National Research Foundation of Korea
European Commission

ACRS – EOPEN Special Session, 17 Oct. 2019

Platform Based Deep-Learning Applications as a Solution for Satellite Big Data Analysis : The Case of Rice Paddy Detection in South Korea

HyunWoo Jo
e-GIS/RS Center, Korea Univ.

European Commission | Horizon 2020 European Union funding for Research & Innovation

National Research Foundation of Korea

EOPEN

Overview

- Background
- Data Preparation
- Method
- Results
- EOPEN Platform

From Data to Information

- Decision makers need “information” not data
  - Proper data processing under researchers’ understanding can change data into information
  - Handling data is being harder as the volume and complexity of data increases
- Big Data consists of massive satellite images
  - Increasing spatio-temporal resolution
  - Collecting data from wide range of area

Limitations for Big data Processing

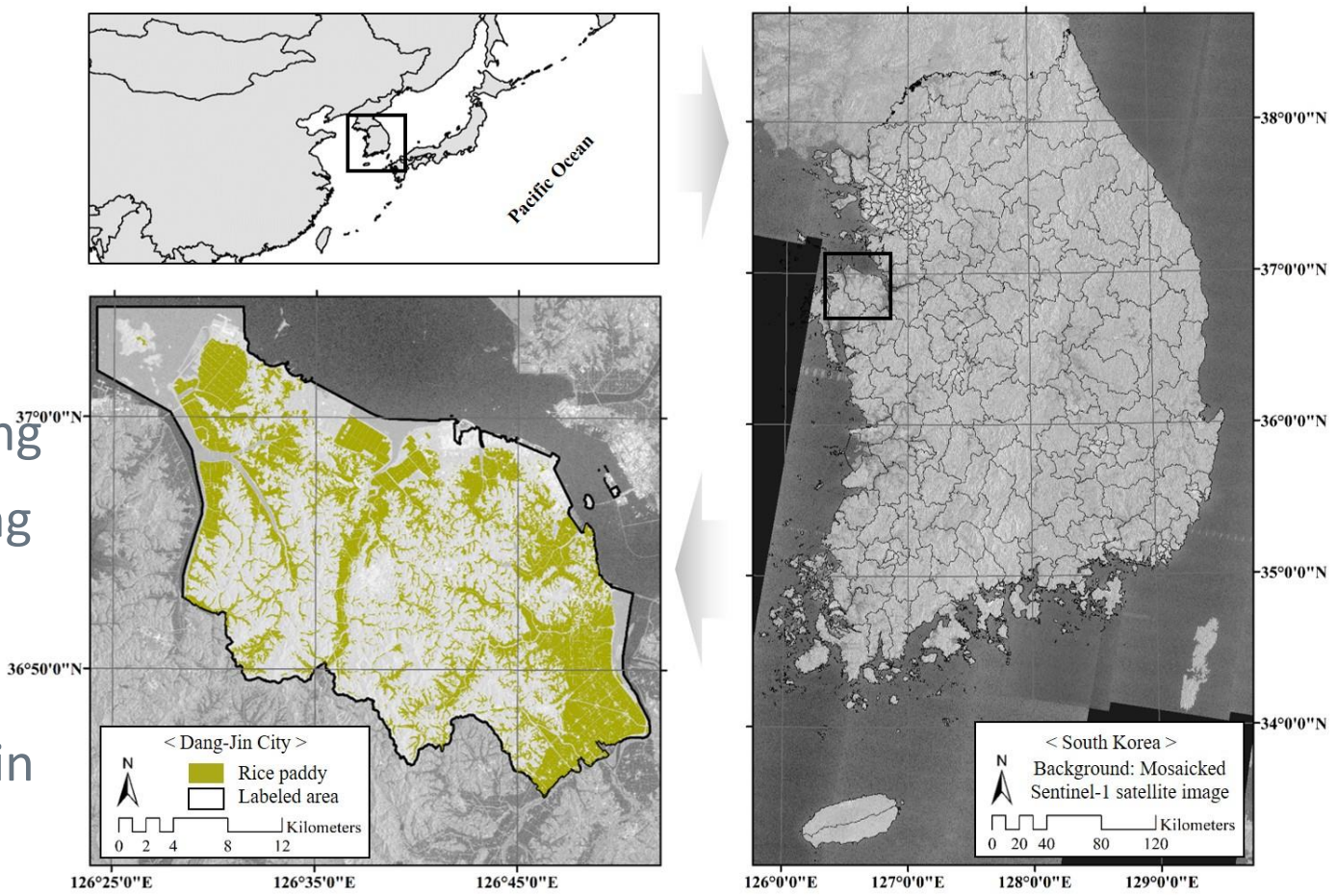
- Technical barriers
  - High performance hardware for Big Data processing
  - Infrastructure for downloading/storing massive data
- Imbalance between Big Data and labeling data
  - Data should be labeled to be trained in the model
  - Satellite images are acquired in near-real-time over wide areas, while labeling data (land cover map etc.) are not updated rapidly
  - Using incorrect labeling data could increase error

Solutions for Big Satellite Data Processing

- Platform based data downloading / storing / processing
  - Sharing high performance server which individual cannot afford
  - Supporting big data processing with the advanced ICT solutions
- Deep-learning analysis with Data science techniques
  - Extract information from big data with neural networks which provides complex, non-linear modeling
  - Appling techniques to overcome the shortage of labeling data

Rice Paddy Detection in South Korea

- Ideal testbed for suggested Big Data processing solution
- Fragmented rice paddy in Asia & Europe needs high-resolution monitoring
- Non-linear modeling of deep-learning is effective for exploiting rice phenology appears in time series dataset



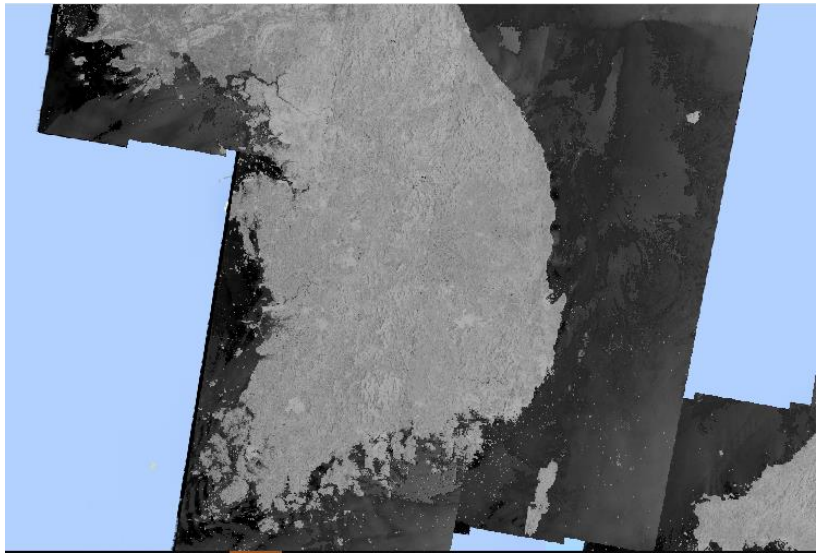
Rice paddy in Dang-Jin, Level-3 land cover map

