NVIDIA Artificial Intelligence Lab &
Competence Center for Deep Learning
German Research Center for
Artificial Intelligence (DFKI)

Opportunities and Future Directions in Land Use and Land Cover Classification with Sentinel-2

Patrick Helber, Benjamin Bischke, Damian Borth, Andreas Dengel

DFKI-German Research Center for Artificial Intelligence

- Largest AI research center in the world
- About 900 employees
  - more than 510 staff members and
  - more than 400 part-time student researchers
- about 210 ongoing projects


Saarbrücken


Kaiserslautern


Bremen


Berlin


Osnabrück

DFKI is a Joint Venture of...

SAP
HARTING
John Deere
intel
Google
VSE Vorweg gehen
Nuance
Deutsche Telekom
software AG
EMPOLIS Information Management
Microsoft Deutschland GmbH
RICOH
Deutsche Messe AG Hannover · Germany
KIBG
Fraunhofer

BMW Group
VOLKSWAGEN
BOSCH
AIRBUS GROUP
CLAAS
Universität des Saarlandes
Universität Bremen
Technische Universität Kaiserslautern

> 80 Startup and Spin-Off Companies

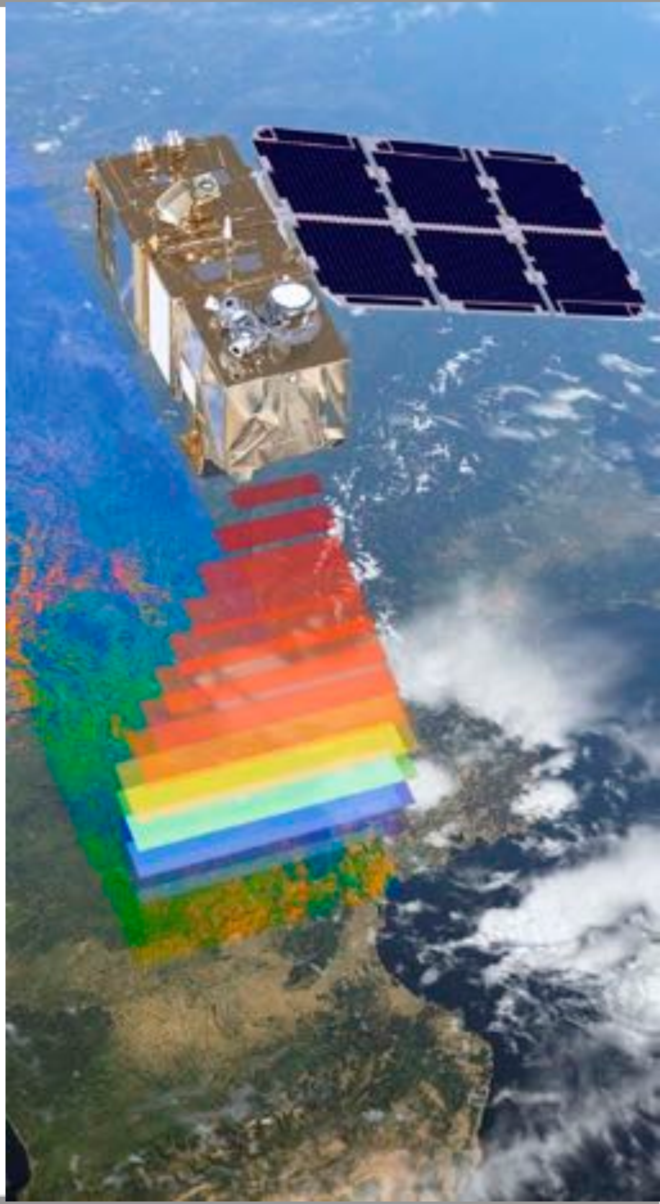
acrolinx
Ascella
b4
brainbot
Camelot IDPro
CELI
Chusable
climbtrack
CONENO
dathemis FINANCIAL DATA ANALYTICS
DHC Dr. Herterich & Consultants
digipen technologies
DNC Information Management
EMPOLIS INFORMATION MANAGEMENT
ententee
eyeled
graphicsmedia.net
Ground Truth Robotics
ifx-design
IMAGO
imc Advanced Learning Solutions
insiders technologies
iOXP cogn live AR
iSol
KIANA SYSTEMS
kielements
KRAKEN
LEVERTON
moana
MINERVIS
mineway informed decisions
motama
NEOCOSMO
STAUFEN.DIGITAL NEONEX
PBFE
PIONEO pioneering technologies
Plan Software
pxio
QUERTEX GmbH social software solutions
schwartz&stahl
SemanticEdge
SEMATELL INTELLIGENT ANSWERS
semvox
sevenRE
SIEDA
sis software
SociovestixLabs
SQFIND
softgarden
SUPPORTINGyou!
TNM
Transfer Ventures
VISION IMPULSE

Sustainable Development Goals + EO Data

- Vast amount of Earth observation data
- Suitable to address **global challenges** and foster **innovative applications**

- Manual analysis practically impossible
  → Automatic analysis necessary

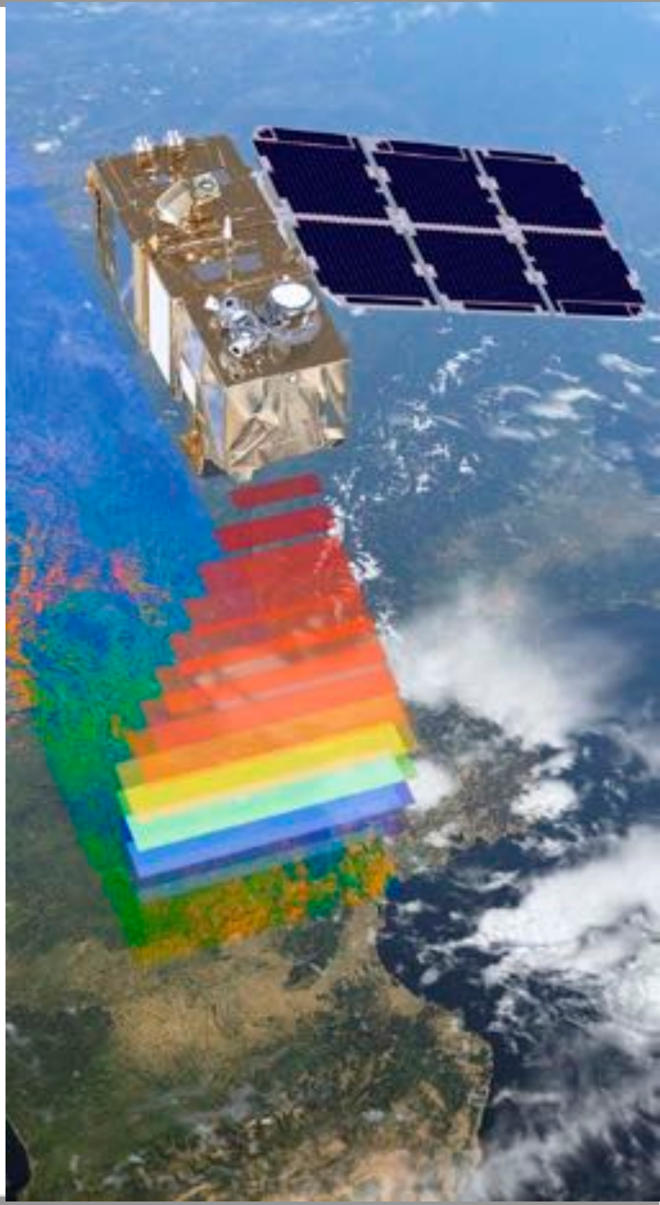
→ **AI can deal with large-scale data**



Sustainable Cities and Communities

Half of humanity – 3.5 billion people – lives in cities today and 5 billion people are projected to live in cities by 2030

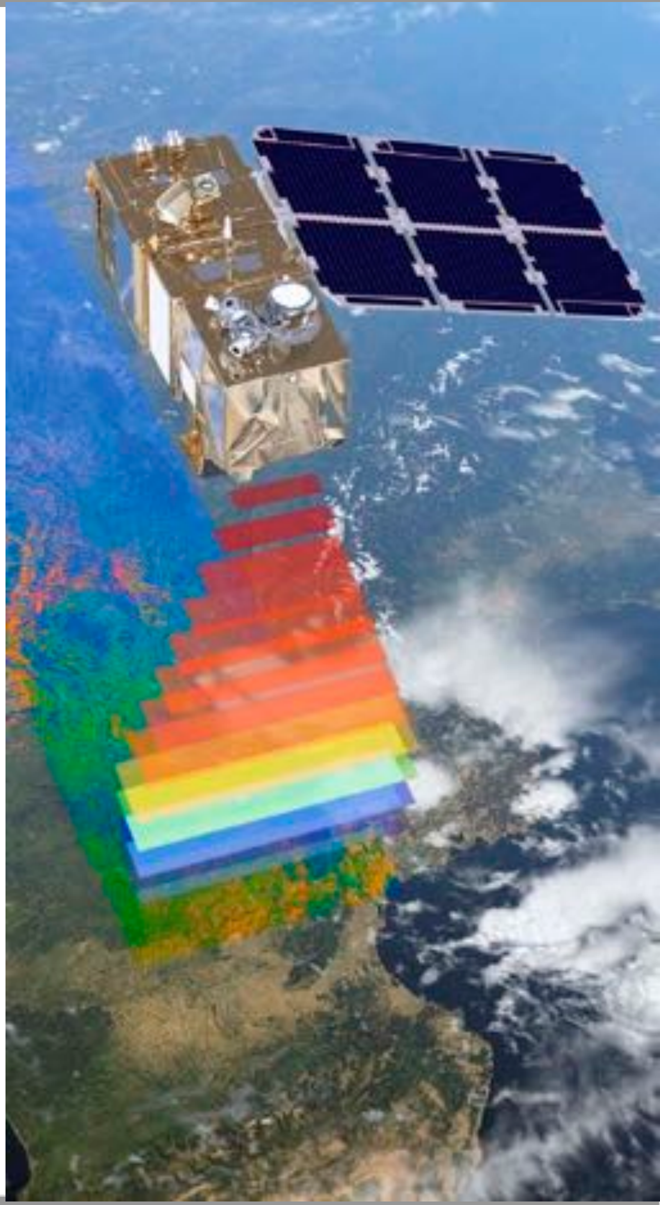
95 per cent of urban expansion in the next decades will take place in developing world



Sustainable Cities and Communities

883 million people live in slums today and most them are found in Eastern and South-Eastern Asia.

Rapid urbanization is exerting pressure on fresh water supplies, sewage, the living environment, and public health



Overview - Deep Learning in Earth Observation

Application Oriented Research

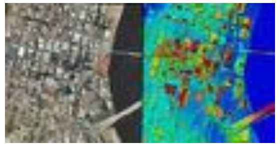
Land-Use and Land-Cover Classification (Patch based)

