

Data Science for Social Impact

제1회 SDGs 데이터 혁신 포럼 (통계청 통계개발원)

날짜

2022.05.25

발표자

차 미 영 교수

소속

KAIST 전산학부 & IBS 기초과학연구원



서울시 버스 노선의 개선



미래의 식량문제



- 유엔기구 글로벌 펄스 (UN Global Pulse Lab) 협업을 통해 섬나라 인도네시아의 주요 생필품 소고기, 닭고기, 양파, 고추 등 가격을 실시간 예측하는 알고리즘 개발
- 기존 통계대비 10일 빠른 나우캐스팅 알고리즘과 프로토타입 기술을 무상 제공

관세행정의 디지털 전환



Customs Description	Quantity	Declared Value	
Cotton T-shirts	2	\$ 50.00	x
Jeans Pants	1	\$ 100.00	x
Customs Description	1	\$ Value	x
+ Add line			



면세범위 초과물품, 위장 반입, 원산지 조작 등 세관에서 벌어지는 불법 행위를 적발하는 AI기술 개발



Agenda

01. Data Science Vision

02. Poverty Mapping

03. Climate Action

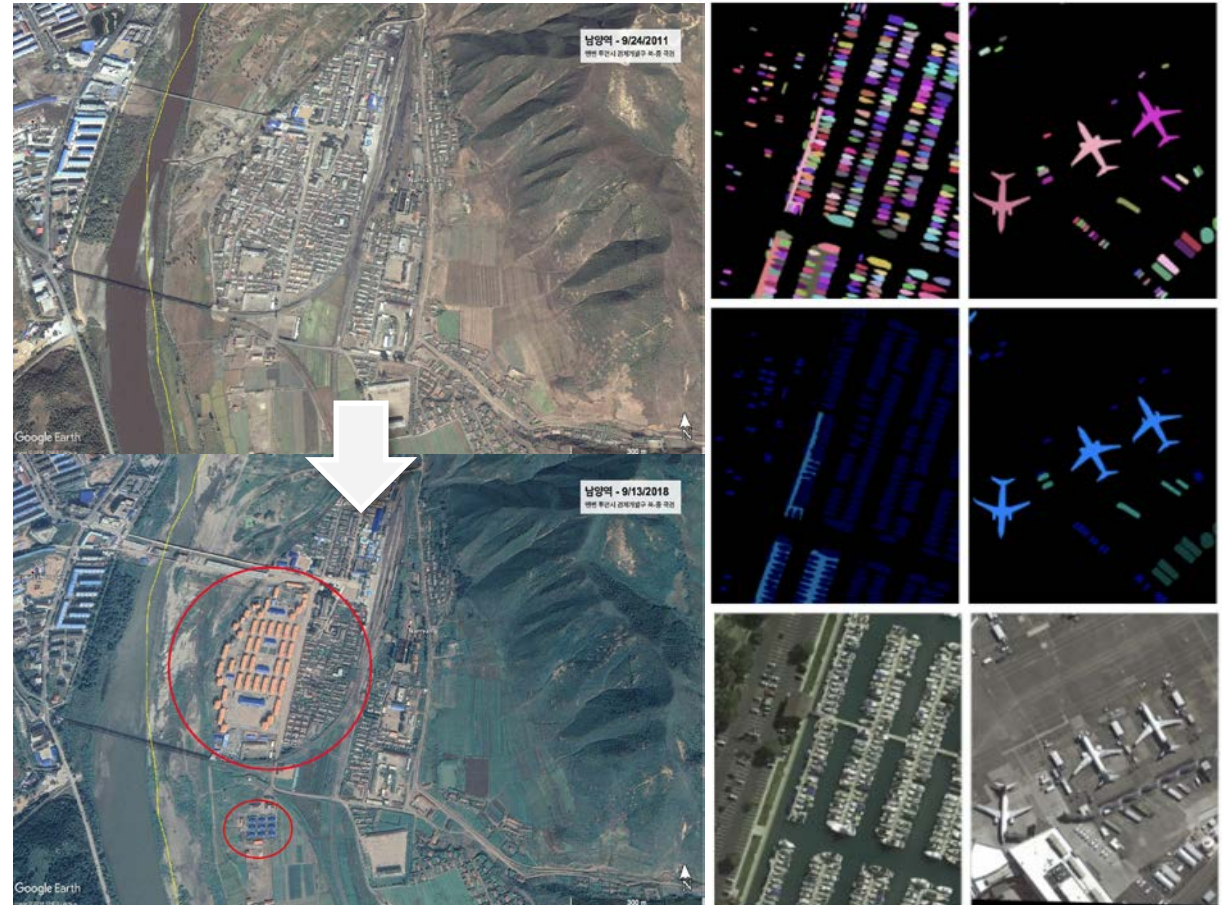


코로나19로 인한 시장소득 급감의 영향을
정부 재정 지원으로 상쇄하고 있으나 노인
빈곤율은 여전히 높은 수준

위성영상을 활용한 인문지리학 및 원격탐사의 혁신

야간 영상에서 주간영상으로 발전

- 주간 위성영상은 고도화된 딥러닝 기술과 접목 하여 인간 거주 패턴과 지역단위의 경제 규모까지 추적 가능



합성곱신경망(CNN)과 전이학습(transfer learning)이 위성영상에 활용되어 빈곤지역 예측에 활용

RESEARCH

Science
AAAS

RESEARCH ARTICLES

ECONOMICS

Combining satellite imagery and machine learning to predict poverty

Neal Jean,^{1,2*} Marshall Burke,^{3,4,5*}† Michael Xie,¹ W. Matthew Davis,⁴
David B. Lobell,^{3,4} Stefano Ermon¹

Reliable data on economic livelihoods remain scarce in the developing world, hampering efforts to study these outcomes and to design policies that improve them. Here we demonstrate an accurate, inexpensive, and scalable method for estimating consumption expenditure and asset wealth from high-resolution satellite imagery. Using survey and satellite data from five African countries—Nigeria, Tanzania, Uganda, Malawi, and Rwanda—we show how a convolutional neural network can be trained to identify image features that can explain up to 75% of the variation in local-level economic outcomes. Our method, which requires only publicly available data, could transform efforts to track and target poverty in developing countries. It also demonstrates how powerful machine learning techniques can be applied in a setting with limited training data, suggesting broad potential application across many scientific domains.