Sentinel – 1 Acquired from Google Earth Engine

- Active sensing (available for all-weather)
  - Advantage on acquiring time series data
- Sensitive to textural information
  - Extreme contrast of reflectance according to rice existence

- Preprocessed products are downloadable

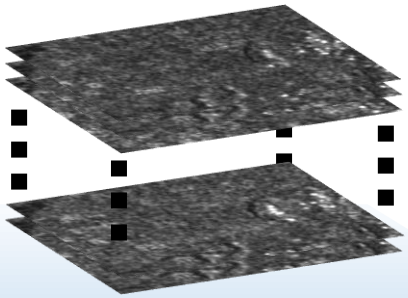
1. Apply orbit file
2. GRD border noise removal
3. Thermal noise removal
4. Radiometric calibration
5. Terrain correction (ortho rectification)

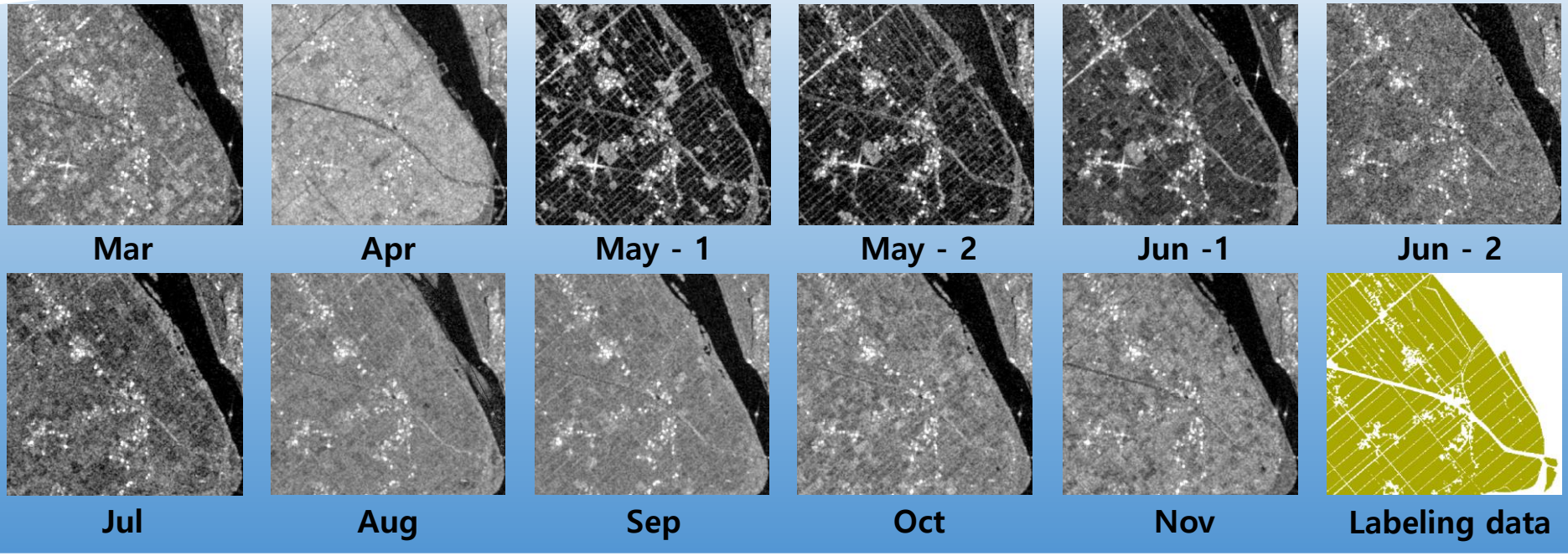


Modified Code removing outlier (Border line)

Time Series Dataset

- Monthly data are acquired with mean value
- A year of time series data set with rice paddy labeling is learned in deep-learning model