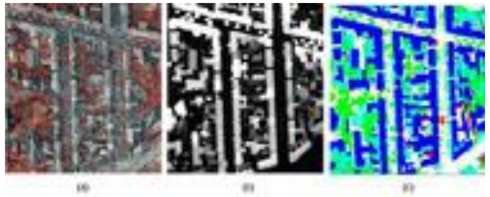
DeepEye for Natural Disasters

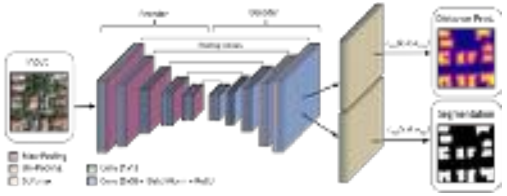
Multimedia Satellite Task 2017/2018

Estimation of (Micro)-Economic Factors

Land Use and Land-Cover Segmentation







Missing Data during Inference

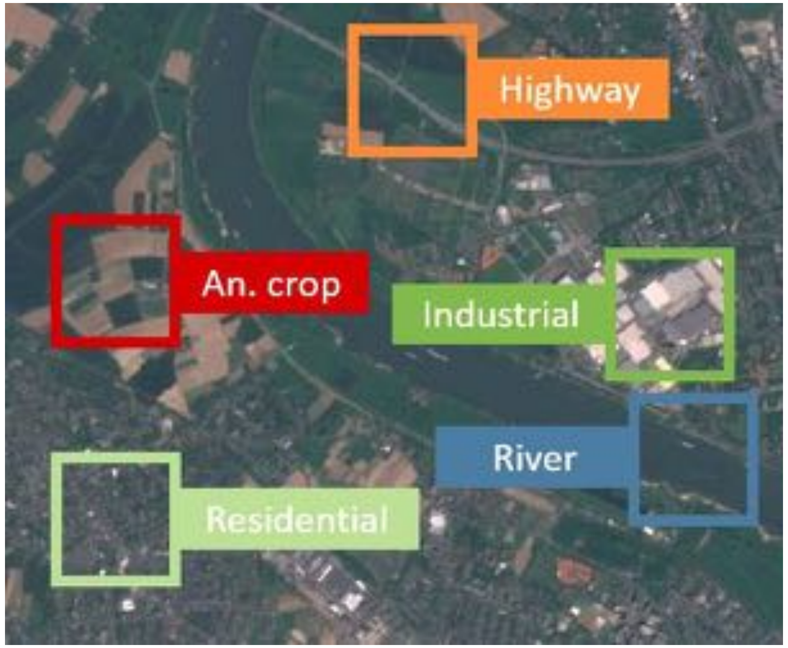
Multimodal Fusion of Satellite Data

Multi-Task Learning for Sem. Segmentation

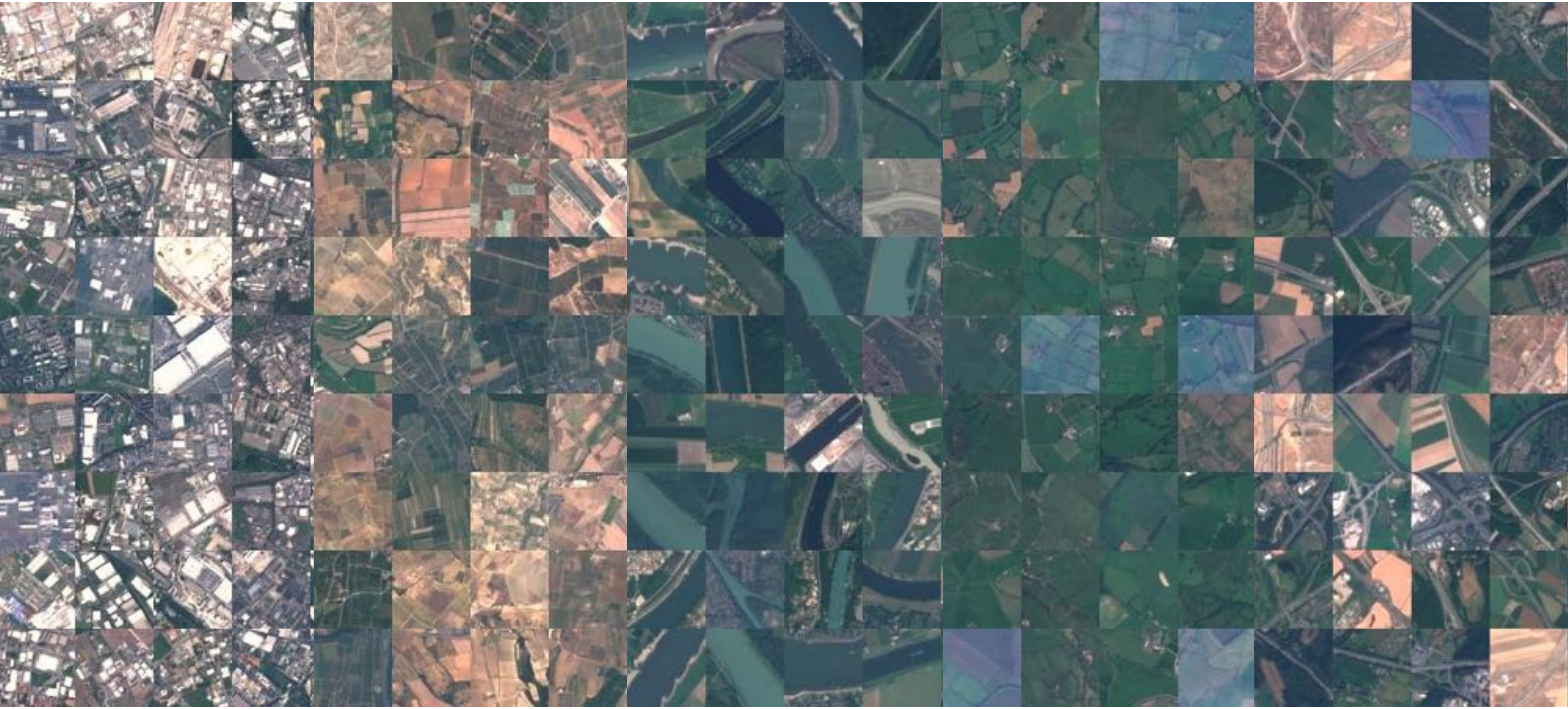
Fundamental Research

Sentinel-2 for Land Use and Land Cover Dynamics

- Two-satellite constellation
  - 5 days revisit time
- Spatial resolution of up to **10 meter per pixel**
- **13 spectral bands**
- Global land surface coverage
  - Onshore
  - Large islands
  - Inland and coastal waters

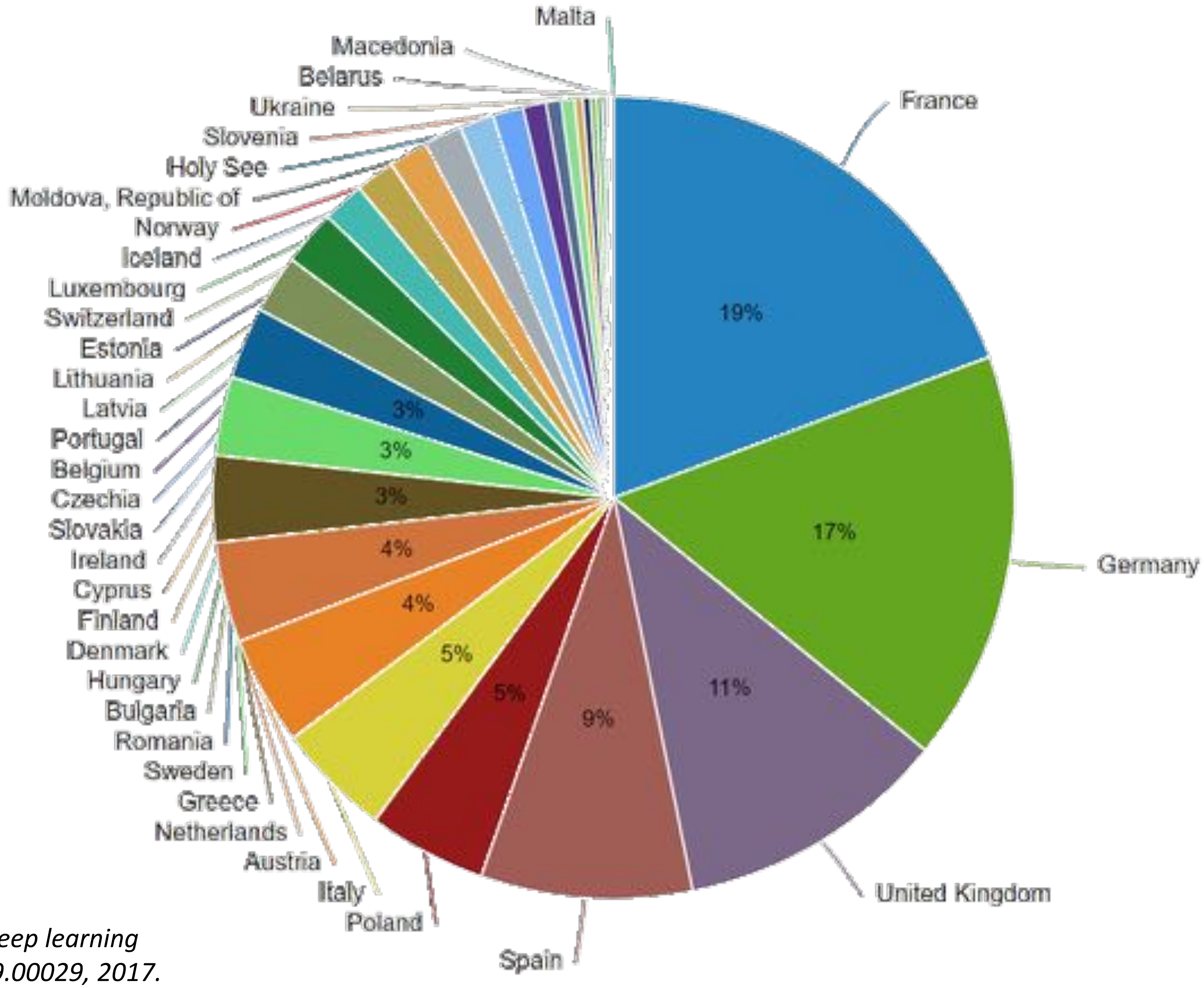


EuroSAT Publicly Released



EuroSAT - Distributed Over 30 Countries

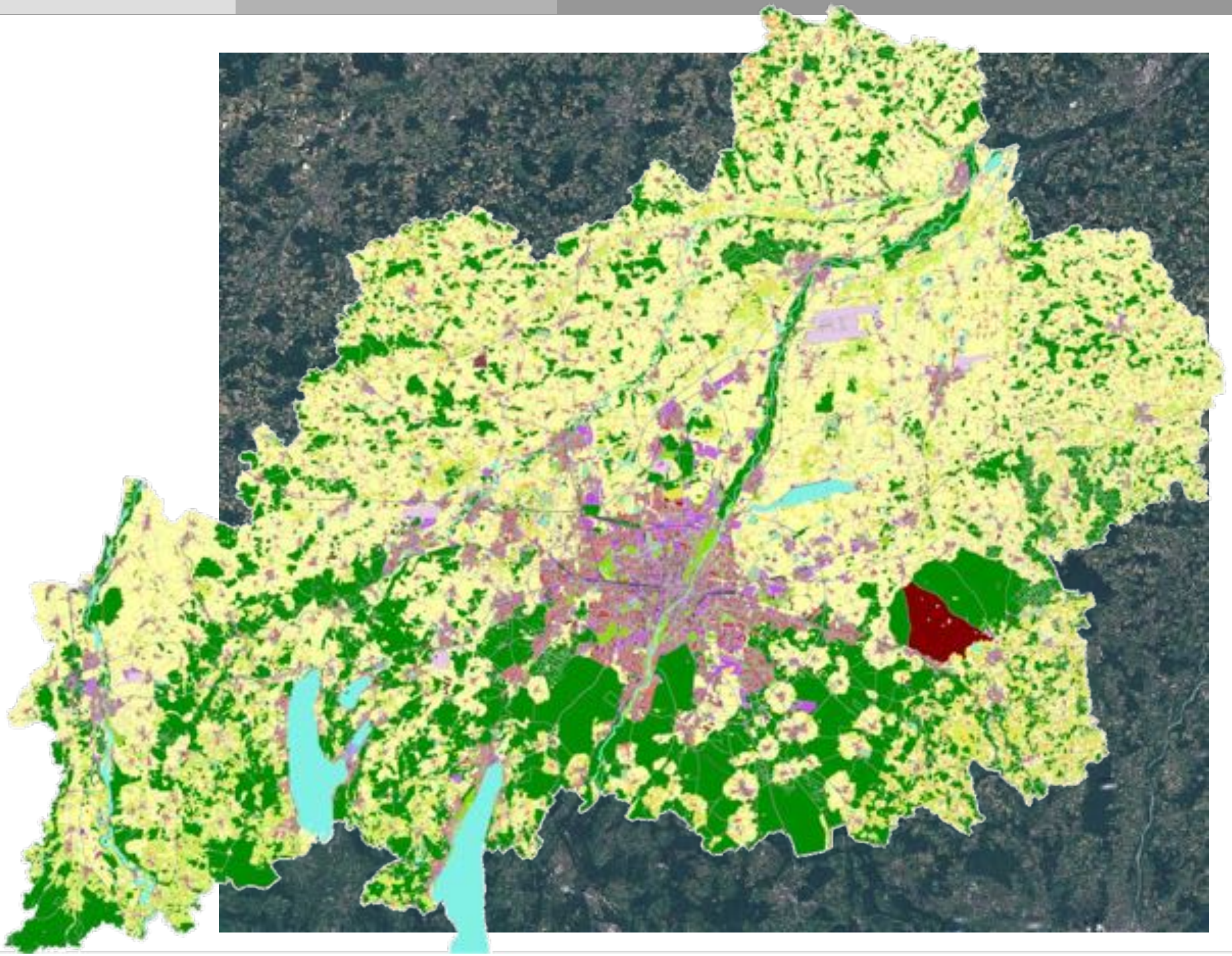
- **10 classes**
- **27,000 geo-referenced images**
- **64 x 64 images**
- 2,000 – 3,000 images per class
- **13 spectral bands**
  - Spatial resolution: 10 m per pixel



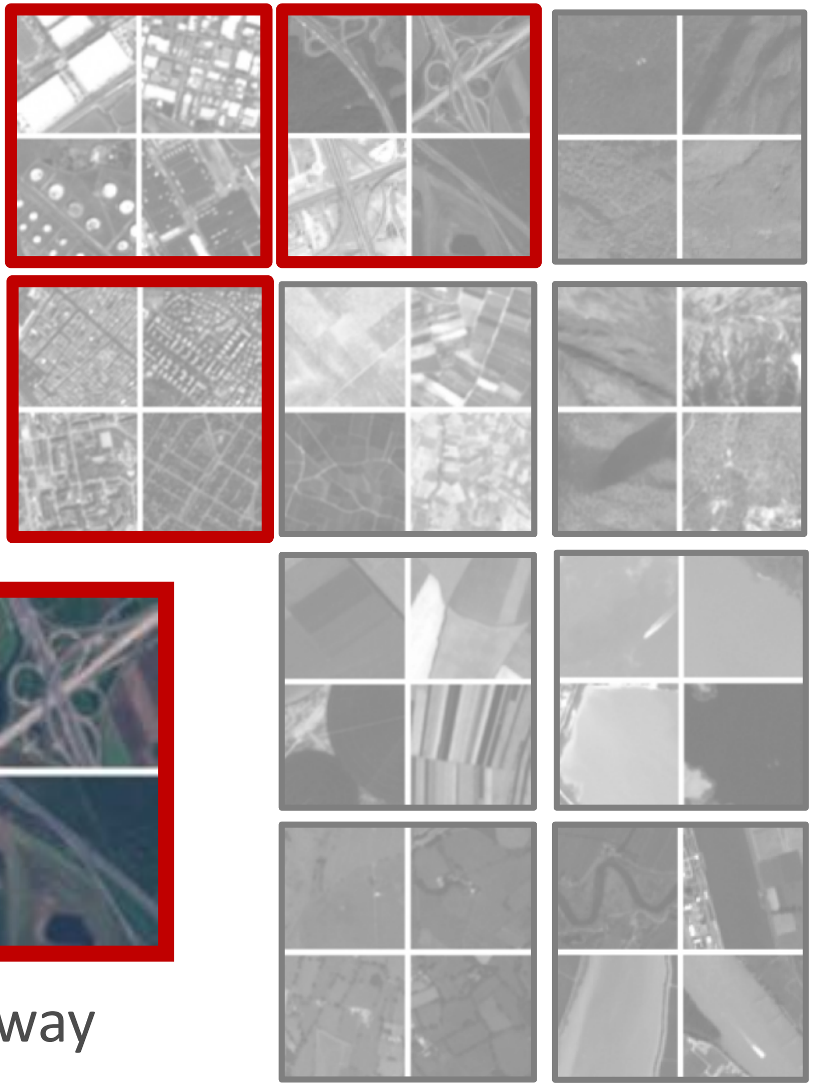
***P. Helber, B. Bischke, A. Dengel, and D. Borth**, "Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification," arXiv preprint arXiv:1709.00029, 2017.*