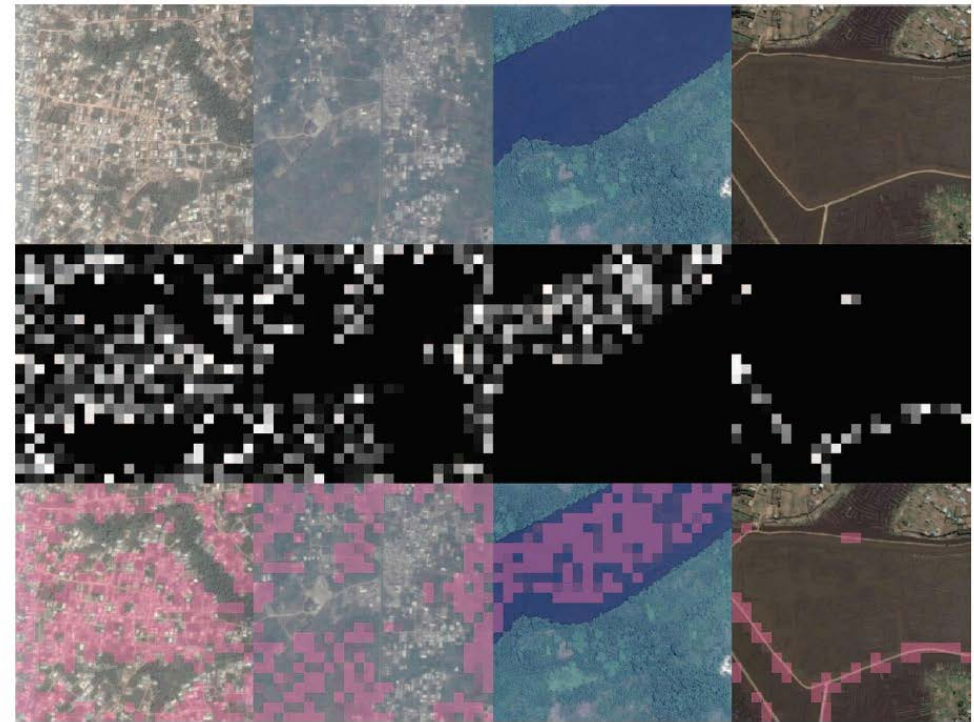


Fig. 2. Visualization of features. By column: Four different convolutional filters (which identify, from left to right, features corresponding to urban areas, nonurban areas, water, and roads) in the convolutional neural network model used for extracting features. Each filter “highlights” the parts of the image that activate it, shown in pink. By row: Original daytime satellite images from Google Static Maps, filter activation maps, and overlay of activation maps onto original images

인공지능 학습 문제로 전환

머신이 이미지를 분석하여 빈곤정도를 유추할 수 있을까?