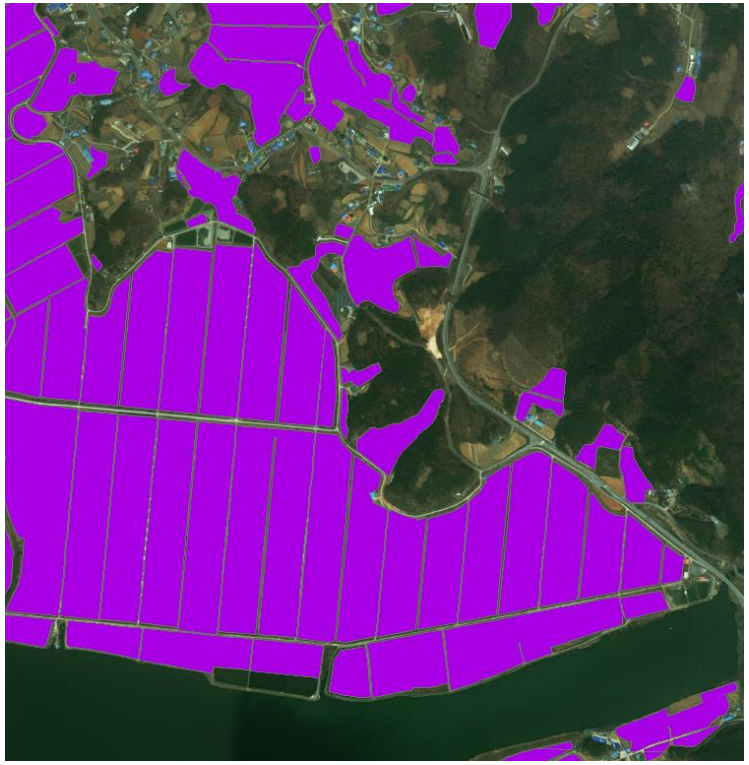




Labeling Data: level-3 land cover map

- Rice paddy mapped in parcel level
- Only on-demand regional update

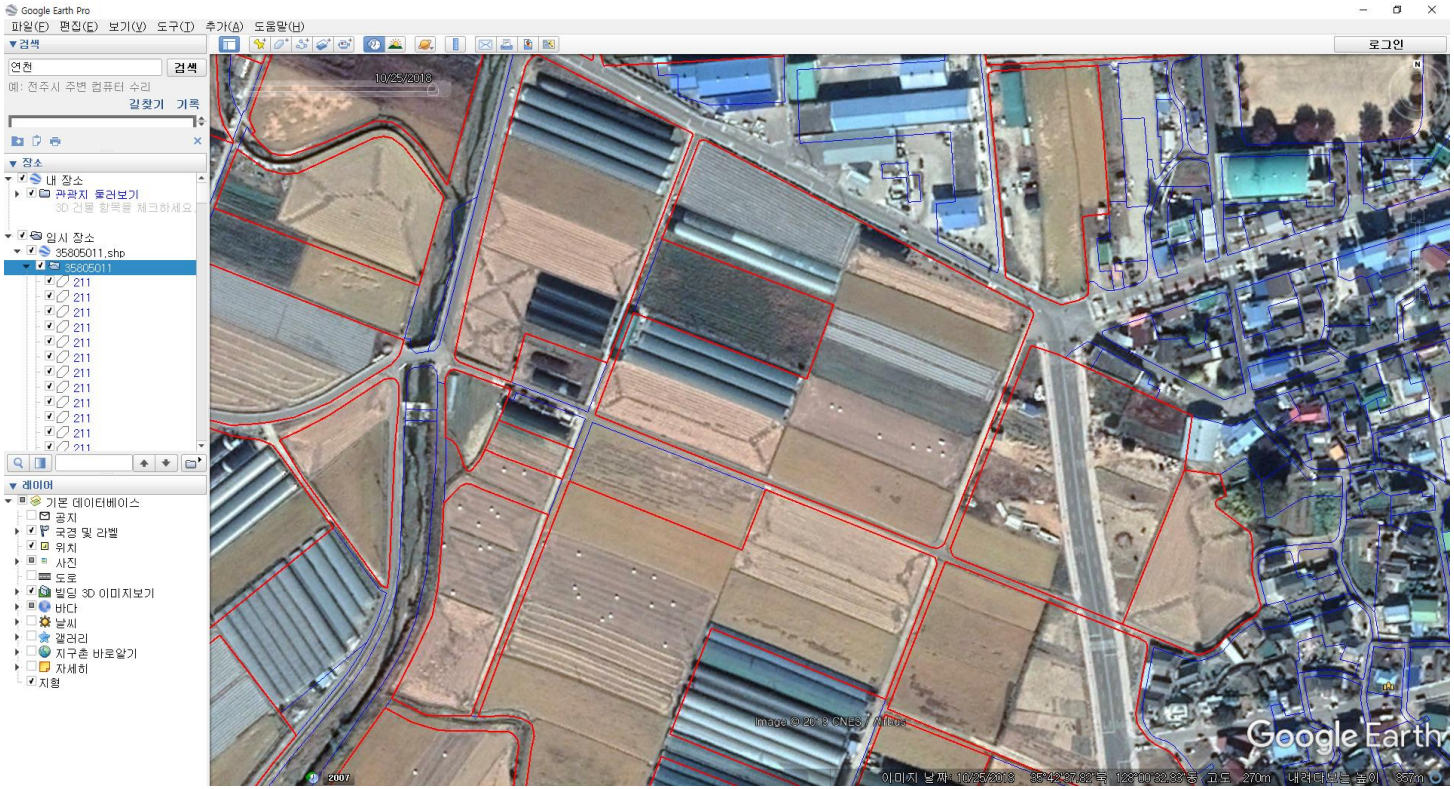
Level	Scale (resolution)	Format	Class	Materials	Update
Level – 1	1 : 50,000 (30m)	Raster	7	Landsat TM	Every 10 years (whole nation)
Level – 2	1 : 25,000 (5m)	Shape	22	Landsat TM / SPOT5 / KOMPSAT-2	About every 4 years (regionally, province level)
Level – 3	1 : 5,000 (1m)	Shape	41	KOMPSAT-2 / IKONOS / Arial photo	Every years (regionally, city level)