Segmentation Masks for 785 cities

- European Urban Atlas
  - Detailed mapping for **785 cities** distributed over **30 European countries**
  - Released August 2016
  - Covered time period: 2011-2013



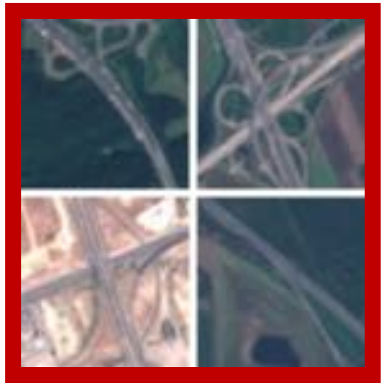
P. Helber, B. Bischke, A. Dengel, and D. Borth, *“Eurosat: A novel dataset and deep learning benchmark for land use and land cover classification,” arXiv preprint arXiv:1709.00029, 2017.*



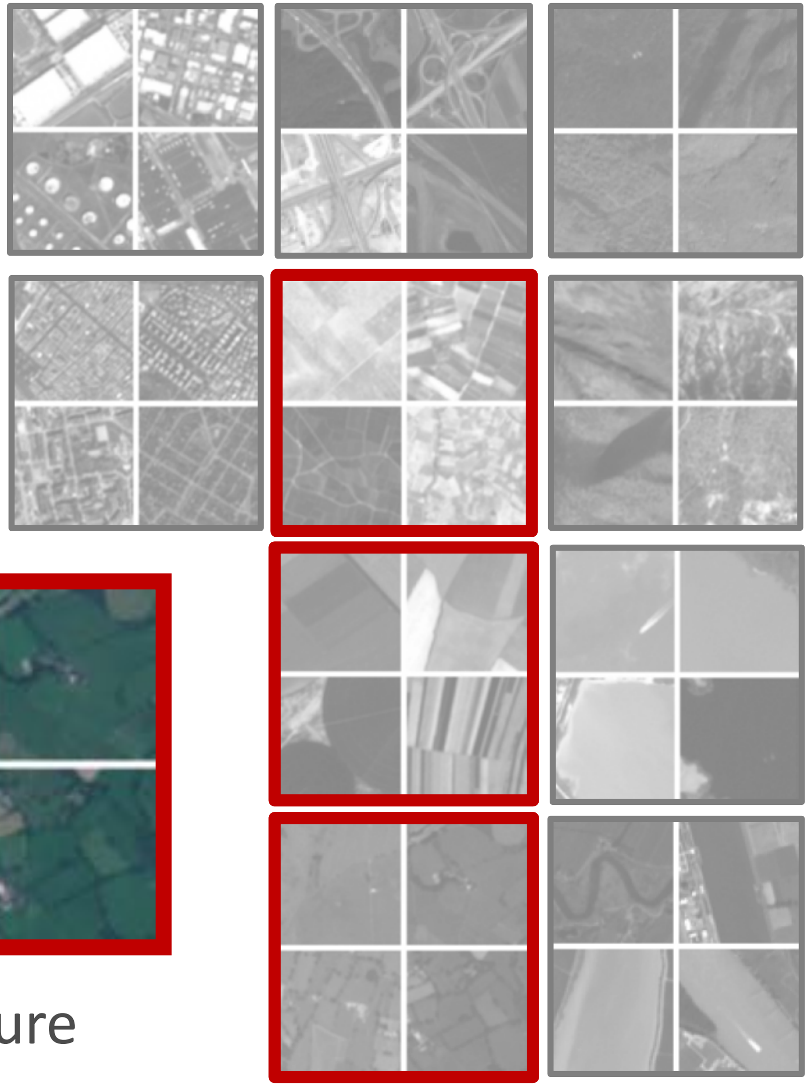
- **Built-up Areas**

Industrial

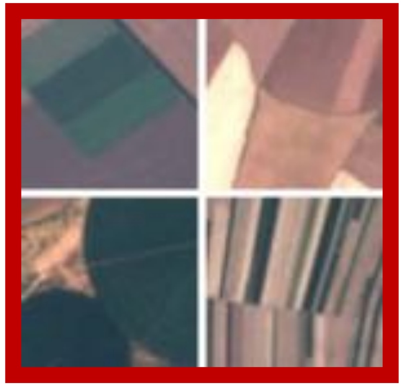
Residential



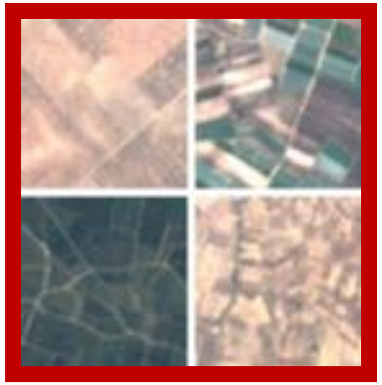
Highway



- **Agricultural Land**



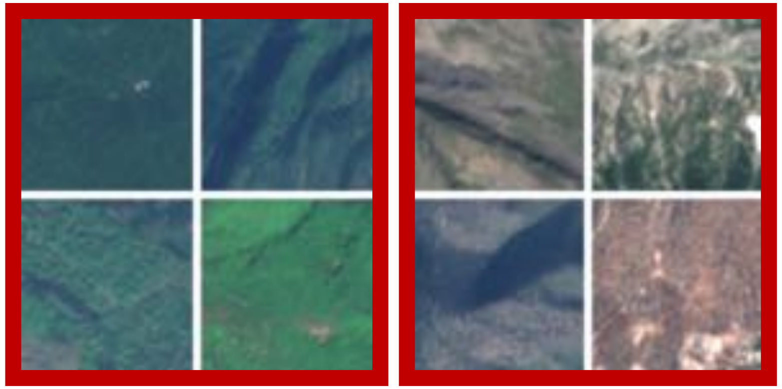
Annual Crop



Permanent Crop

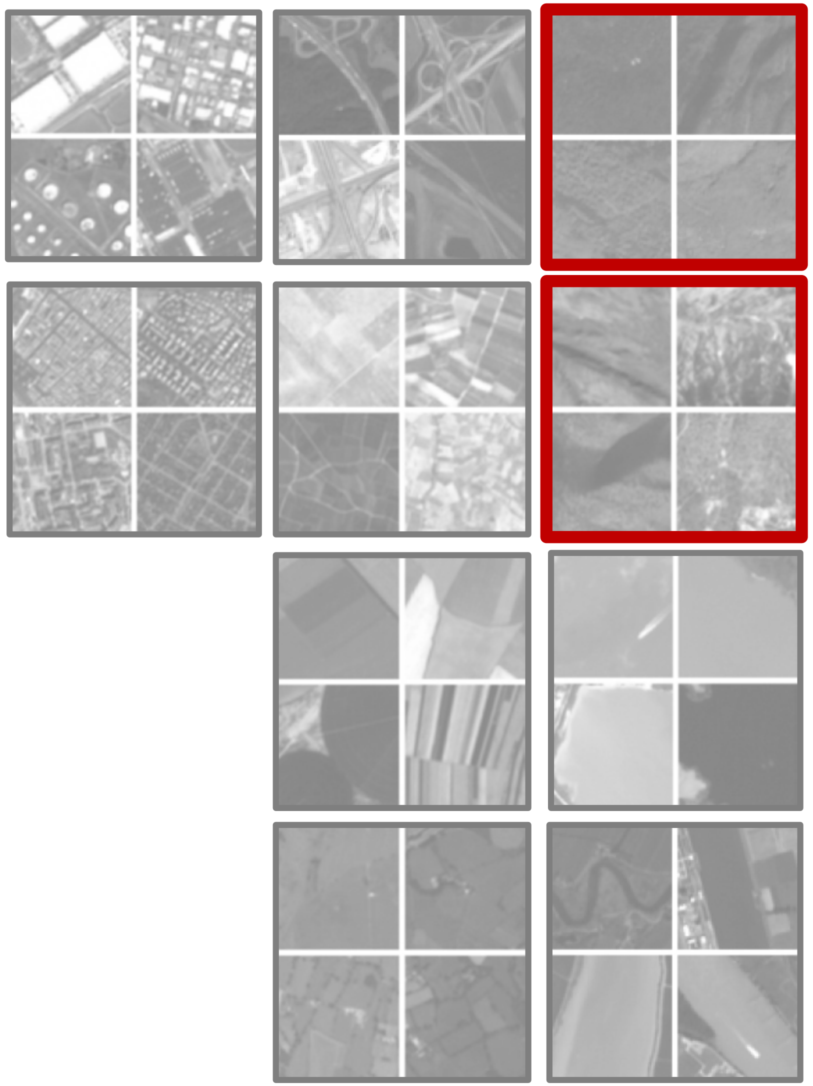
Pasture

- **Undeveloped Land**



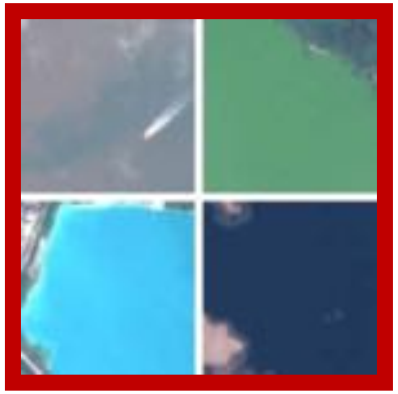
Forest

Herbaceous vegetation



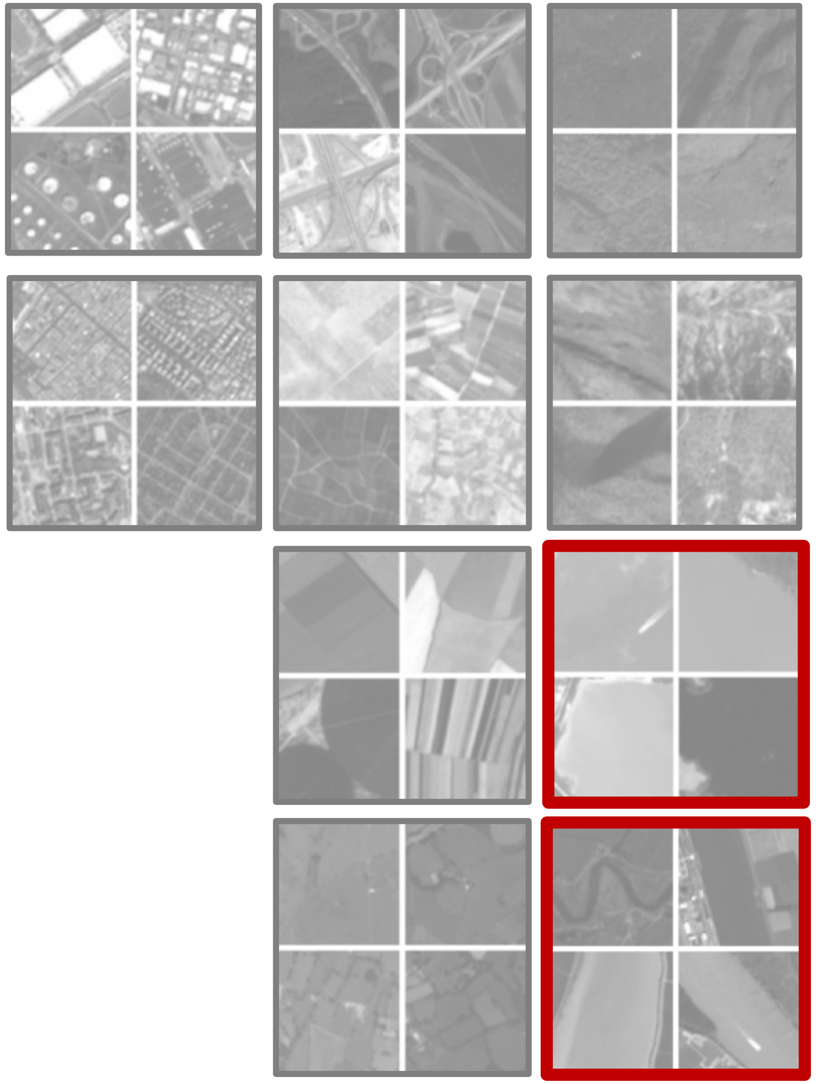
Published Dataset: EuroSAT

- **Water Bodies**



Sea & Lake

River

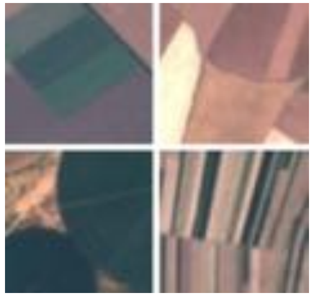


Land-Use and Land-Cover Classification

Industrial

Residential

Highway



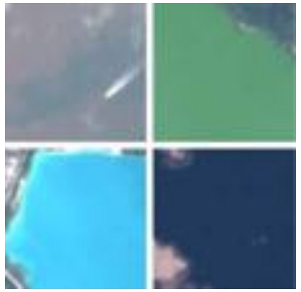
Annual Crop

Permanent Crop

Pasture

Forest

Herbaceous vegetation



Sea & Lake

River

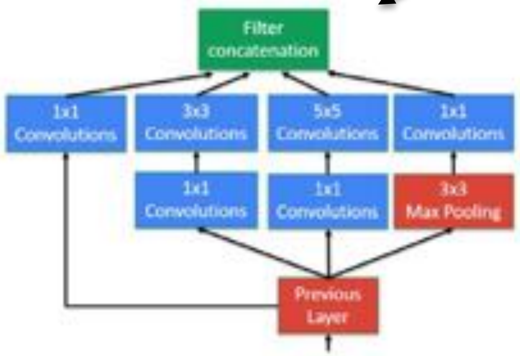
AI-based Satellite Image Analysis

Classification Pipeline Using Deep CNNs

Deep Convolutional Neural Networks

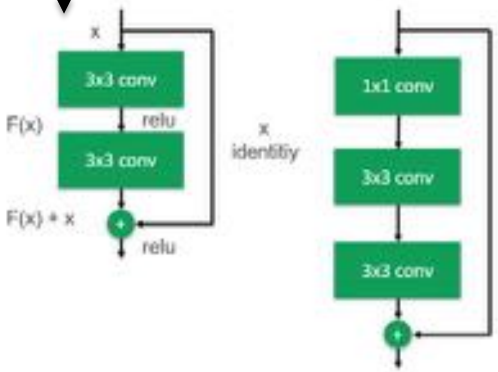
GoogLeNet

Inception Module



ResNet

Residual Building Block



Models Pre-Trained on ImageNet

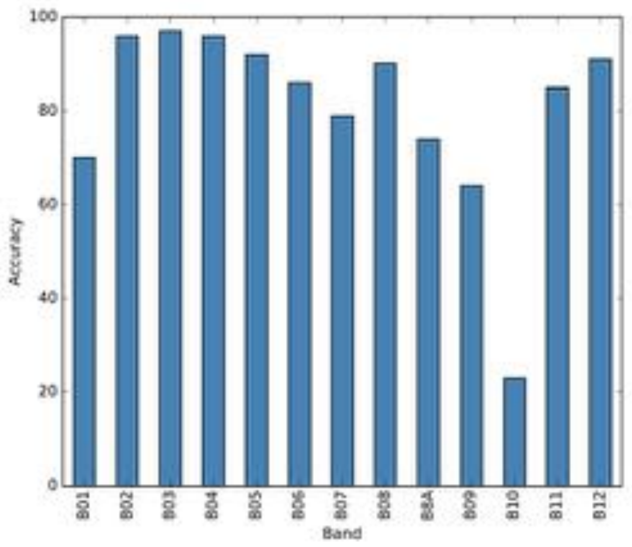
- Dataset split
  - 80% Training and 20% Testing
- Transfer learning
  - Fine-tuning of pretrained networks
- Pretrained on ImageNet
  - ILSVRC-2012 image classification challenge
- Initial learning rate:
  - 0.01 – 0.0001
- Optimizer:
  - RMSProp

Band combination	RGB	SWIR	CI