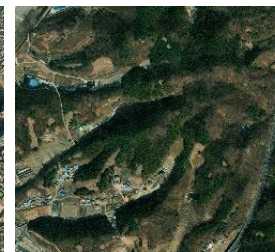


Population: 30,000
Income: \$1.8B
Inequality (Gini): 0.3

- 필요한 정보:
 - 위성영상
 - 기초가 되는 경제지표
 - 학습을 위한 이미지별 경제 라벨



Urban



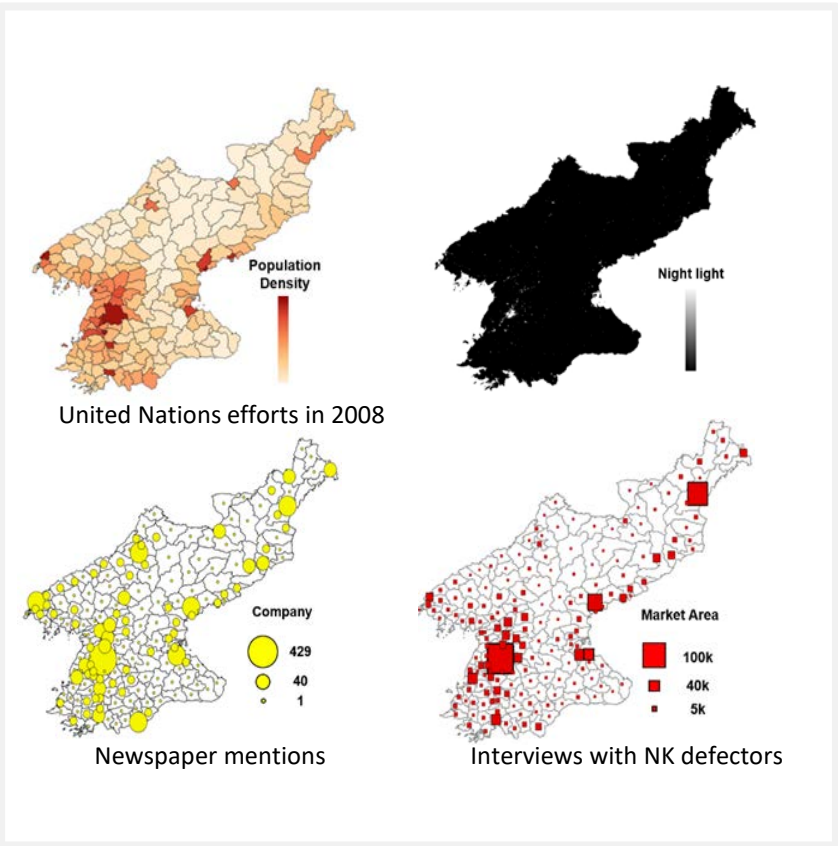
Rural



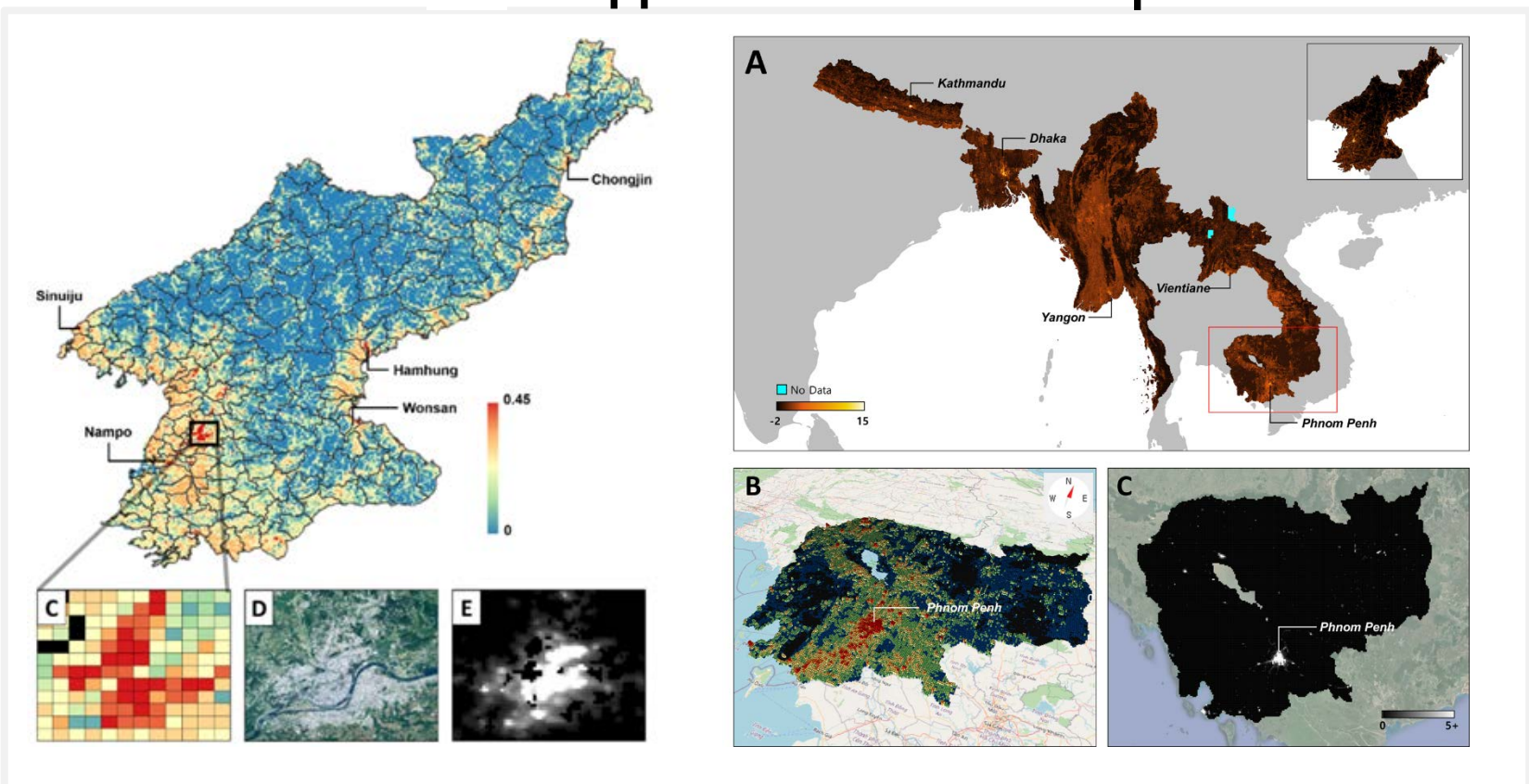
Uninhabited

1km² 격자 단위의 세밀한 빈곤지수의 계산

Existing statistics



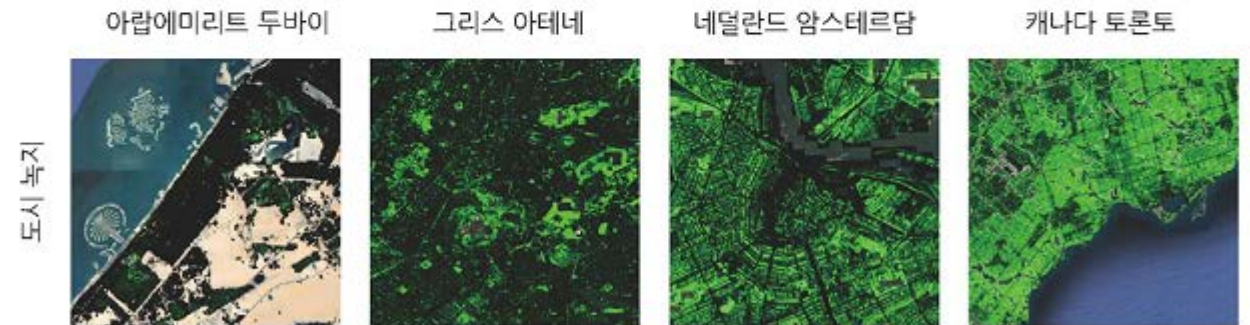
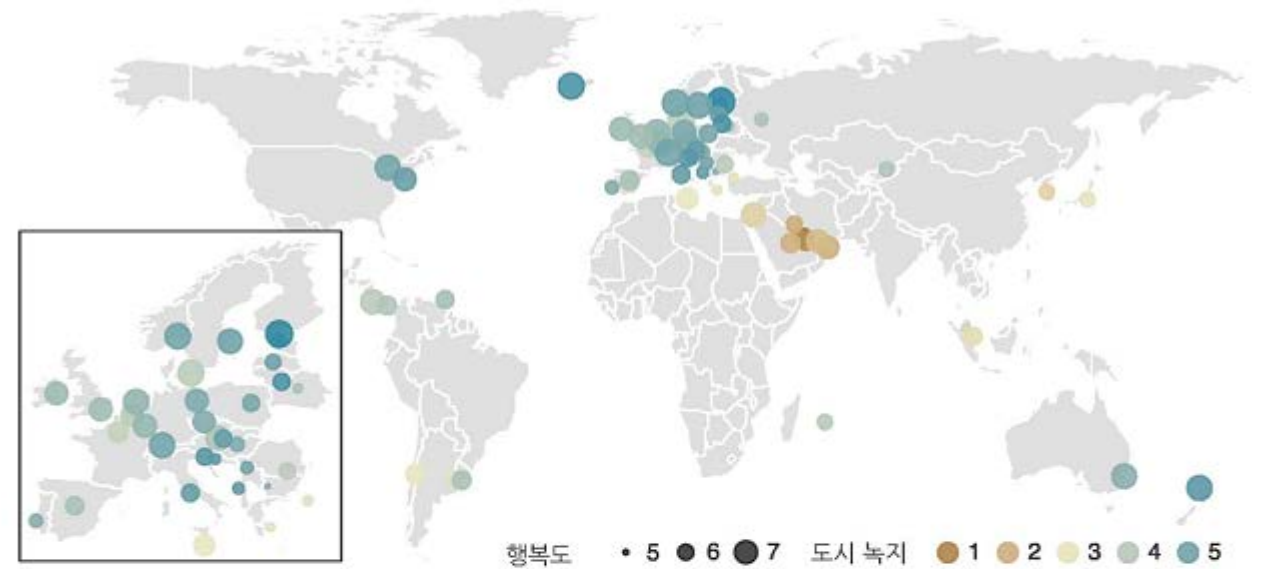
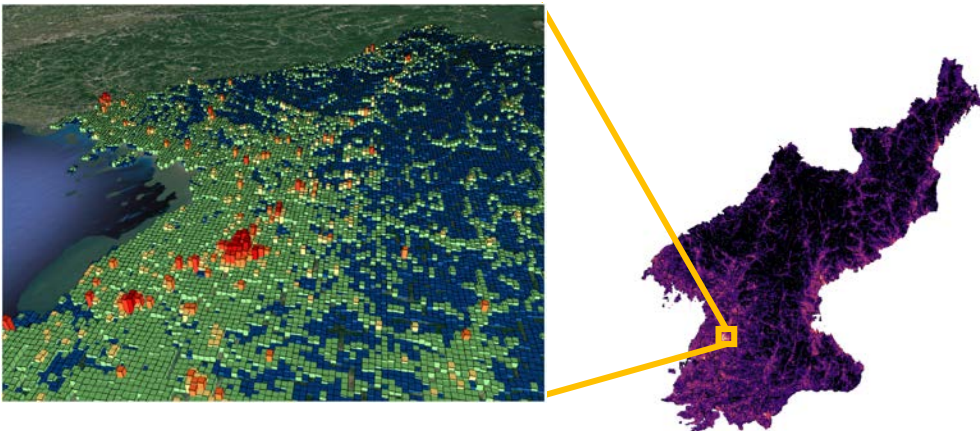
Human-machine model applied on least developed countries



To create the Partial Order Graph (POG) of North Korean images for 2016 ~ 2019, we hired 10 experts who examined 21 satellite clusters given by algorithms and determined their rank order. Expert groups included North Korea defectors, economists, and GIS researchers.

위성영상의 딥러닝 분석으로 빈곤지수와 녹지지수 계산 (SDG 관련 분석 및 지표 개발)

- 컴퓨터비전 합성곱신경망(CNN)과 전이학습(transfer learning) 기반 위성영상 분석 기술을 개발
- 북한을 비롯한 아시아 지역 빈곤 예측과 도심 내 녹지 측정으로 행복지수 연구