Level-3 land cover map Rice Paddy

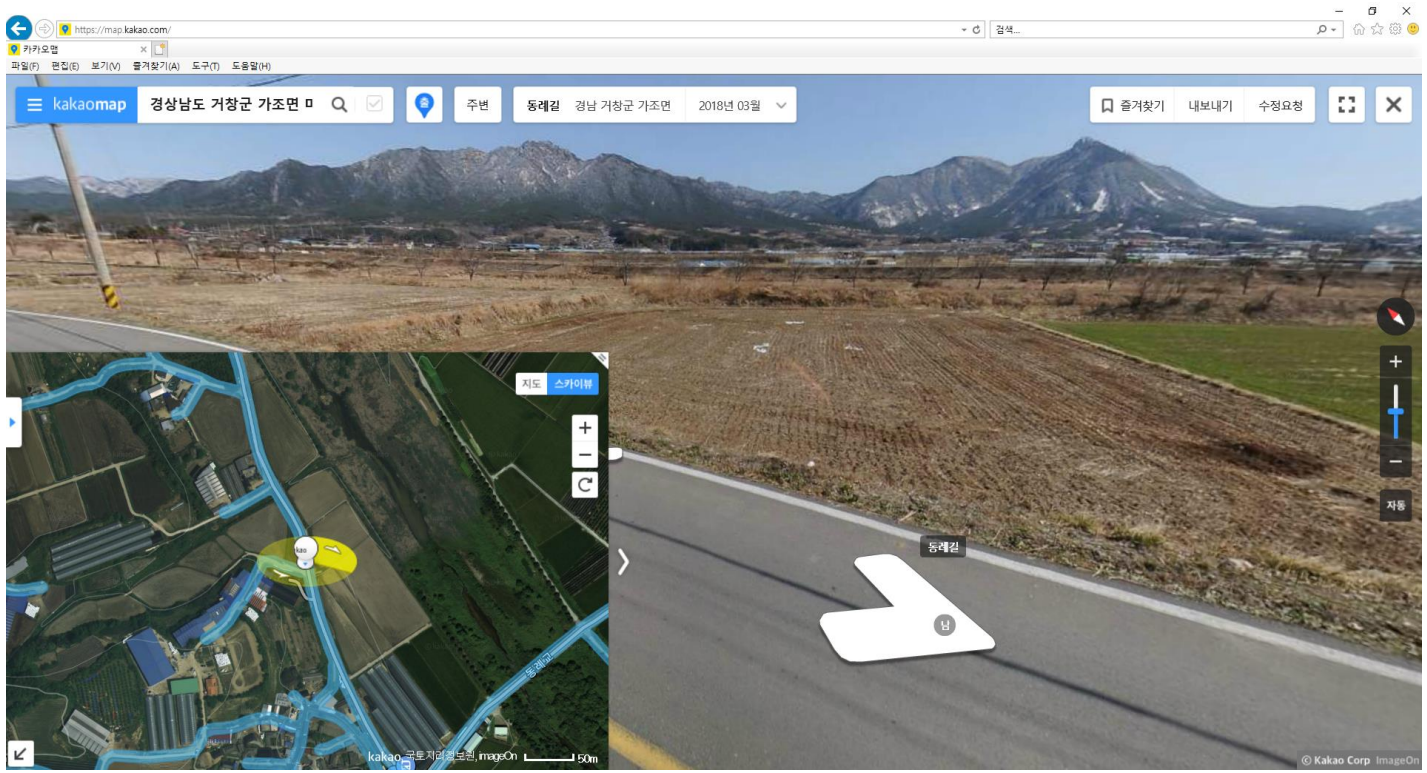
Updating Labeling Data: Google Earth

- Land cover map (paddy area) is updated with on-screen digitizing



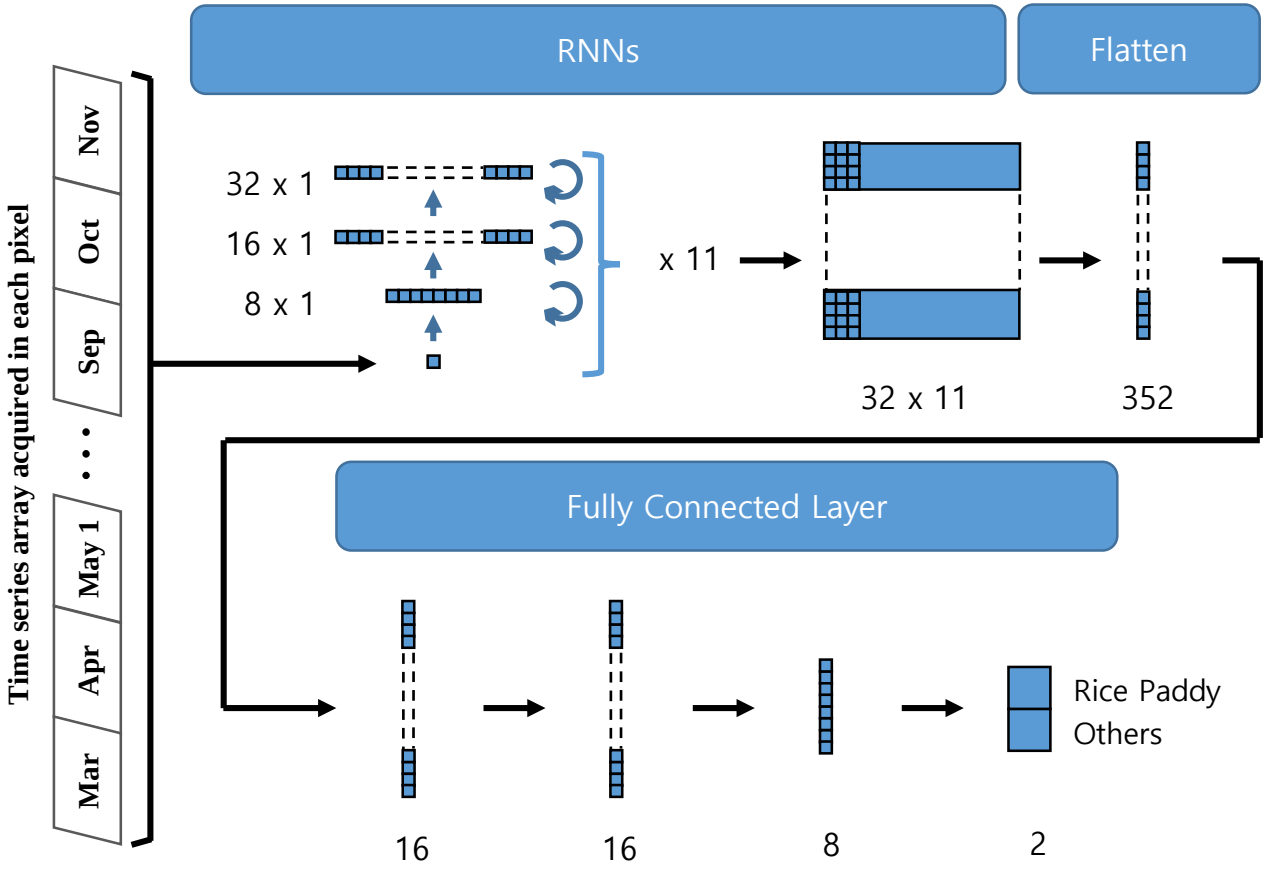
Updating Labeling Data: Domestic street view service

- Uncertain areas are explored with domestic street view services for more accurate update



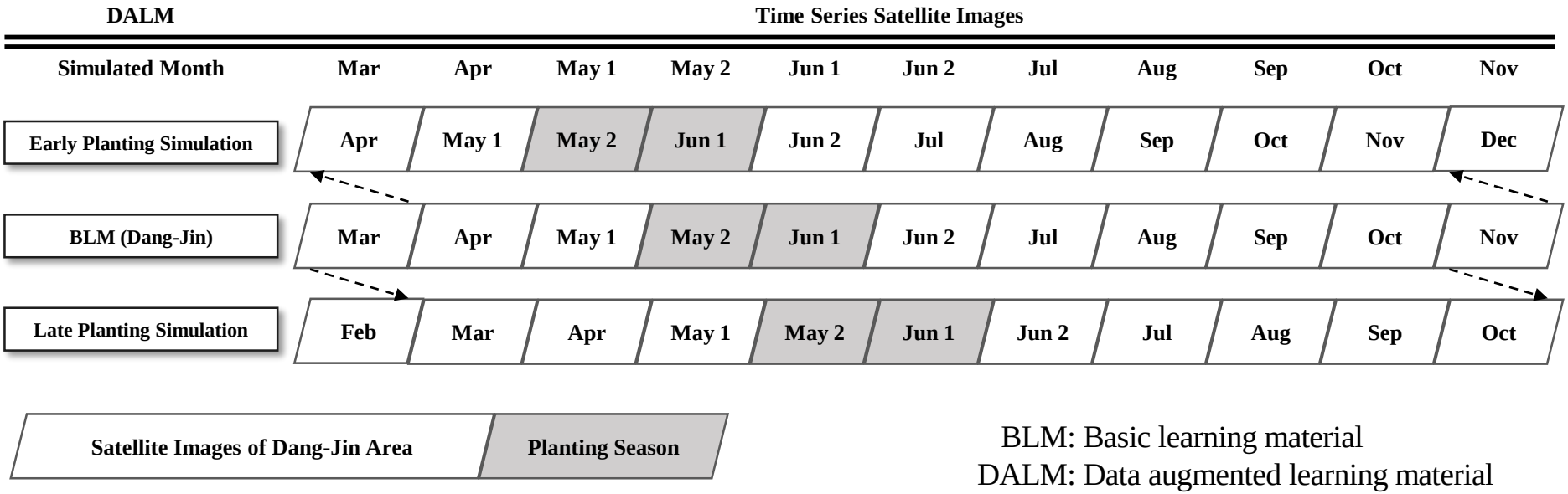
Recurrent Neural Networks (RNNs)