Accuracy	**0.9857**	0.9705	0.9830

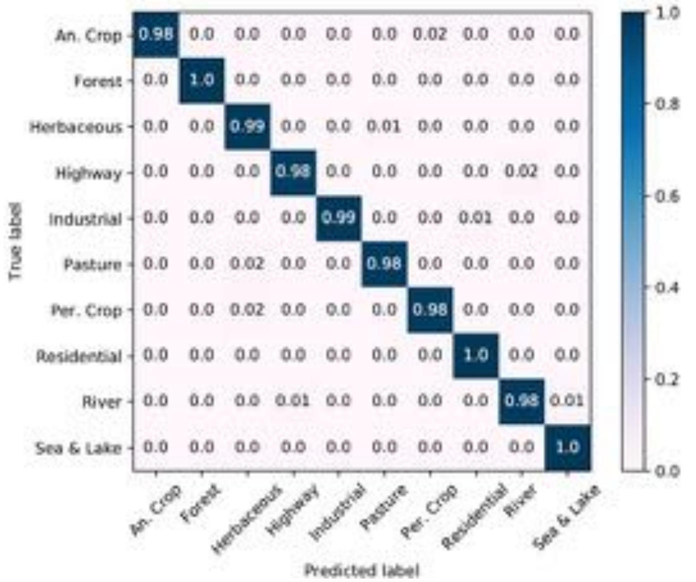
- Transfer learning
  - Fine-tuning of pre-trained networks
- Dataset split
  - 80% Training and 20% Testing
- Pre-trained on ImageNet
  - ILSVRC-2012

Band Comb.	RGB	SWIR	CI
Accuracy	**0.9857**	0.9705	0.9830

SWIR = Short-Wave-Infrared
CI = Color-Infrared



Band
B01 - Aerosols
B02 - Blue
B03 - Green
B04 - Red
B05 - Red Edge 1
B06 - Red Edge 2
B07 - Red Edge 3
B08 - NIR
B08A - Red Edge 4
B09 - Water vapor
B10 - Cirrus
B11 - SWIR 1
B12 - SWIR 2



Time-Series Analysis: Land Change Detection

Sentinel 2 – 64x64 image crops - 10 months (2017)

Spotting Land Use Changes

- **Large-scale scanning and monitoring**
  - **Time component**
    - High frequency
    - Near-real-time
    - Future availability
- **Innovative applications**
  - Building systems for future real-time applications

Residential Area Built Up
in Dallas, USA

August 2015

March 2017

Spotting Land Type Changes

Industrial Buildings Demolished
in Shanghai, China

December 2015

December 2016

Deforestation (Forest Clearing)
in Villamontes, Bolivia

October 2015

September 2016

Melbourne, Australia

Shanghai, China

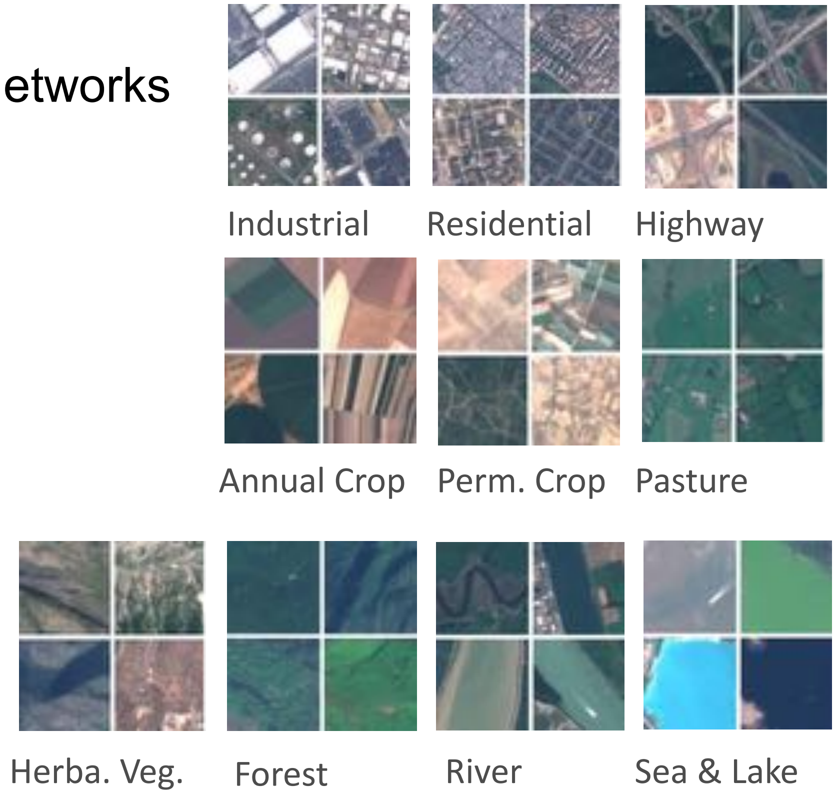
Usage of land use classification for verification of available mapping data as seen in Australia vs. China depicting industrial areas

Land-Use & Land-Cover Classification - Overview

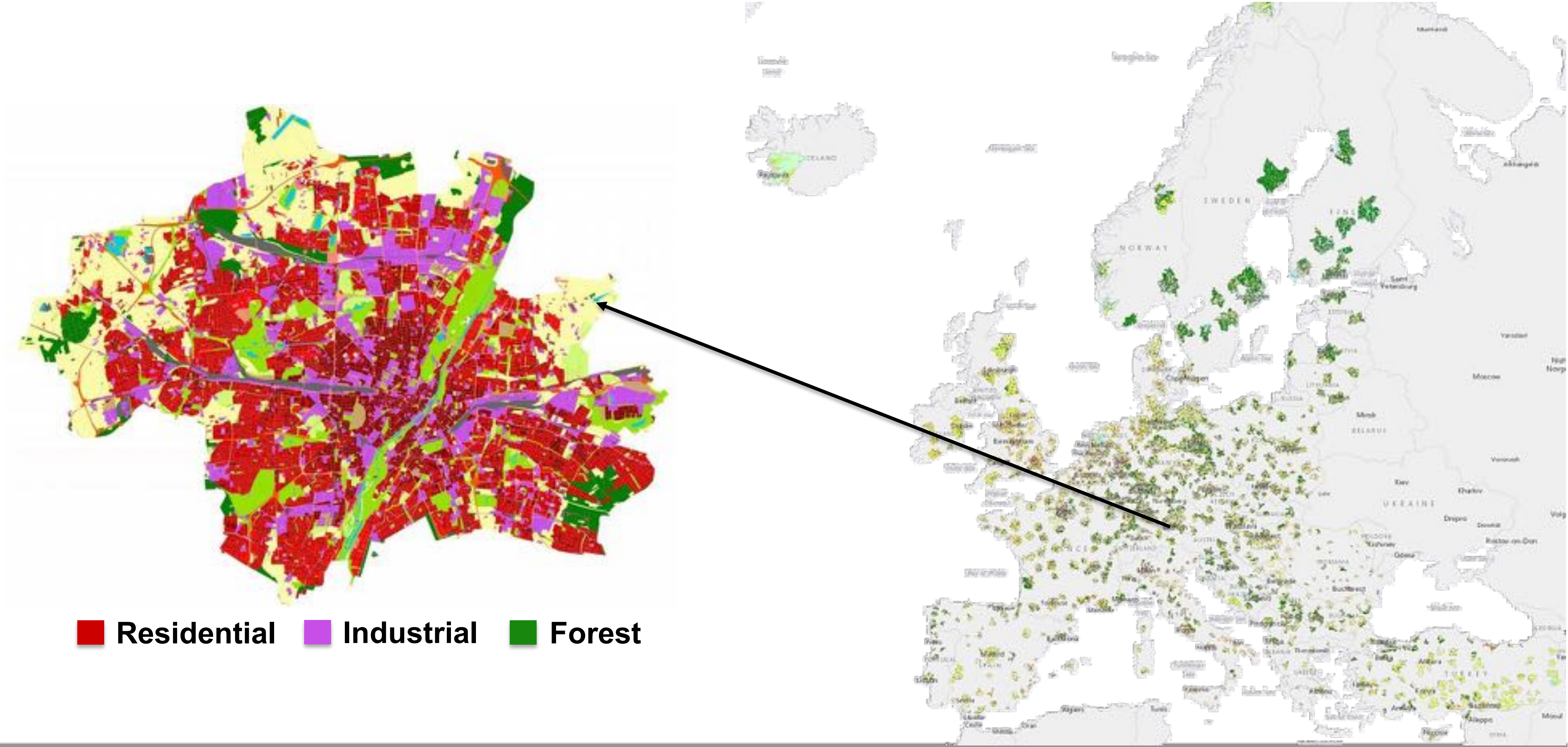
Classification Pipeline Using Deep CNNs

- Patch-Classification with state-of-the-art networks
- Satellite Images from Sentinel2 (ESA)
  - 13 spectral bands
  - Spatial resolution 10 m per pixel
- 10 Classes & 27,000 images
- Classification Accuracies above 95%

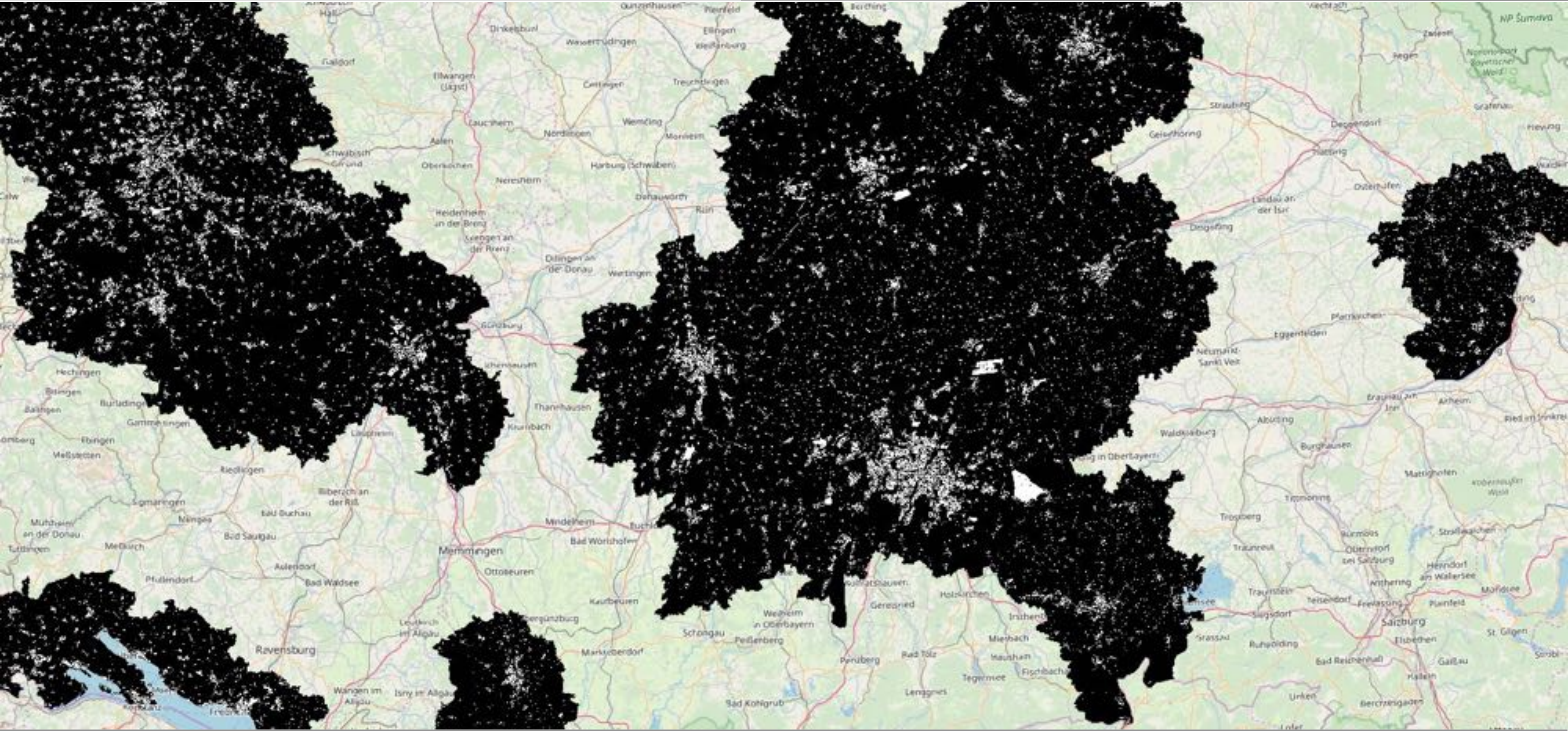
Band combination	**Accuracy (ResNet-50)**
RGB	**98.57 %**
SWIR	97.05 %
CI	98.30 %



