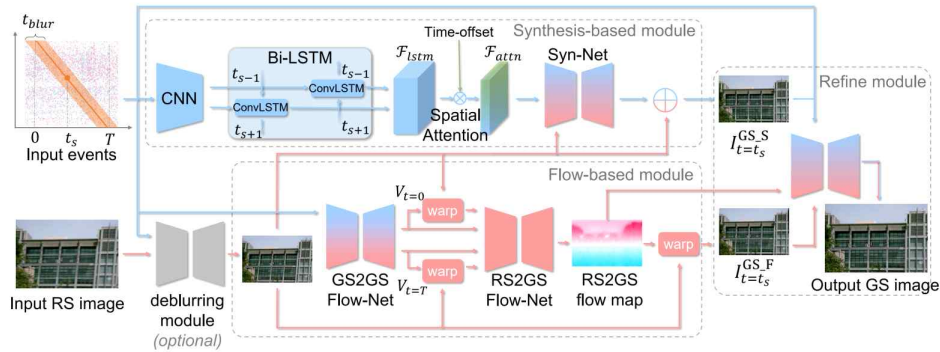
### ○ (Poster Session 4.1) Low-Level Vision

- “All-in-One Image Restoration for Unknown Corruption”에서는 알려지지 않은 다양한 손상 유형 및 수준에서 이미지를 복구할 수 있는 일체형 방법론을 개발하는 연구를 수행하였음
- 이를 위해 CBDE(Contrastive-Based Degraded Encoder)와 DGRN(Degradation-Guided Restoration Network)이라는 두 개의 신경망 모듈로 구성된 일체형 이미지 복원 네트워크인 AirNet을 제안하였음
- 연구에서는 하나의 신경망에서 다양한 손상 이미지를 복구할 수 있고, 손상 유형 및 손상 정도에서 자유로운 장점이 있다고 하였으며, 손상에 대한 사전 정보를 알기 어렵고 성능 저하가 공간과 시간에 따라 변하는 실제 시나리오에서 더 나은 유연성을 가진다고 하였음

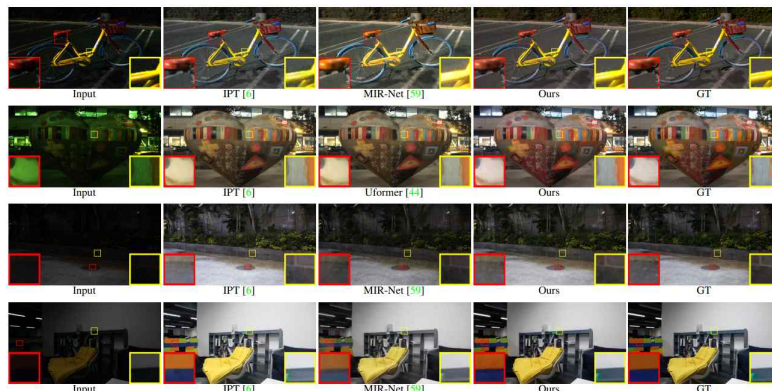
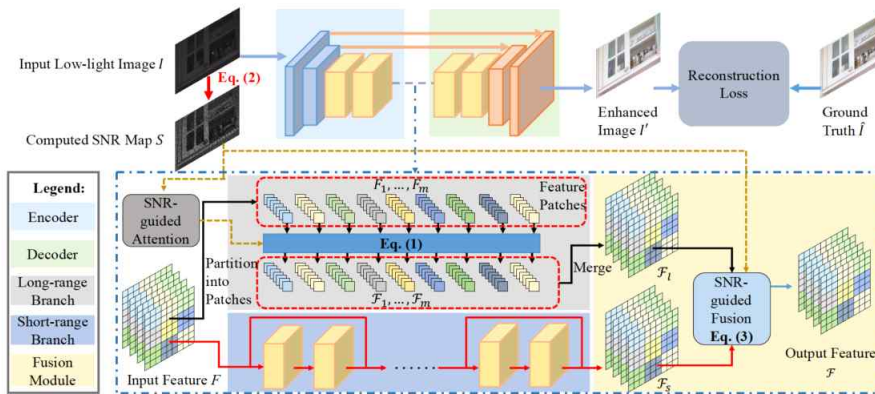


### ○ (Poster Session 4.1) Computational Photography

- 드론이나 스마트폰의 카메라의 경우 롤링 셔터(Rolling-shutter) 방식으로 작동하며, 물체의 속도나 센서의 진동 때문에 젤로 현상(jello effect)이 발생하고, 이미지가 뭉개지는 문제가 발생함
- “EvUnroll: Neuromorphic Events Based Rolling Shutter Image Correction”에서는 문제를 해결하고자 RS 센서와 이벤트 센서로 구성된 새로운 컴퓨터 이미징 설정을 소개하고 EvUnroll이라는 신경망을 제안함
- 이벤트 센서를 사용하여 일반 카메라에 쓰이는 글로벌 셔터와 융합하는 모듈을 설계하고, 융합된 하이브리드 카메라 시스템으로 취득한 영상들을 학습데이터로 활용하여 신경망을 테스트하였으며, 좋은 성능을 내는 것으로 결과를 도출하였음