- Deep-learning model that is specialized in time-series analysis



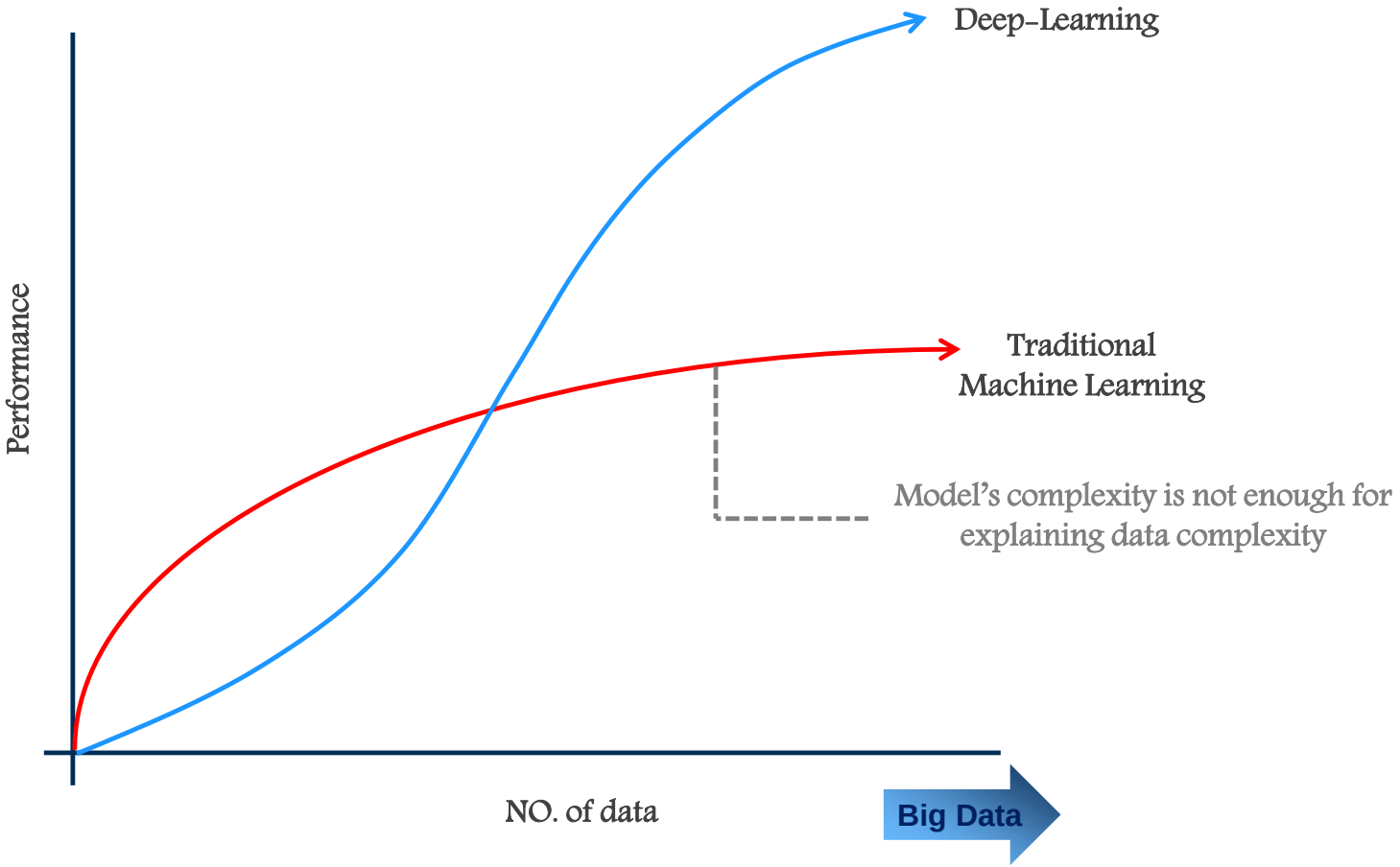
Data Augmentation

- Increase the volume of training data
- Overcome the shortage of labeling data in confined area by simulating rice phenology in other regions
- Move back and forth the time series of Dang-Jin area