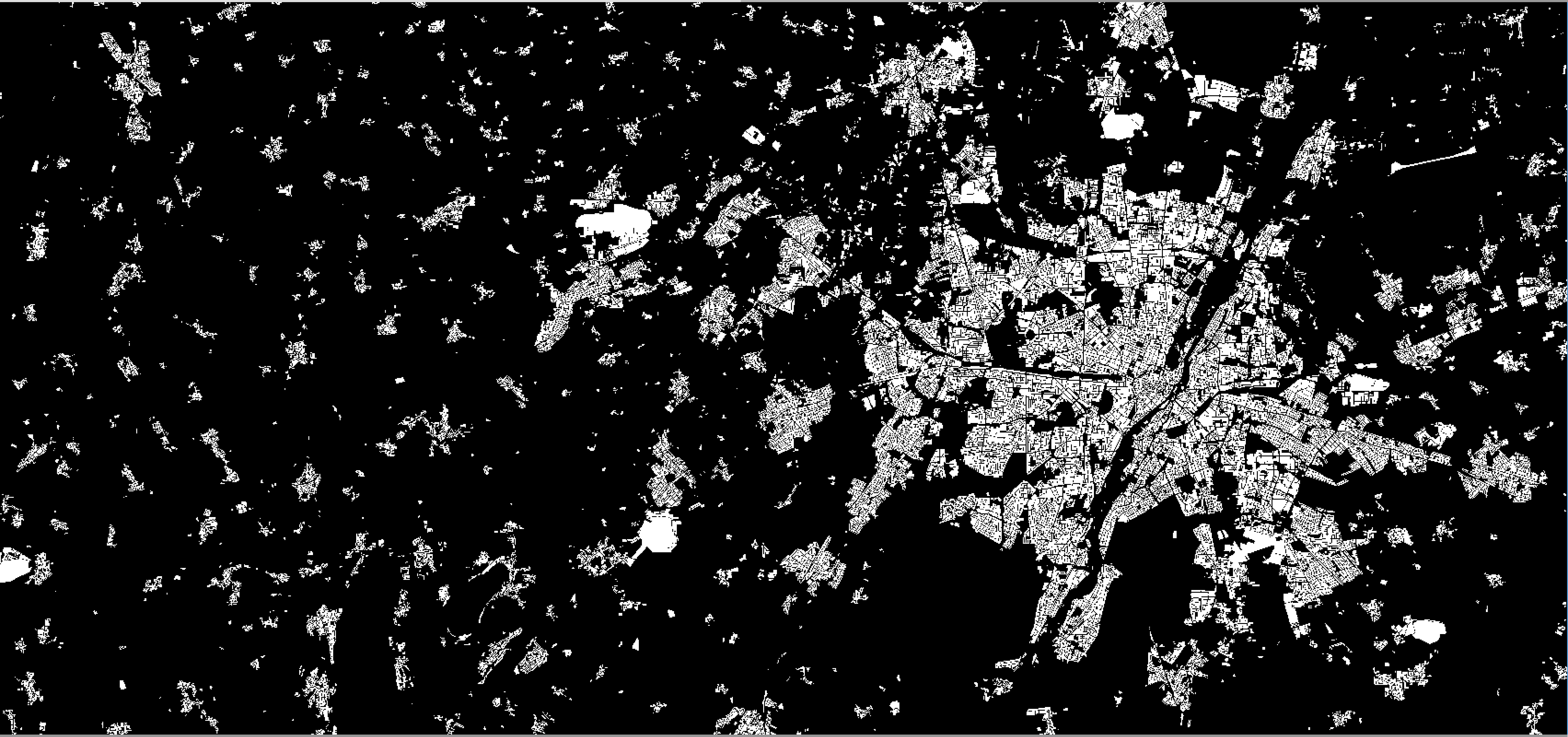
Semantic Segmentation - Mapping of 785 cities



AI-based Human Settlement Layer

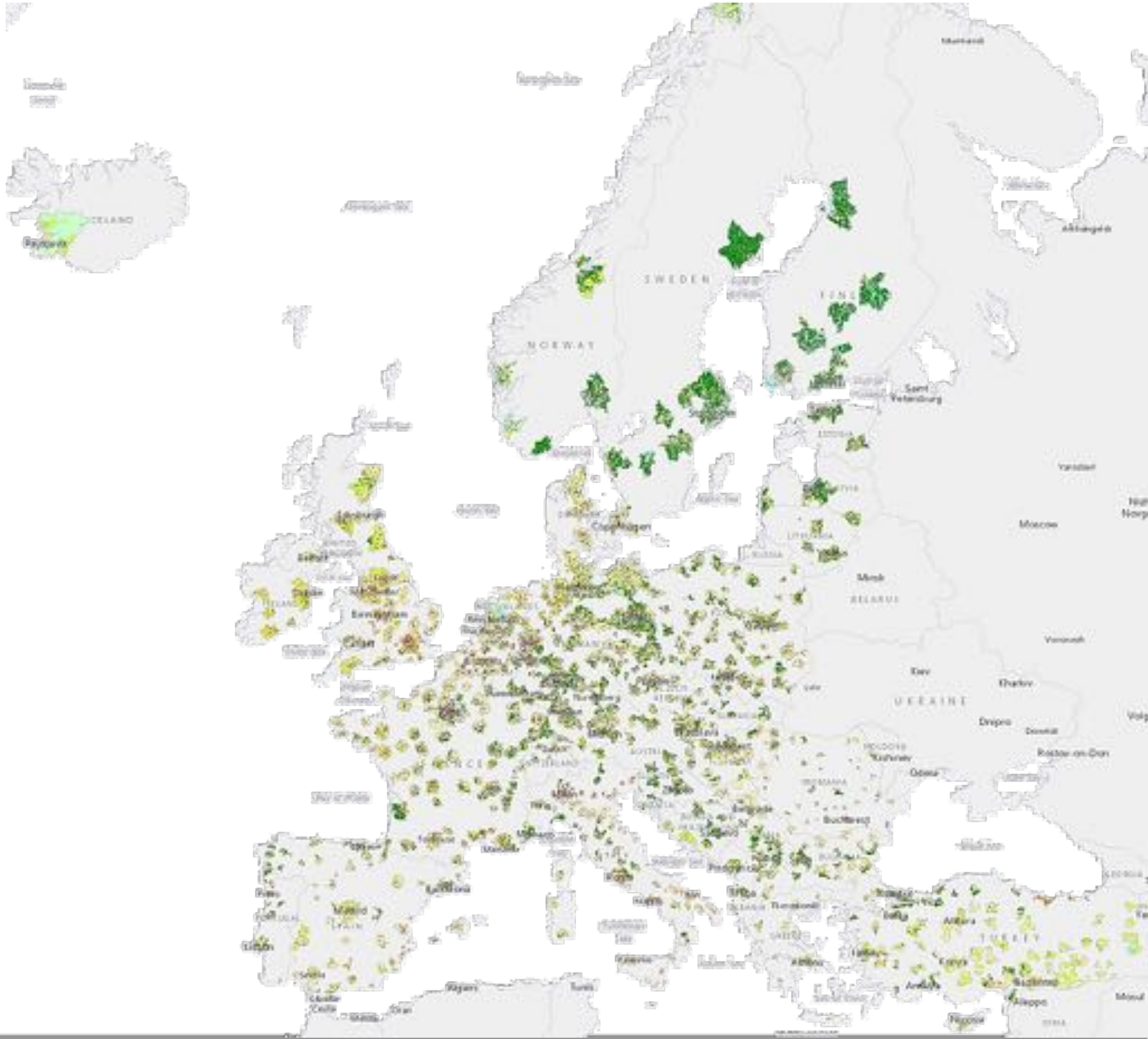


AI-based Human Settlement Layer

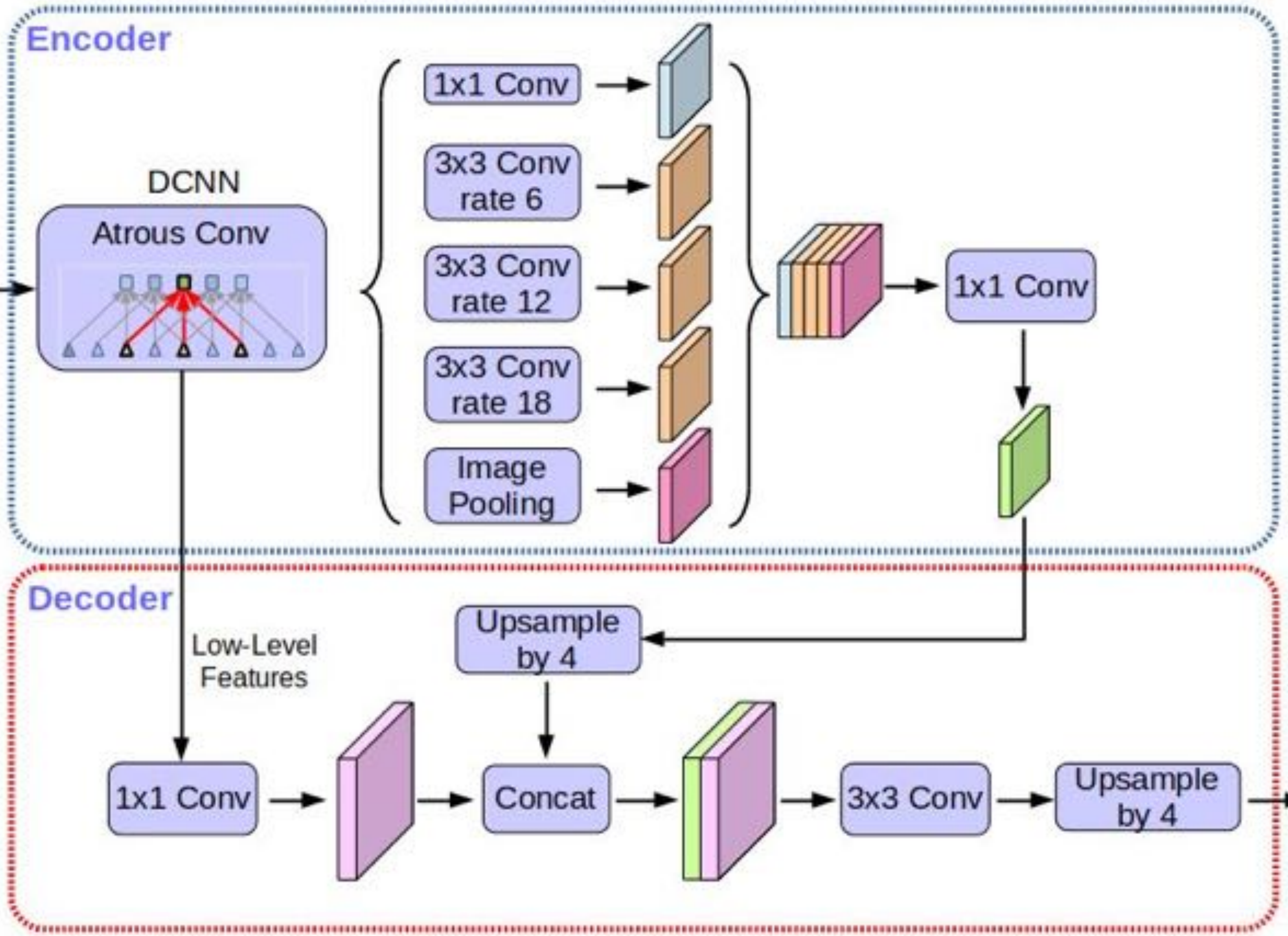


AI-based Human Settlement Layer

- ca. 90.000 RGB images
- Spatial resolution: 512 x 512
- 785 cities



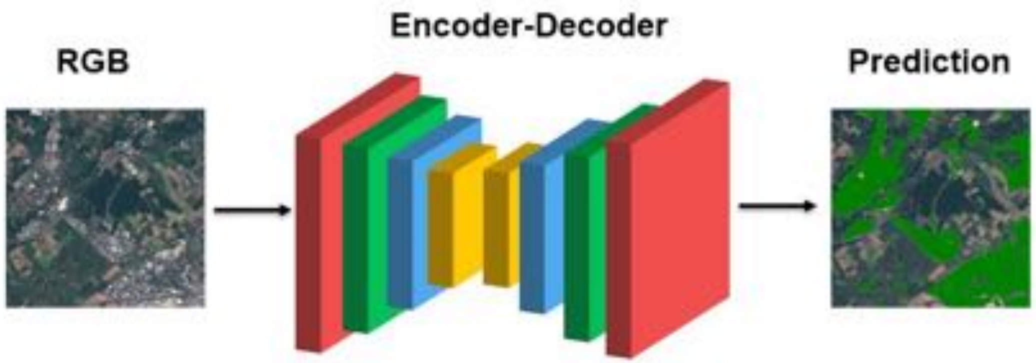
Encoder-Decoder-based Semantic Image Segmentation