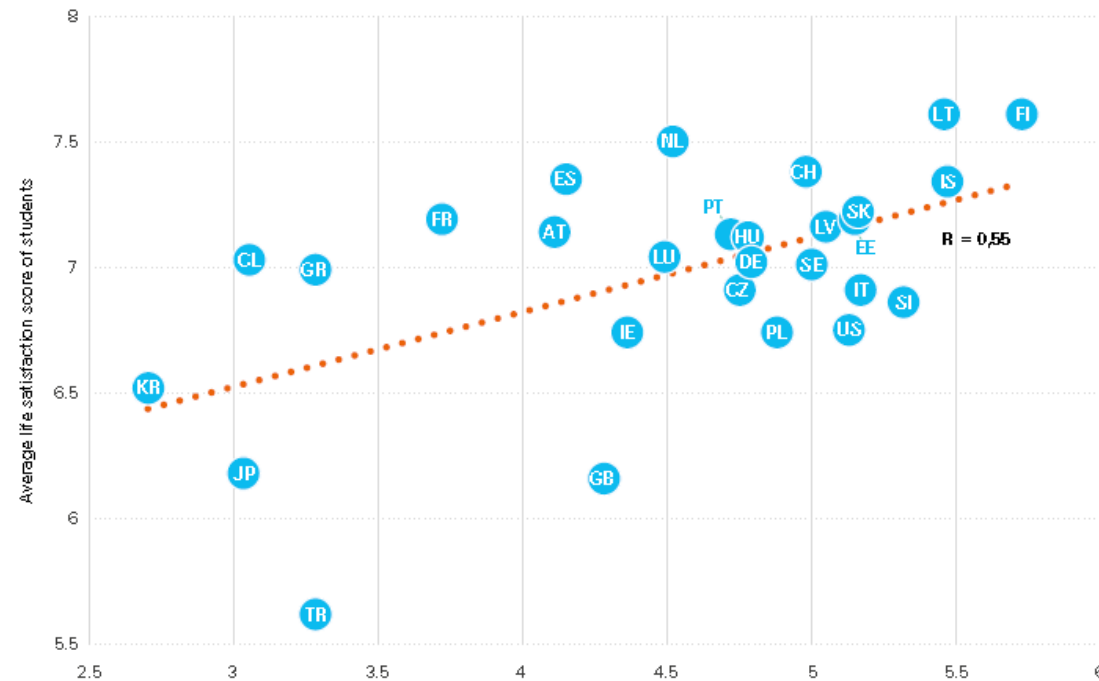


국가 별 행복도와 도심 녹지 비율: 전 세계 60개국에 대해 도시의 녹지 비율(원 색)과 행복도 조사 결과(원 크기)를 비교하면 둘 사이의 상관관계가 포착된다. 왼쪽 아래는 박스는 유럽 국가에 대해 상세한 결과를 보여준다. 그간 녹지 실태 조사가 현장 방문 혹은 항공사진을 기반으로 했다면 이번 연구에서는 여름철 위성영상 자료를 활용하여 더 광범위한 지역에 대한 분석이 가능했다.

연구팀 결과가 UNICEF 2022 보고서에서 주요 결과로 사용 및 Slovakia 정부 브리핑

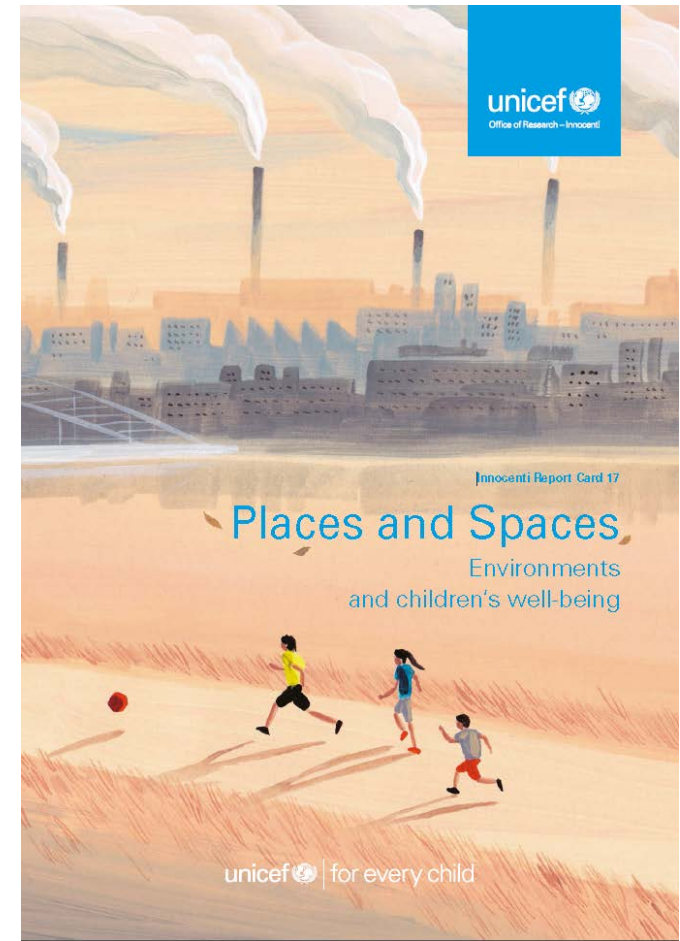
- 연구팀의 60개국 위성영상 녹지 계산 데이터의 비교에 근거해 Finland, Iceland, Lithuania는 한국이나 이스라엘보다 두 배 이상 인구당 녹지를 보유함을 확인하며 연구팀의 지표를 적극 활용
- 15세 인구 및 성인 인구의 행복도와 녹지도 높은 상관관계 확인하며 녹지 확보의 정책적 함의 제공

Figure 19: Children in greener countries are more satisfied with life
Life satisfaction of 15-year-olds (2018) and urban green spaces (2018)



Source: PISA 2018 and Kwon et al. (2021).

Urban Green Spaces Index



Agenda

01. Data Science Vision

02. Poverty Mapping

03. Climate Action



통합수자원관리 수준은 2017년에 비해 2020년에 향상되었으나 OECD 국가 중 중하위권이며 물 스트레스는 높은 수준



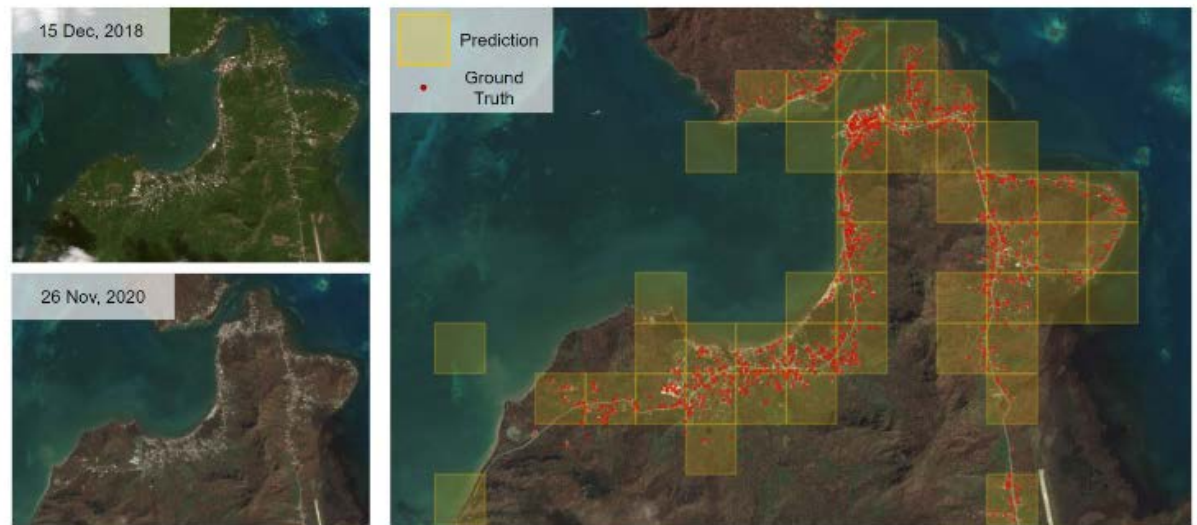
온실가스 배출량 증가율은 둔화하는 추세
속 2019년 배출량은 전년 대비 감소