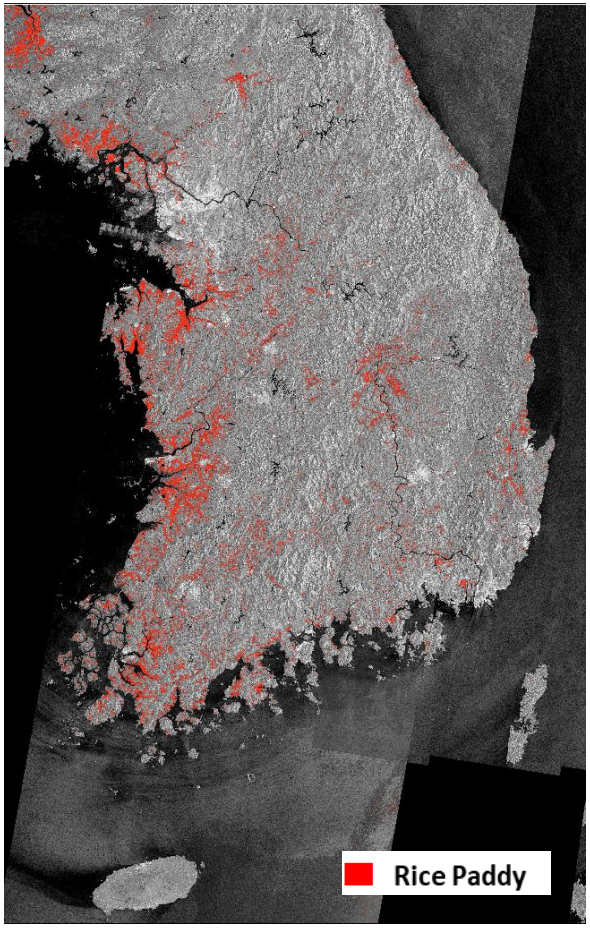


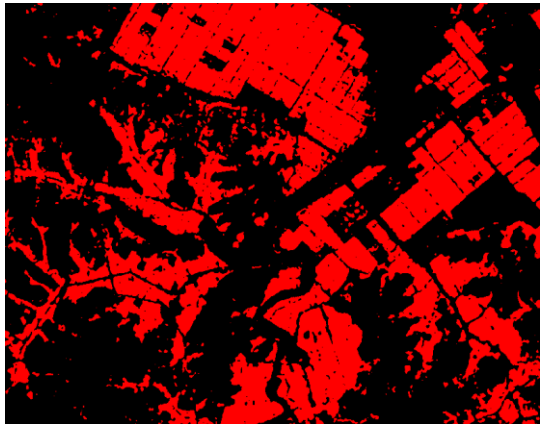
BLM: Basic learning material
DALM: Data augmented learning material