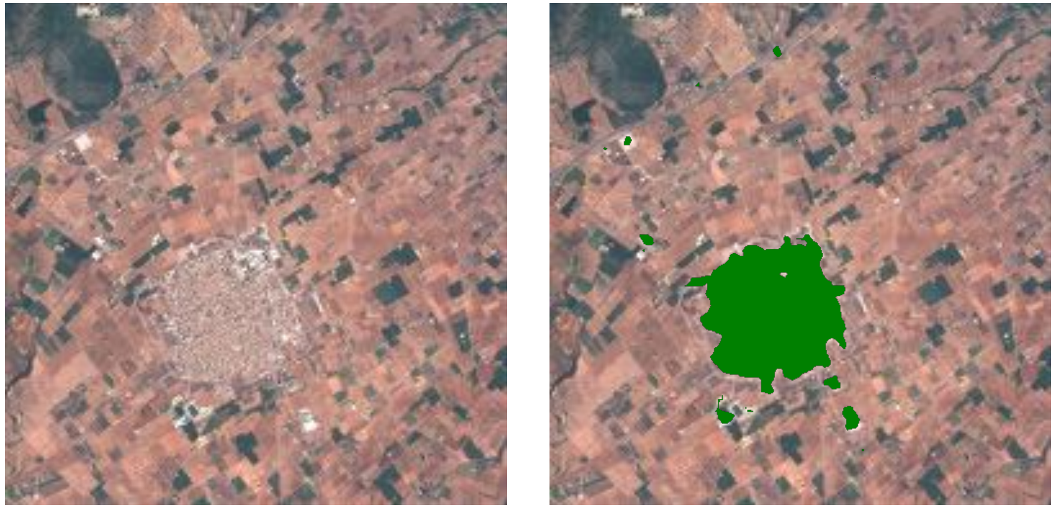


NN Input	mIOU
RGB	0.7657

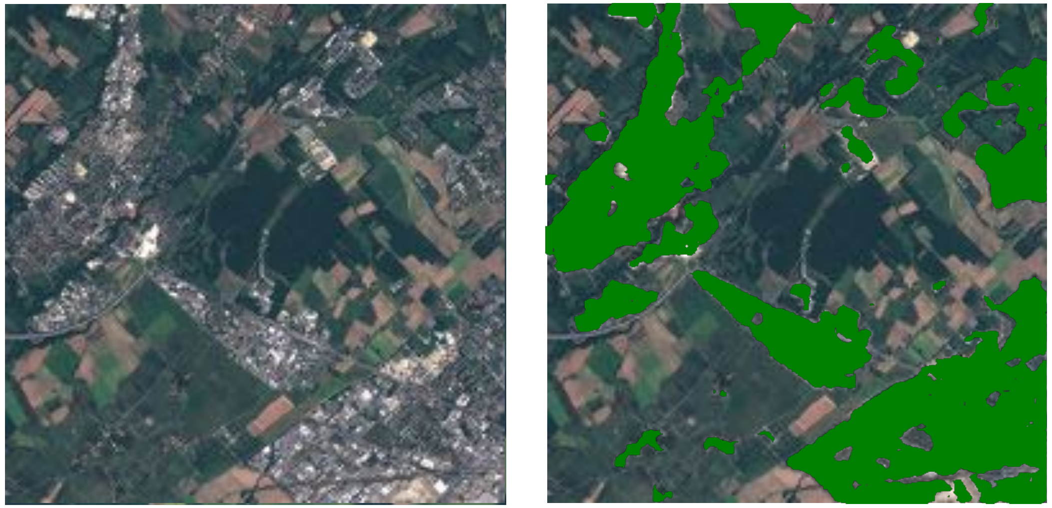
Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation*, Liang-Chieh Chen et al., arXiv: 1802.02611, 2018.*

Encoder-Decoder-based Semantic Image Segmentation



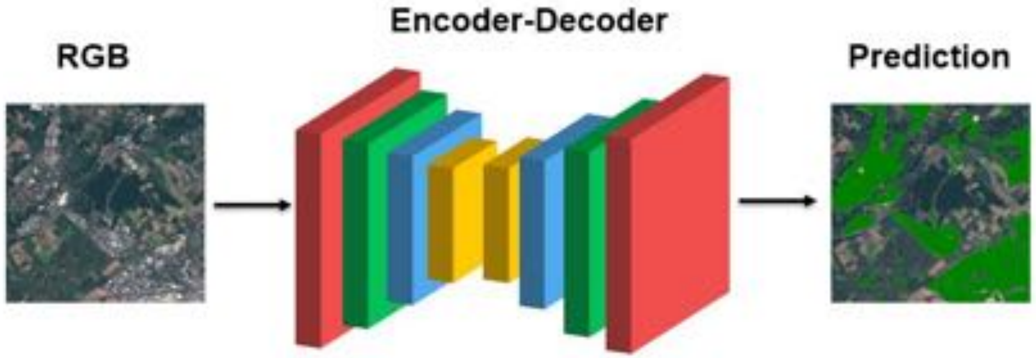


Example 1 : Urban Areas



Example 2 : Urban Areas

Encoder-Decoder-based Semantic Image Segmentation



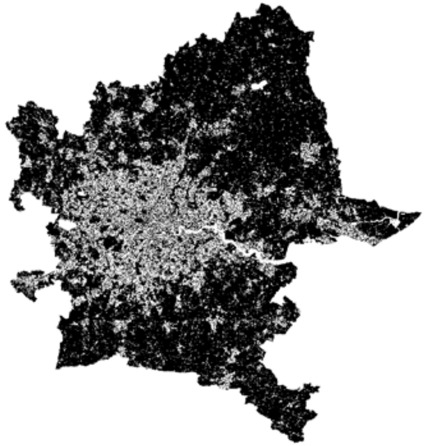
Example 1 : Rural Areas

Example 2 : Rural Areas

AI-based Human Settlement Layer

Munich

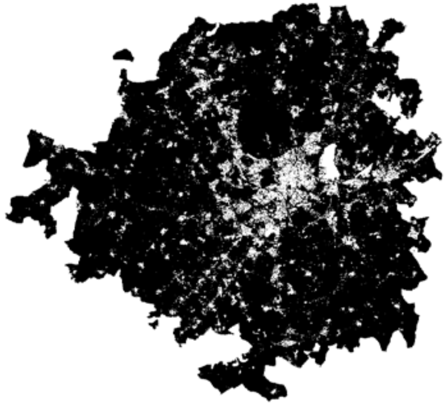
Paris



London

Roma

Brussel



Madrid

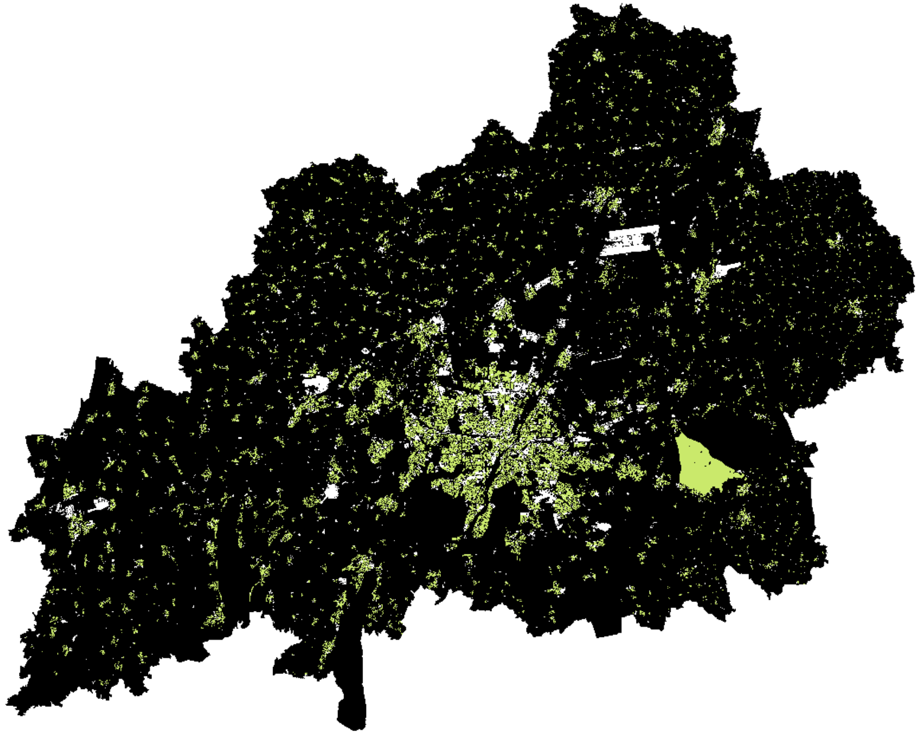
Warszawa

Sofia

AI-based Human Settlement Layer – Next Steps?

Human Settlements

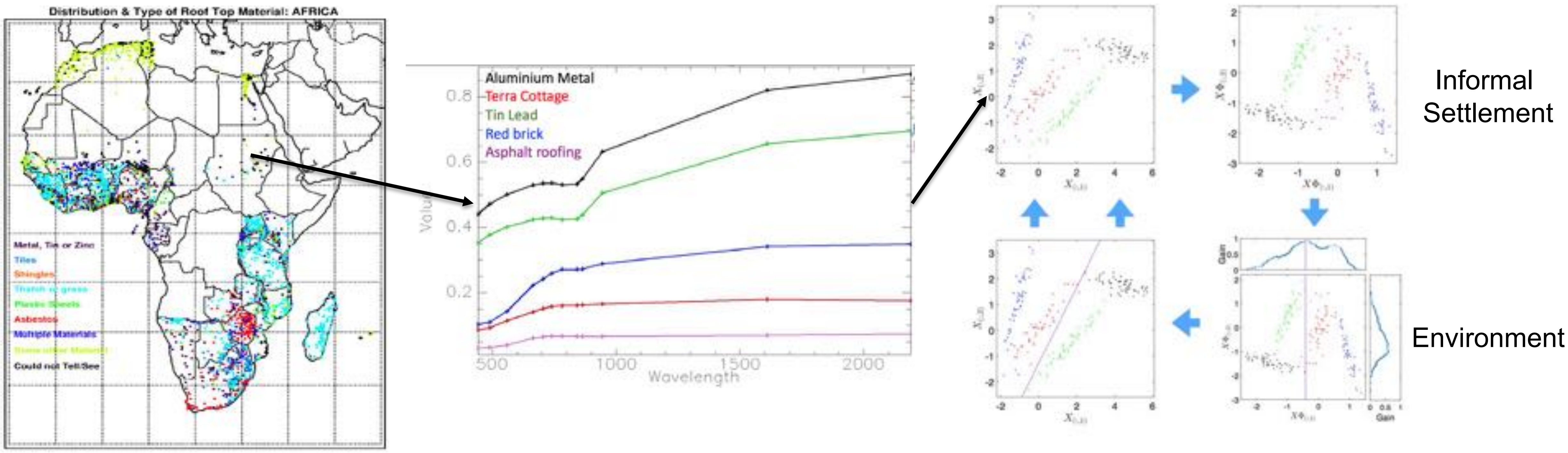
Population Estimation?



Residential vs. Industrial

Different Forms of Residential Areas?

- **Mapping and Detecting the Locations of Informal Settlements**
  - Session: FDL Europe ESA AI4EO Accelerator (ID: **301**)
  - Wed, 14.11.2018 AI4EO (Part5)
  - 09:35 - 09:50, MAGELLAN



Thanks!

NVIDIA AI Lab partner
all networks trained on DGX-1

Overview - Deep Learning in Earth Observation

Application Oriented Research

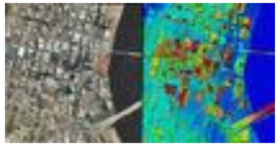
DeepEye for Natural Disasters

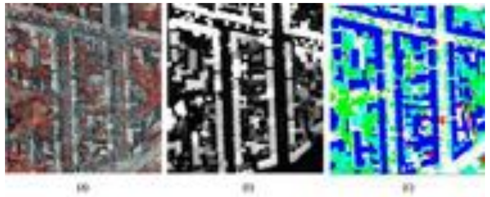
Multimedia Satellite Task 2017/2018

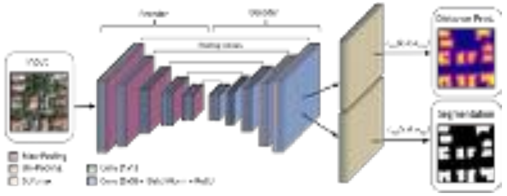
Land-Use and Land-Cover Classification (Patch based)

Estimation of (Micro)-Economic Factors

Land Use and Land-Cover Segmentation







Missing Data during Inference

Multimodal Fusion of Satellite Data

Multi-Task Learning for Sem. Segmentation

Fundamental Research

Overview - Deep Learning in Remote

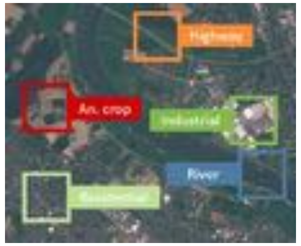
Application Oriented Research

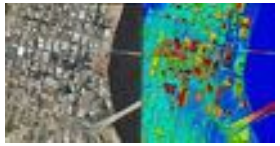
DeepEye for
Natural Disasters

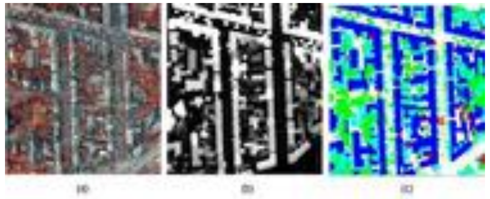
Multimedia Satellite
Task 2017/2018

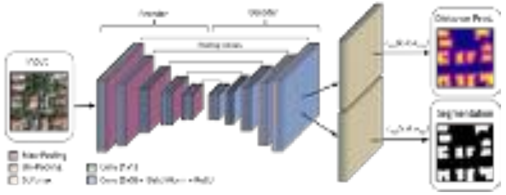
Land-Use and Land-
Cover Classification
(Patch based)

Estimation of (Micro)-
Economic Factors









Missing Data during
Inference

Multimodal Fusion of
Satellite Data

Multi-Task Learning
for Sem. Segmentation

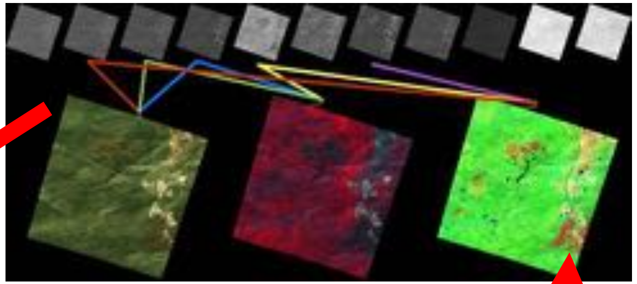
Fundamental Research

DeepEye - Social Media and Satellite Imagery

- Combination of Social Media Analysis and Satellite Image Processing for Natural Disasters with NASA's Landsat 8 Satellite

Deep Eye Visualisation Browser

Bischke, B., Borth, D., Schulze, C., and Dengel, A., 2016. *Contextual enrichment of remote-sensed events with social media streams*. In Proceedings of the ACM Multimedia Conference (Amsterdam, Netherlands 15-19 October 2016).