해양쓰레기 총량 증가하고 있으며,
해수수질도 2018년부터 상대적으로 나빠짐

Sustainable development: 기후변화로 인한 자연재해를 모니터링

- 기후변화로 인해 재해의 발생빈도와 그 피해 규모가 급격히 증가
- 재난 발생 이전-이후 위성영상을 분석
격자 단위에서 홍수 피해 여부 추정
(Z=17, 1.2m/pixel ArcGIS World Imagery)
- 딥러닝 모델로 정밀한 피해 지역의
범위 예측(F1=0.86 정확도) 및 효율적
구호 자원 배분에 기여 가능



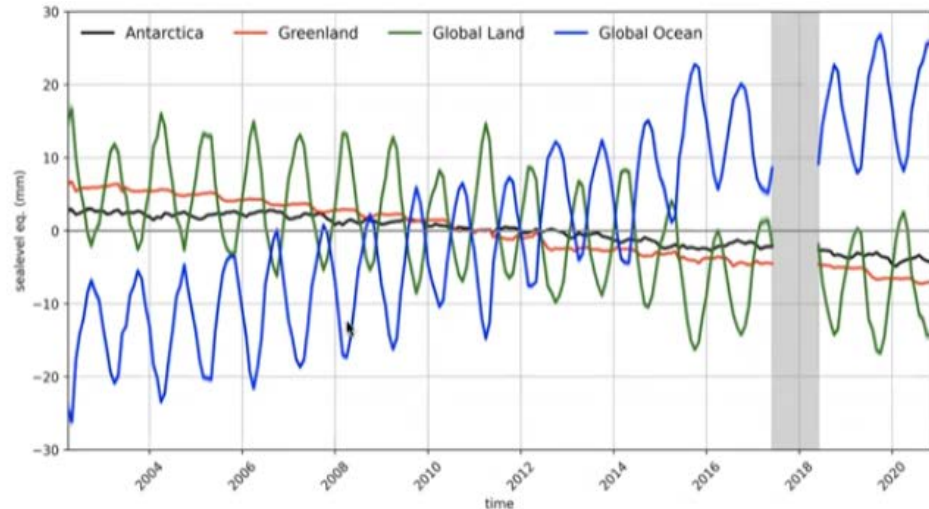
(A) Input satellite imagery

(B) Prediction Result



Global water shortage

- Freshwater (담수)의 감소는 세계적 문제로 캘리포니아주는 물 사용에 대한 법안이 통과
- 빅데이터와 머신러닝에 기반한 위성영상 모니터링을 통해 정확한 담수량 측정과 미래예측



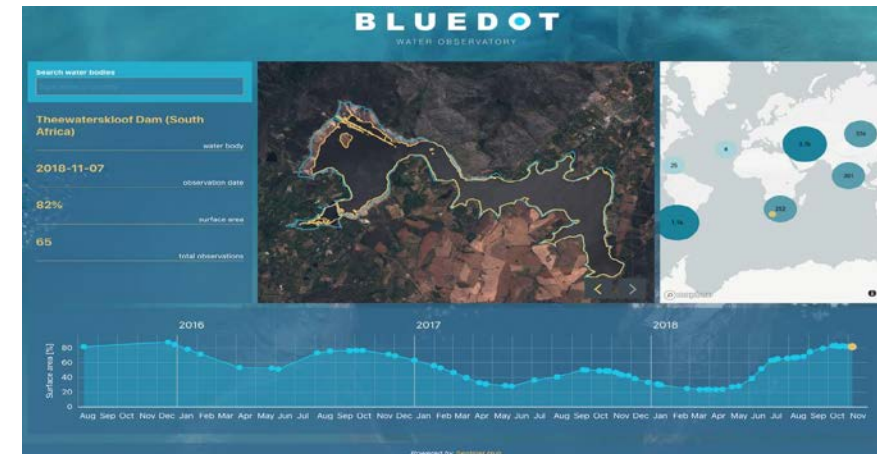
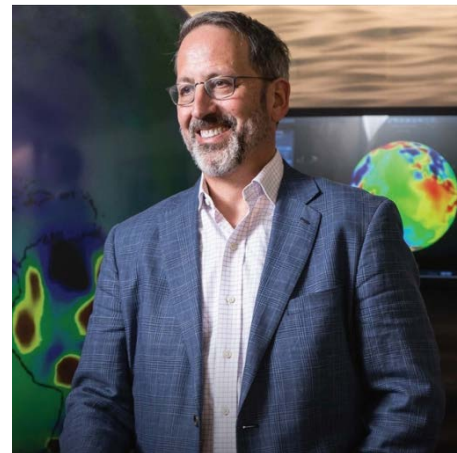
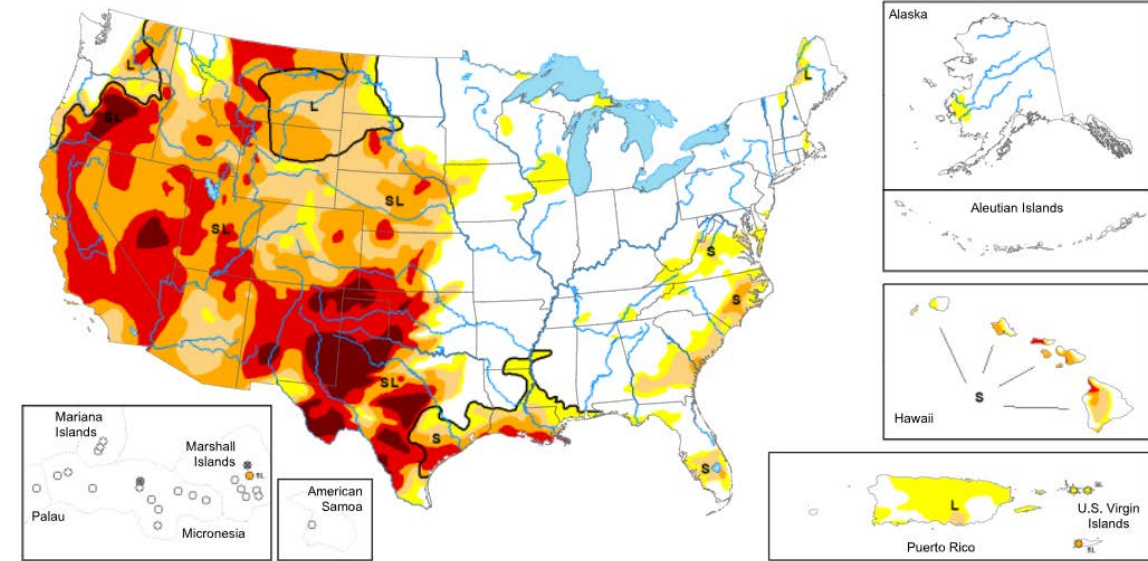
Global Institute for Water Security, Canada (Dr. Jay Famiglietti) <https://water.usask.ca/>