Deep-learning Application

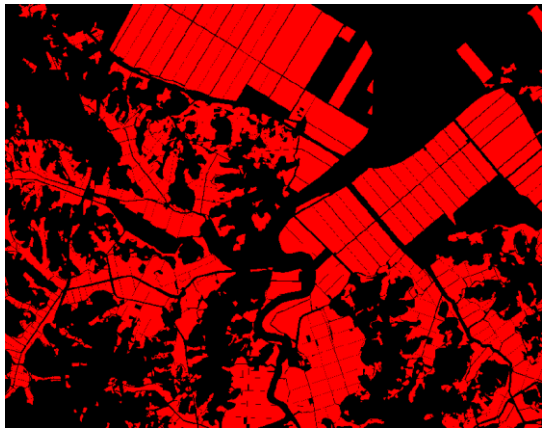


Results: Rice paddy detection in South Korea





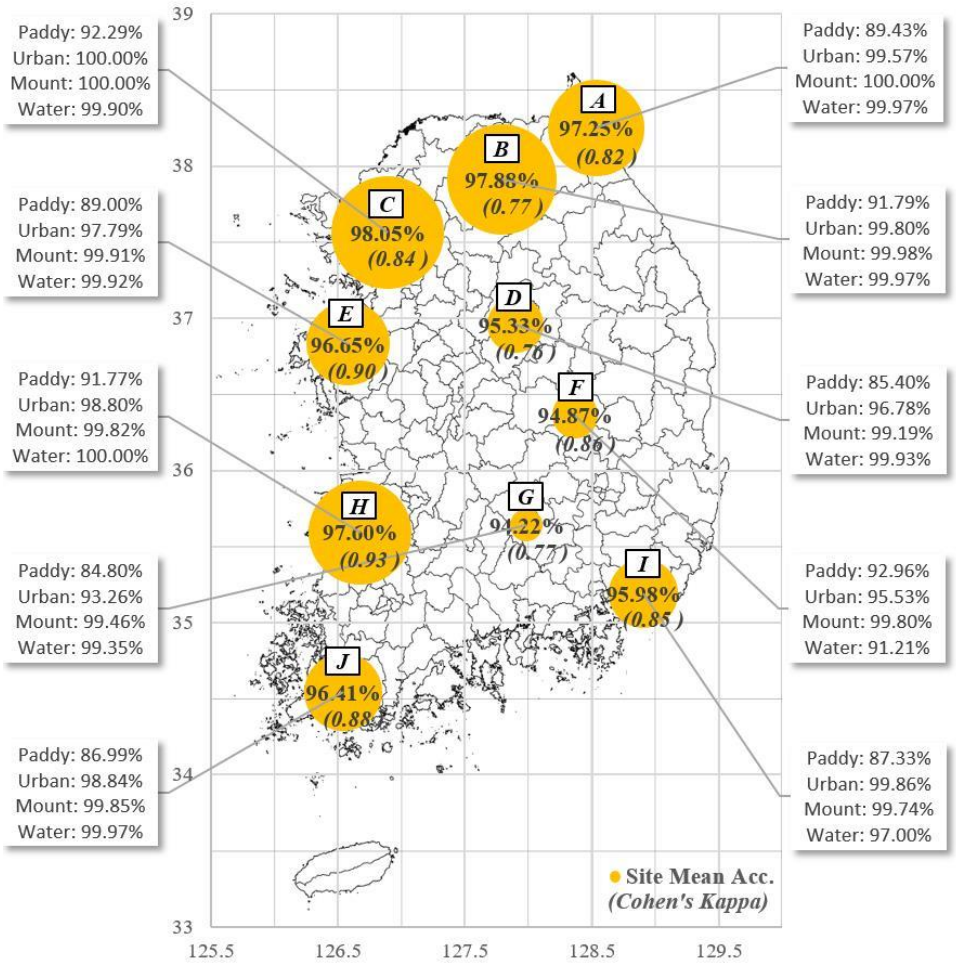
Rice Paddy Detection with RNN



Level-3 Land Cover Map

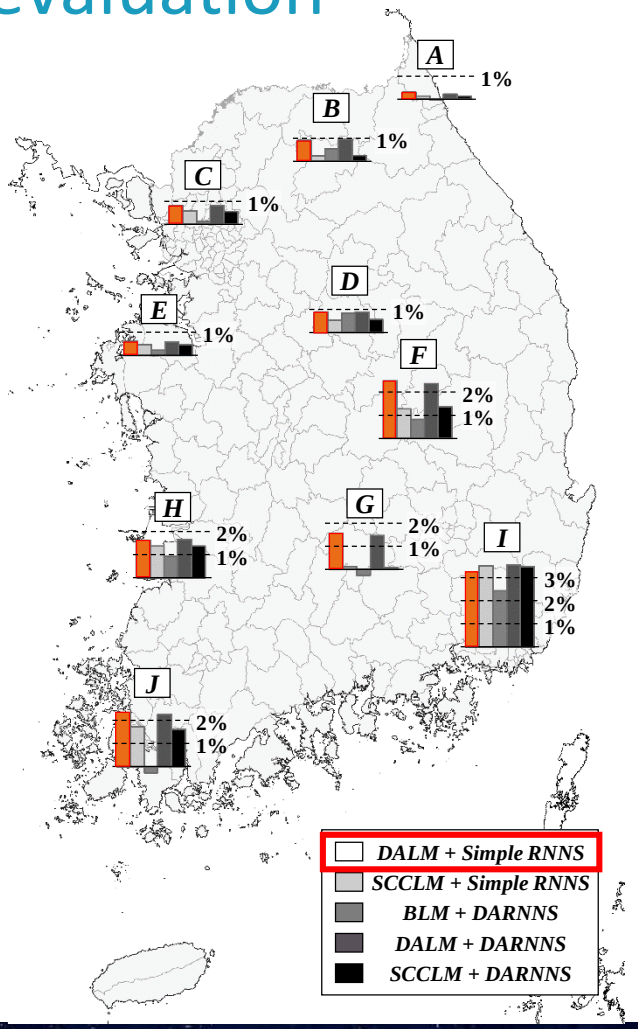
Results: Evaluation on 40 sample plots

- 10 sites were selected from the entire nation
- Each site consists of 4 plots with different landscapes
  : Agricultural, Urban, Mountain, Water
- Mean performance in 40 plots
  : Acc. 96.42%
  : Cohen's Kappa 0.86



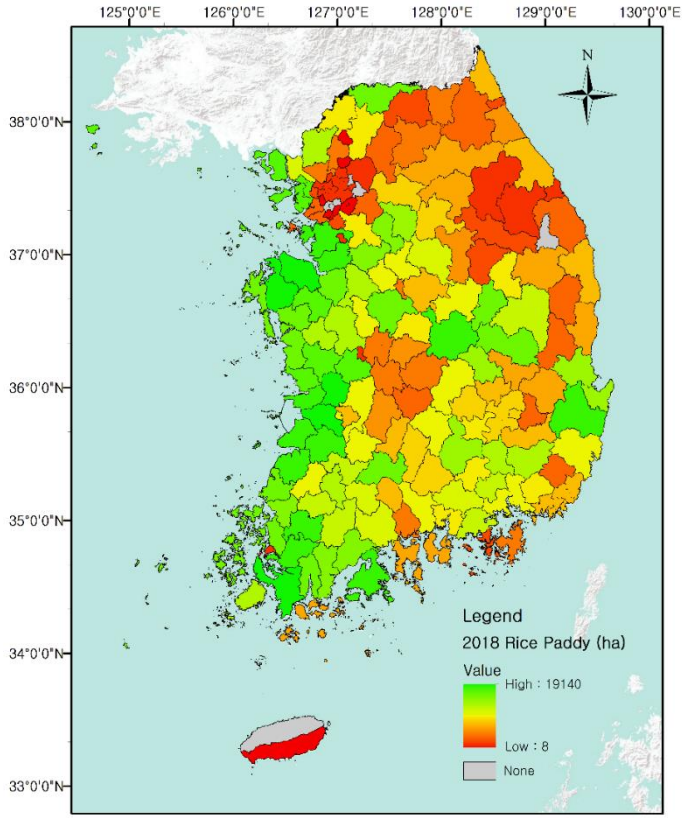
Results: Technique effectiveness evaluation

- Accuracy improvement of data augmentation compared with basic learning material
- Applying data augmentation improved overall accuracy in spite of using confined labeling data in Dang-Jin
- Area far from Dang-Jin showed greater accuracy improvement

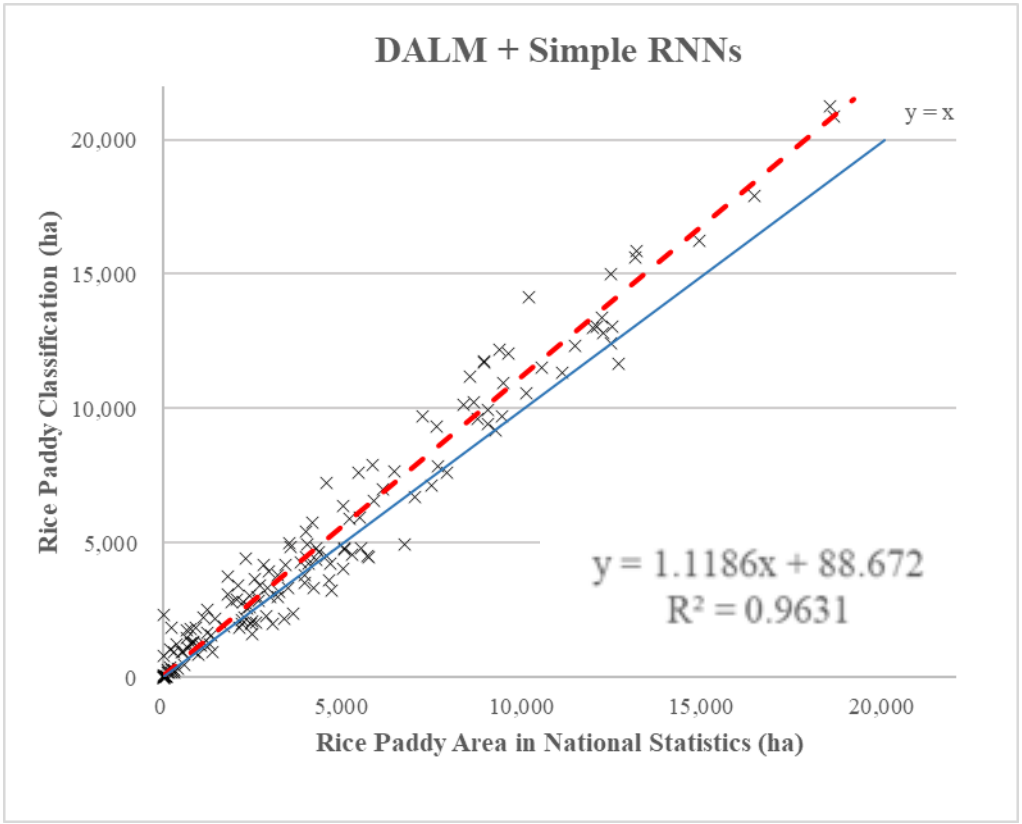


Results: City-level statistic evaluation

- Evaluated in 255 cities



Level-3 Land Cover Map (2015)



EOPEN: opEn interOperable Platform for unified access and analysis of Earth observatioN data