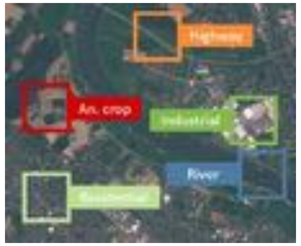
Satellite Band Analysis of Landsat8

NASA, May 2016

Overview - Deep Learning in Remote Sensing

Application Oriented Research

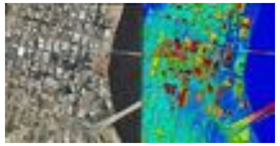
Land-Use and Land-Cover Classification (Patch based)



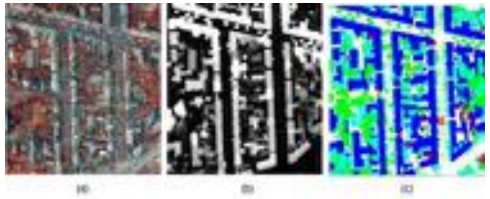
DeepEye for Natural Disasters

Multimedia Satellite Task 2017/2018

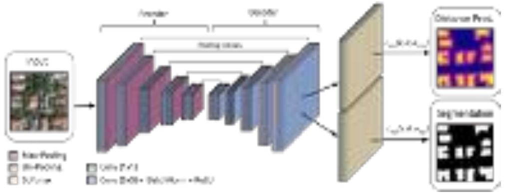
Estimation of (Micro)-Economic Factors



Missing Data during Inference



Multimodal Fusion of Satellite Data



Multi-Task Learning for Sem. Segmentation

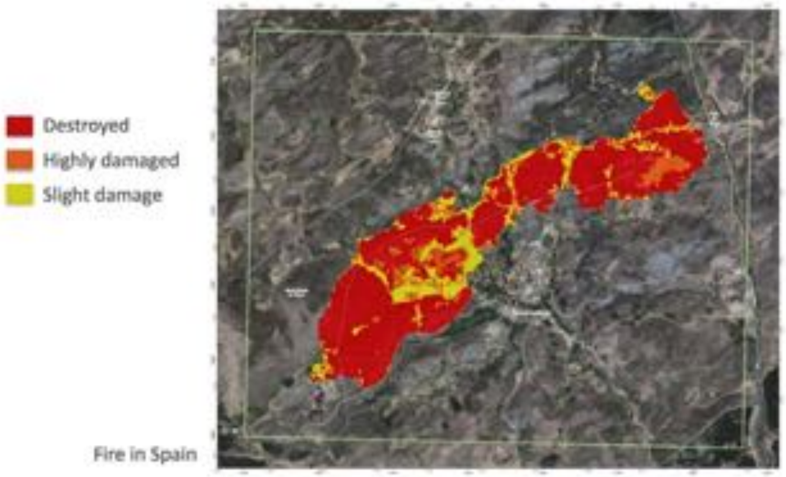
Fundamental Research

Multimedia Satellite Task 2017

- Lead Organisers of the Multimedia Satellite Task 2017 (with Virginia Tech & Queensland Uni.) at Multimedia Eval
- 15 Teams registered from all the world (Brasil, Australia, Greece, Brunei, Italy, UK, Germany, Netherlands, Norway, Pakistan)
- More than 60 submission on two subtasks

MediaEval Workshop

Damage Estimation



EMS - Copernicus

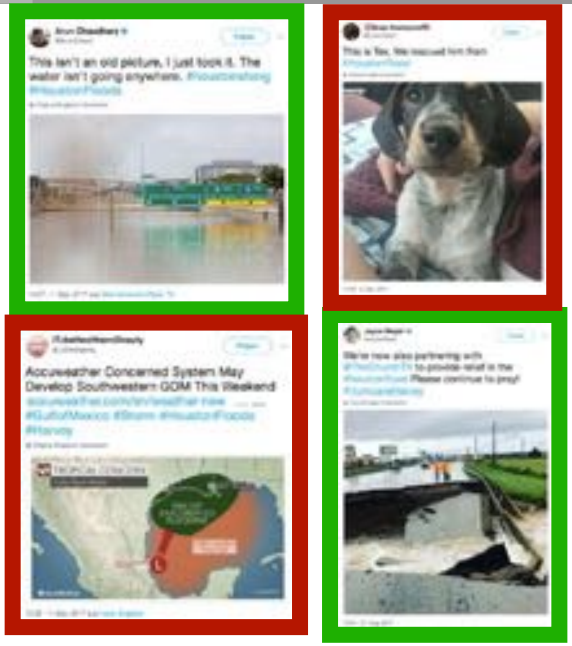
Emergency Response

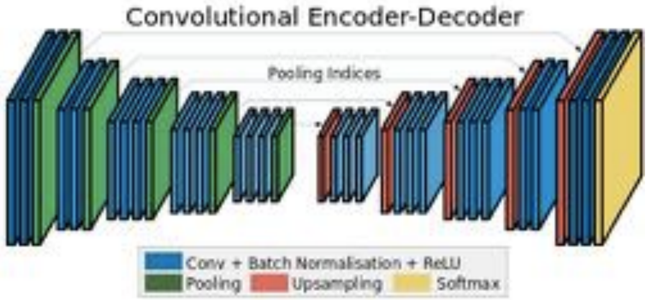
DigitalGlobe, October 2017

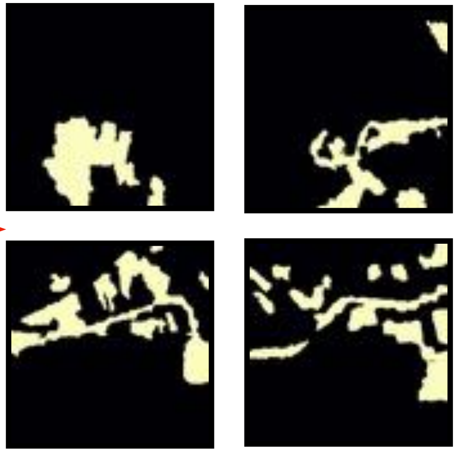
Bischke, Benjamin, et al. "The multimedia satellite task at mediaeval 2017: Emergence response for flooding events." Proc. of the MediaEval 2017 Workshop (Sept. 13-15, 2017). Dublin, Ireland. 2017.

Multimedia Satellite Task 2017

- Main Focus on **Flooding Events**
  - Retrieval of Flood related Reports/Images from Social Media Streams
  - Segmentation of Flooded Areas in Satellite Imagery (Satellite Imagery from Planet) with Deep Neural Networks







Bischke, Benjamin, et al. "Detection of flooding events in social multimedia and satellite imagery using deep neural networks." Working Notes Proc. MediaEval Workshop. 2017.

Multimedia Satellite Task 2018

- Continue with **Flooding Events:**
  - Focus on Impact Estimation of Infrastructure (Road Access, blocked Road)
  - Two Subtasks:
    1. Classification of Road-Access & Passability in Social Multimedia
    2. Semantic Segmentation of Roads/blocked Roads in various multiple Satellite Images (Radar, Optical)

Images in Tweets (Hurricane Harvey)

Satellite Images of Houston (US) for Hurricane Harvey

Combine Result via Geo-Location

Overview - Deep Learning in Remote Sensing

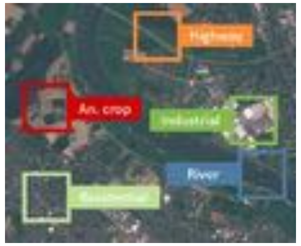
Application Oriented Research

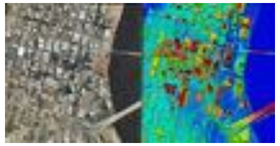
DeepEye for
Natural Disasters

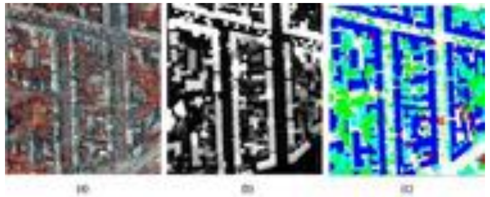
Multimedia Satellite
Task 2017/2018

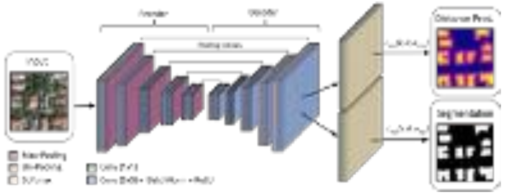
Land-Use and Land-
Cover Classification
(Patch based)

Estimation of (Micro)-
Economic Factors









Missing Data during
Inference

Multimodal Fusion of
Satellite Data

Multi-Task Learning
for Sem. Segmentation

Fundamental Research

Overview - Deep Learning in Remote Sensing

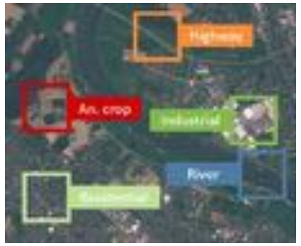
Application Oriented Research

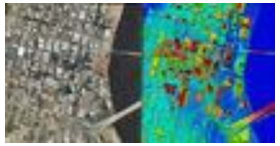
DeepEye for
Natural Disasters

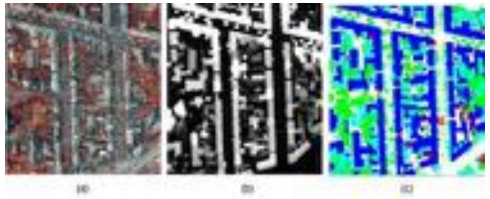
Multimedia Satellite
Task 2017/2018

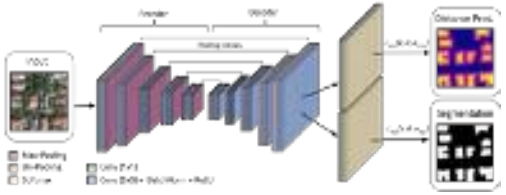
Land-Use and Land-
Cover Classification
(Patch based)

Estimation of (Micro)-
Economic Factors









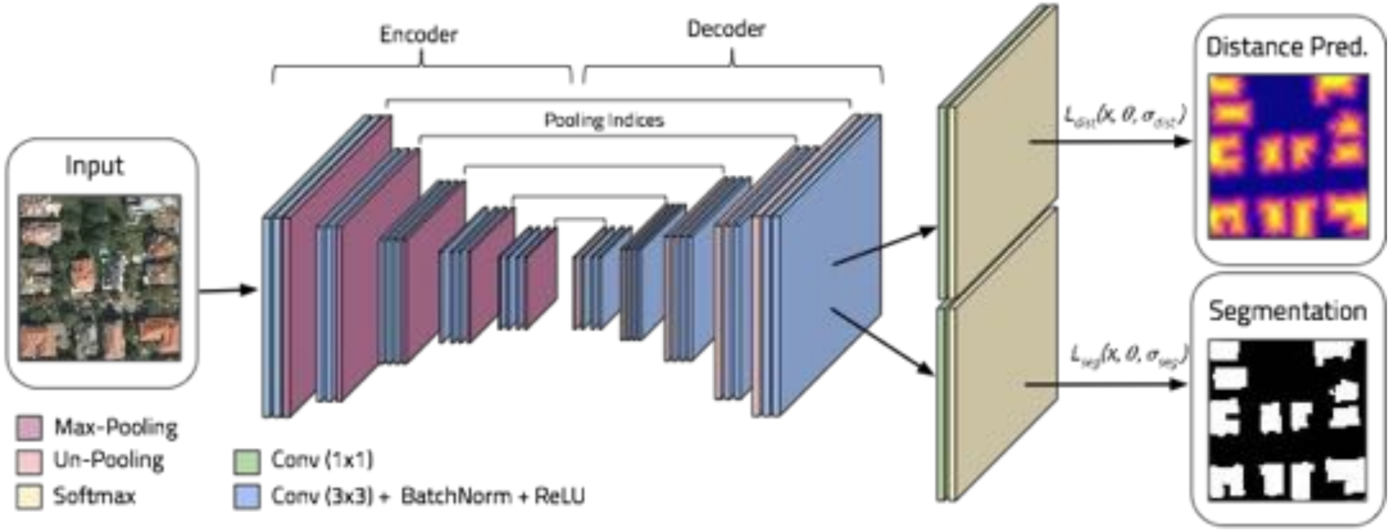
Missing Data during
Inference

Multimodal Fusion of
Satellite Data

Multi-Task Learning
for Sem. Segmentation

Fundamental Research

Multi-Task Learning to Improve Semantic Segmentation



- Multi-Task Learning with multiple output representations
- Learn Task based uncertainty weights
- Improves the semantic segmentation predictions near boundaries

$$L_t(x,\theta,\sigma_t) = \sum_{c=1}^{C} -C_c log P(C_c = 1|x,\theta,\sigma_t)$$

$$= \sum_{c=1}^{C} -C_c log(exp(\frac{1}{\sigma_t^2} f_c(x))) + log \sum_{c'=1}^{C} exp(\frac{1}{\sigma_t^2} f_{c'}(x)) \tag{6}$$