National Drought Mitigation Center, USA <https://droughtmonitor.unl.edu/>

U.S. Drought Monitor

Map released: May 12, 2022

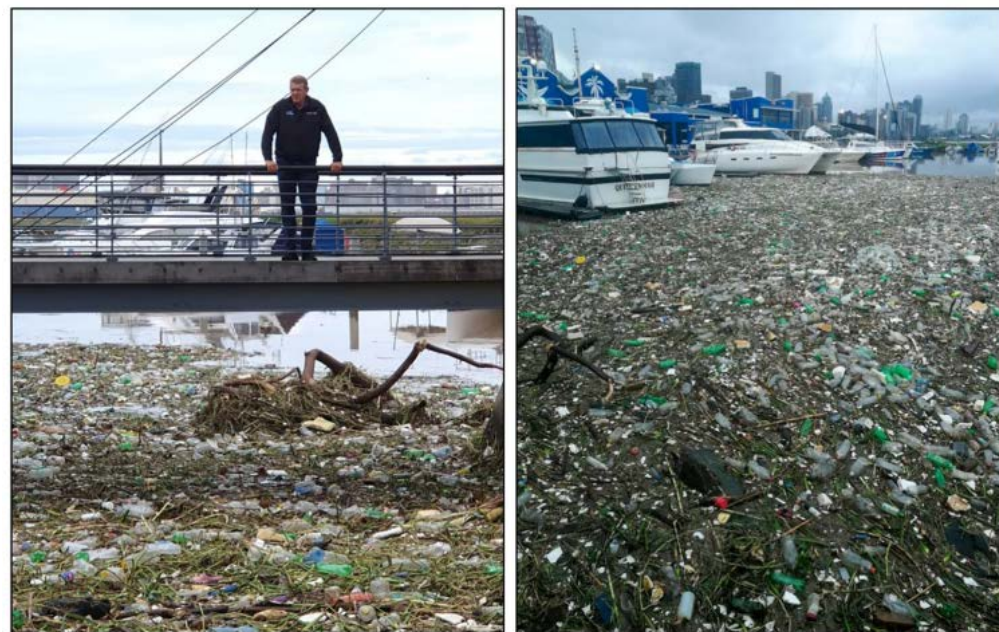
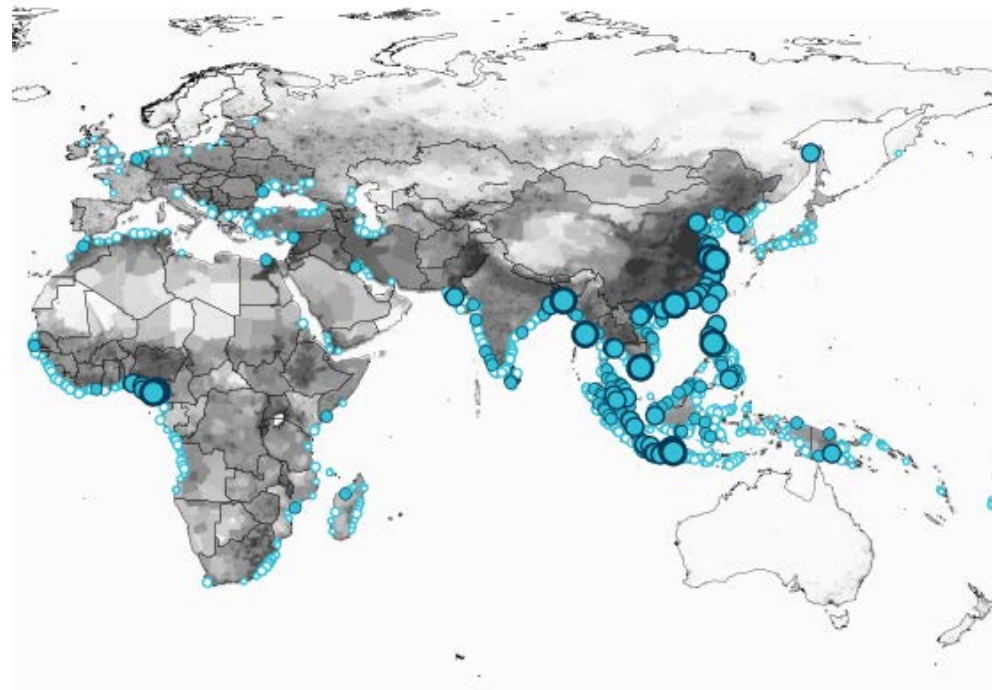
Data valid: May 10, 2022