- “SNR-Aware Low-Light Image Enhancement”에서는 신호 대 잡음비(SNR, Signal-to-Noise-Ratio)를 두가지 측면에서 접근하여 트랜스포머(transformer)와 CNN 모델을 활용하여 저조도 이미지의 품질을 향상시키는 새로운 방법을 제시하고 있음
- 기존의 저조도 영상 품질을 향상시키는 연구들에서는 은닉층에서 컨볼루션 구조를 채택하였으나 노이즈가 많은 영상의 경우 그 효과가 부정확한 경우가 있기 때문에 이 연구에서는 영상 내 서로다른 영역의 픽셀이 상호 보완적으로 작용하여 품질을 향상시키는 방법을 채택하였음
- 광범위한 실험을 통해 연구의 프레임워크가 동일한 구조를 가진 다른 7개의 대표적인 접근 방식보다 일관되게 더 나은 성능을 보여주고 있으며, 향후 영상품질개선 기술 개발에 적용방안에 대해 검토 예정



## IV. 붙임자료(항공권)

최경아 부연구위원

■ 승객성명 Passenger Name  
CHOI/KYOUNGAH MS

1

서울  
ICN

Incheon international  
Terminal No: 2  
17JUN22(금)14:30 (Local Time)



로스엔젤레스  
LAX

Los angeles intl  
Terminal No: B  
Tom Bradley Int'l  
17JUN22(금)09:40 (Local Time)

**KE0017**  
예약번호: 5YH3NH  
Operated by KE  
KOREAN AIR

■ 대한항공은 인천공항 제2여객터미널 에서 운항합니다.

2

로스엔젤레스  
LAX

Los angeles intl  
Terminal No: 3  
17JUN22(금)12:20 (Local Time)



뉴 올리언스  
MSY

Louis armstrong intl  
Terminal No: -  
17JUN22(금)18:12 (Local Time)

**KE7437**  
예약번호: 5YH3NH  
Operated by DL  
DELTA AIR LINES

■ DELTA AIR LINES 항공기로 운항하는 공동운항편입니다.  
DELTA AIR LINES 의 탑승수속 카운터를 이용하시기 바랍니다.  
■ 모바일/웹 체크인은 운항항공사에서 운영합니다.

■ 공동운항(CODE SHARE)편 이용 시 실제 운항 항공사에 따라서 마일 적립 가능 여부가 달라질 수 있습니다.

3

뉴 올리언스  
MSY

Louis armstrong intl  
Terminal No: -  
25JUN22(토)10:15 (Local Time)



아틀란타  
ATL

Hartsfield-jackson intl  
Terminal No: S  
25JUN22(토)12:43 (Local Time)

**KE6974**  
예약번호: 5YH3NH  
Operated by DL  
DELTA AIR LINES

■ DELTA AIR LINES 항공기로 운항하는 공동운항편입니다.  
DELTA AIR LINES 의 탑승수속 카운터를 이용하시기 바랍니다.  
■ 모바일/웹 체크인은 운항항공사에서 운영합니다.

■ 공동운항(CODE SHARE)편 이용 시 실제 운항 항공사에 따라서 마일 적립 가능 여부가 달라질 수 있습니다.

4

아틀란타  
ATL

Hartsfield-jackson intl  
Terminal No: I  
25JUN22(토)15:00 (Local Time)



서울  
ICN

Incheon international  
Terminal No: 2  
26JUN22(일)19:10 (Local Time)

**KE8042**  
예약번호: 5YH3NH  
Operated by KE  
KOREAN AIR



임룡혁 연구원

■ 승객성명 Passenger Name  
**IM/RYUNGHYEOK MR**

1

서울  
**ICN**

Incheon international  
Terminal No: 2

17JUN22(금)14:30 (Local Time)



로스엔젤레스  
**LAX**

Los angeles intl  
Terminal No: B

Tom Bradley Int'l

17JUN22(금)09:40 (Local Time)

**KE0017**

예약번호: 5YH3NH  
Operated by KE  
KOREAN AIR

- 대한항공은 인천공항 제2여객터미널 에서 운항합니다.

2

로스엔젤레스  
**LAX**

Los angeles intl  
Terminal No: 3

17JUN22(금)12:20 (Local Time)



뉴 올리언스  
**MSY**

Louis armstrong intl  
Terminal No: -

17JUN22(금)18:12 (Local Time)

**KE7437**

예약번호: 5YH3NH  
Operated by DL  
DELTA AIR LINES

- DELTA AIR LINES 항공기로 운항하는 공동운항편입니다.  
DELTA AIR LINES 의 탑승수속 카운터를 이용하실 수 있습니다.
- 모바일/웹 체크인은 운항항공사에서 운영합니다.

- 공동운항(CODE SHARE)편 이용 시 실제 운항 항공사에 따라서 마일 적립 가능 여부가 달라질 수 있습니다.

3

뉴 올리언스  
**MSY**

Louis armstrong intl  
Terminal No: -

25JUN22(토)10:15 (Local Time)



아틀란타  
**ATL**

Hartsfield-jackson intl  
Terminal No: S

25JUN22(토)12:43 (Local Time)

**KE6974**

예약번호: 5YH3NH  
Operated by DL  
DELTA AIR LINES

- DELTA AIR LINES 항공기로 운항하는 공동운항편입니다.  
DELTA AIR LINES 의 탑승수속 카운터를 이용하실 수 있습니다.
- 모바일/웹 체크인은 운항항공사에서 운영합니다.

- 공동운항(CODE SHARE)편 이용 시 실제 운항 항공사에 따라서 마일 적립 가능 여부가 달라질 수 있습니다.

4

아틀란타  
**ATL**

Hartsfield-jackson intl  
Terminal No: I

25JUN22(토)15:00 (Local Time)



서울  
**ICN**

Incheon international  
Terminal No: 2

26JUN22(일)19:10 (Local Time)

**KE8042**

예약번호: 5YH3NH  
Operated by KE  
KOREAN AIR