Applying the same assumption as in [25]:

$$\frac{1}{\sigma^2} \sum_{c'} exp(\frac{1}{\sigma^2} f_{c'}(x)) \approx (\sum_{c'} exp(f_{c'}(x)))^{\frac{1}{\sigma^2}} \tag{7}$$

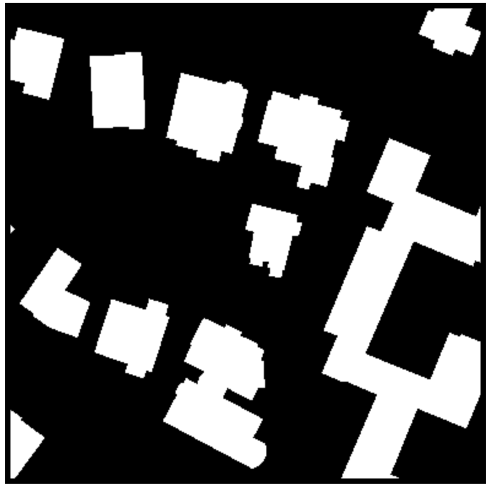
allows to simplify Eq. 6 to:

$$L_t(x,\theta,\sigma_t) \approx \frac{1}{\sigma_t^2} \sum_{c=1}^{C} -C_c log P(C_c = 1|x,\theta) + log(\sigma_t^2) \tag{8}$$

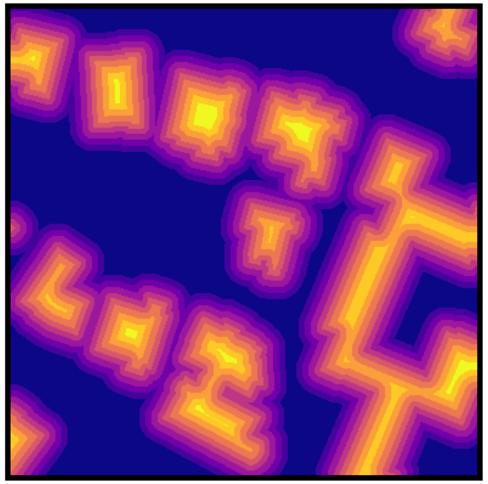
Bischke, B., Helber, P., Folz, J., Borth, D., & Dengel, A. (2017). Multi-Task Learning for Segmentation of Building Footprints with Deep Neural Networks. *arXiv preprint arXiv:1709.05932*.

Multi-Task Learning - Qualitative Results

Input

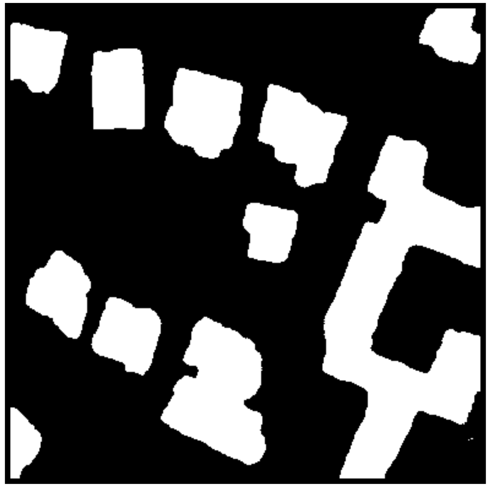


Ground Truth

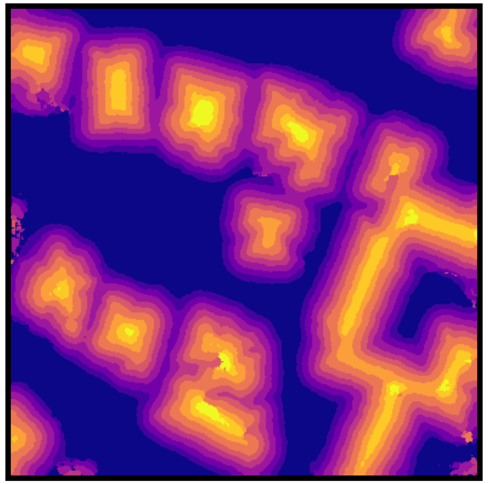


Ground Truth (Distances)

Predicted SegNet



Predicted MultiTaskNet



Predicted MultiTaskNet

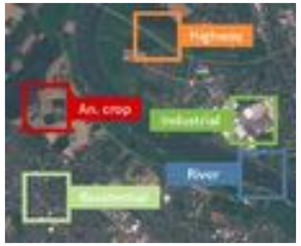
Overview - Deep Learning in Remote Sensing

Application Oriented Research

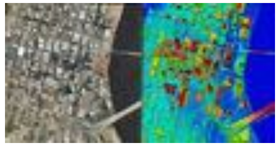
DeepEye for
Natural Disasters

Multimedia Satellite
Task 2017/2018

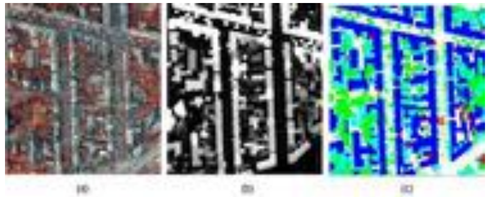
Land-Use and Land-
Cover Classification
(Patch based)



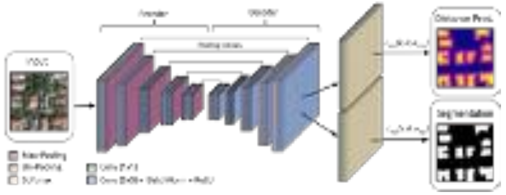
Estimation of (Micro)-
Economic Factors



Missing Data during
Inference



Multimodal Fusion of
Satellite Data

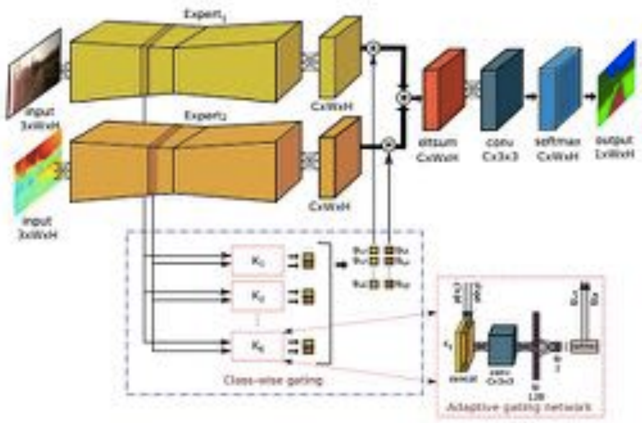


Multi-Task Learning
for Sem. Segmentation

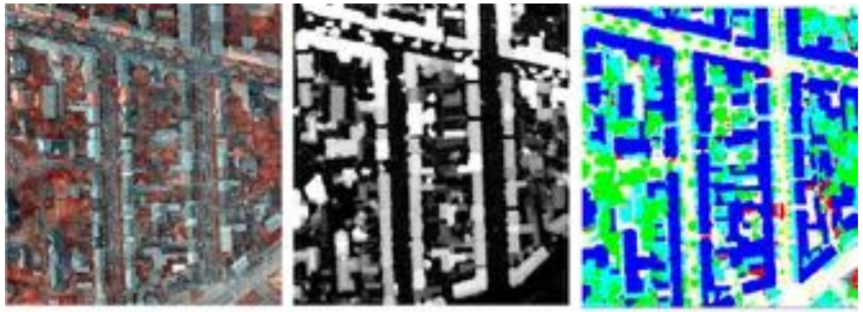
Fundamental Research

Multimodal Fusion in Deep Neural Networks

- How-to fuse multiple views of a particular region?
  - Multiple Satellites (Optical, Radar)
  - Multiple Sensors (Depth, RGB)
  - Domain knowledge (False-Color Images)
- Research on novel approaches for Network Fusion
  - Unsupervised Methods
  - Attention Guided Methods

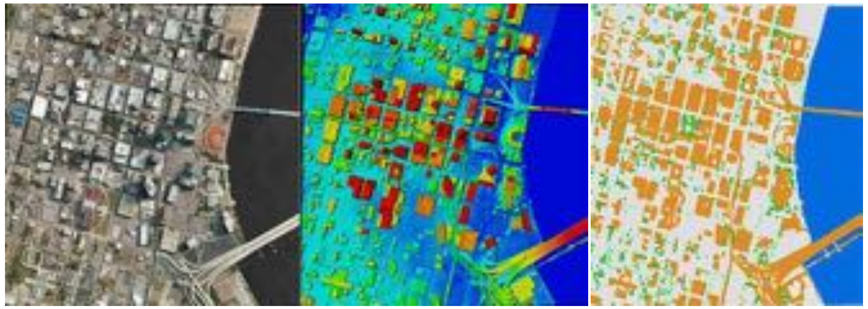


NIR Depth Labels



(a) (b) (c)

RGB 3D-Depth Labels



Overview - Deep Learning in Remote Sensing

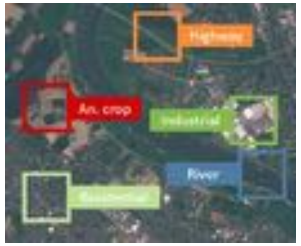
Application Oriented Research

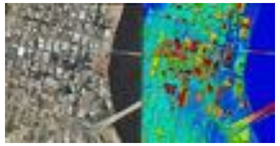
DeepEye for
Natural Disasters

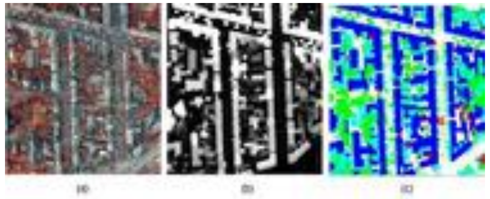
Multimedia Satellite
Task 2017/2018

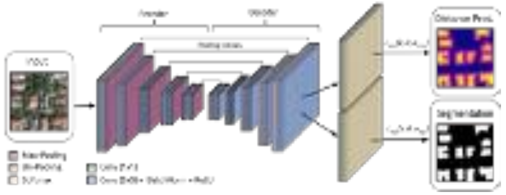
Land-Use and Land-
Cover Classification
(Patch based)

Estimation of (Micro)-
Economic Factors







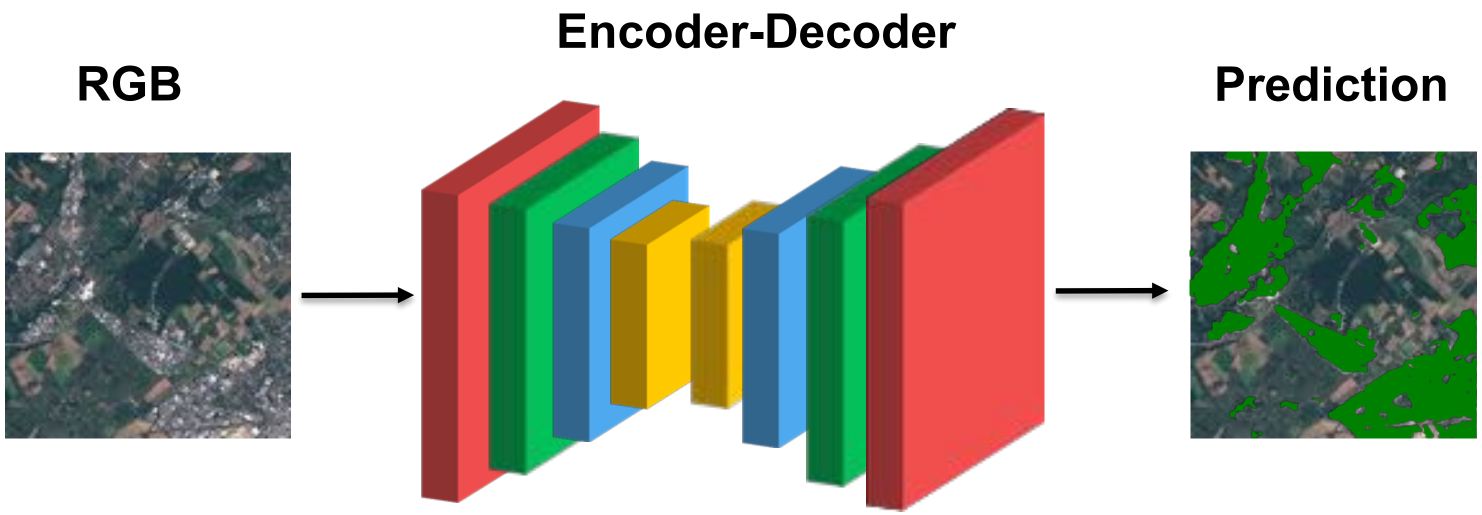


Missing Data during
Inference

Multimodal Fusion of
Satellite Data

Multi-Task Learning
for Sem. Segmentation

Fundamental Research


Encoder-Decoder
RGB
Prediction