BlueDot <https://water.blue-dot-observatory.com/>

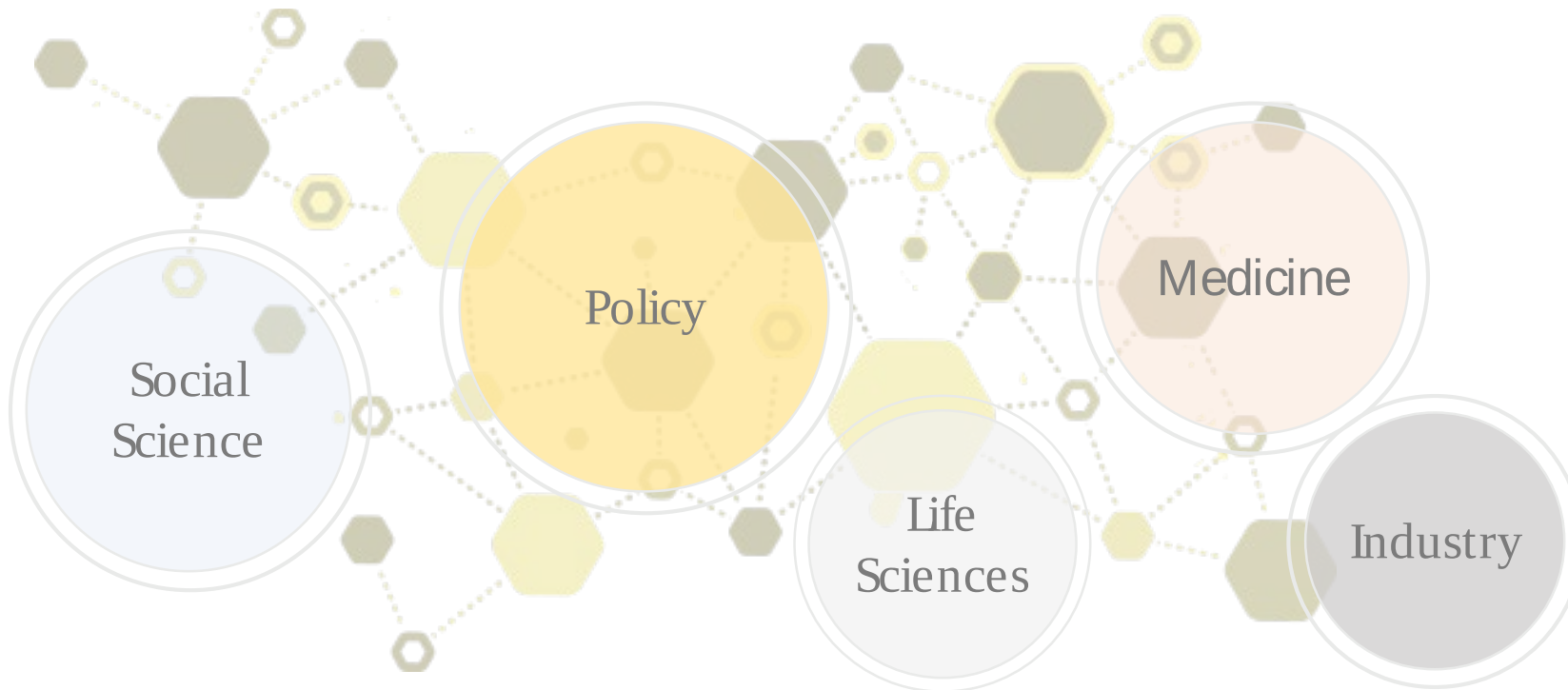
Marine plastics: 플라스틱 해양쓰레기

- 플라스틱 해양쓰레기는 주요 환경문제로 우려가 커지고 있음. 연간 115~241만 톤의 플라스틱이 강을 통해 바다로 유입되고 있는 것으로 파악
- 특히 아시아 지역의 급격한 도시화로 인해, 태평양 지역의 플라스틱 문제가 큰 위협이 될 것으로 예상
- 위성영상과 인공지능 탐지기술로 해양 플라스틱을 추적하는 기술이 최근 주목을 받고 있음



새로운 기회

이종(heterogeneous) 빅데이터의 결합과 새로운 인공지능 기반 계산과학 방법의 적용으로
SDG 목표 달성과 계획 수립에 데이터 사이언스 기반 도약이 세계적으로 시작되고 있음



IBS Team & Collaborators



Meeyoung Cha

Chief Investigator, IBS Data Science Group
Associate Professor
Kaist, School of Computing



Sangyoon Park, Dr.

University of Hong Kong,
Economics



Jihee Kim, Dr.

Kaist, College of Business



Donghee Yvette Wohn, Dr.

Njit, Informatics



Sooyoung Bae, Dr.

University of Massachusetts
Amherst, Communication



Cheng-Te Li, Dr.

NCKU, Statistics



**Krishna Gummadi,
Dr.**

MPI-SWS



Yu-I Ha, Dr.

Kaist Post-doc Fellow



Kunwoo Park, Dr.

Soongsil Univ, School of AI Convergence



Jiyoung Han, Dr.

Seoul National University Post-
doc Fellow



Jeesurk Yang

Seoul National University
M.S student, Geography



Sungkyu Park, Dr.

IBS Research Fellow



Hyojin Chin, Dr.

IBS Research Fellow



Bayu Adhi Tama, Dr.

IBS Research Fellow



Kyungshin Park, Dr.

Korea University LAW



Kyomin Jung, Dr.

Seoul National University



Hyunjoo Yang, Dr.

Incheon National University
School of Northeast Asian Studies



Karandeep Singh, Dr.

IBS Research Fellow



Sejong Kwon, Dr.

Samsung Research



**Sundong Kim,
Dr.**

IBS Research Fellow





IBS DATA SCIENCE GROUP

기초과학연구원 데이터사이언스그룹