Consortium Members

serco

HLRS

iti Information Technologies Institute

FINNISH METEOROLOGICAL INSTITUTE

Distretto delle Alpi Orientali

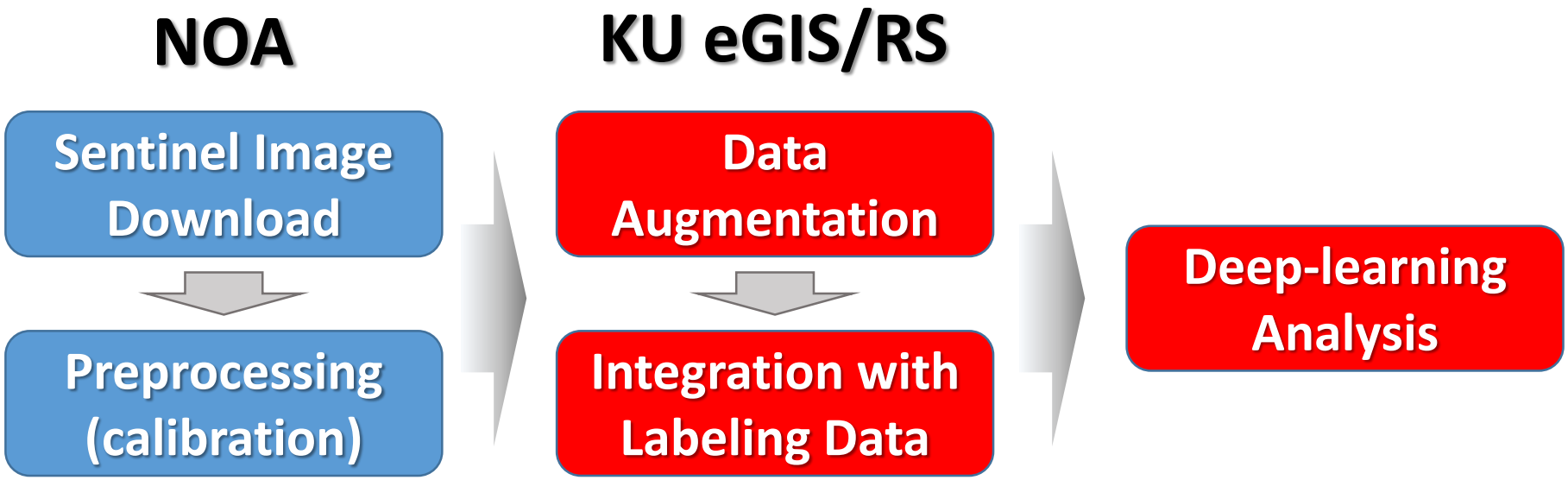
KOREA UNIVERSITY

SUNDOSOFT

spaceapplications SERVICES

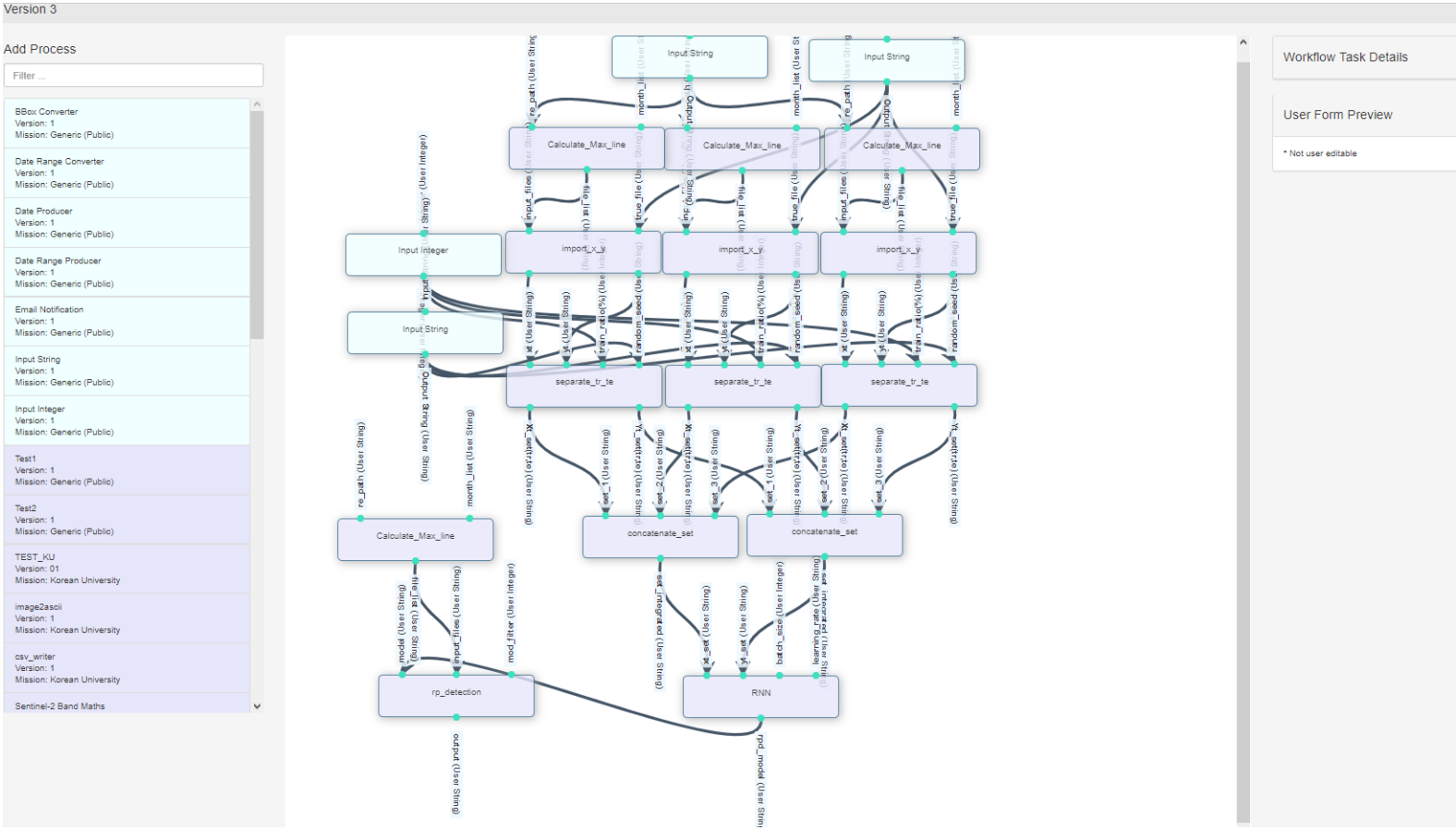
EOPEN Developer Platform

- From data download to analysis under the platform environment
- Perform all processes in this presentation with EOPEN platform



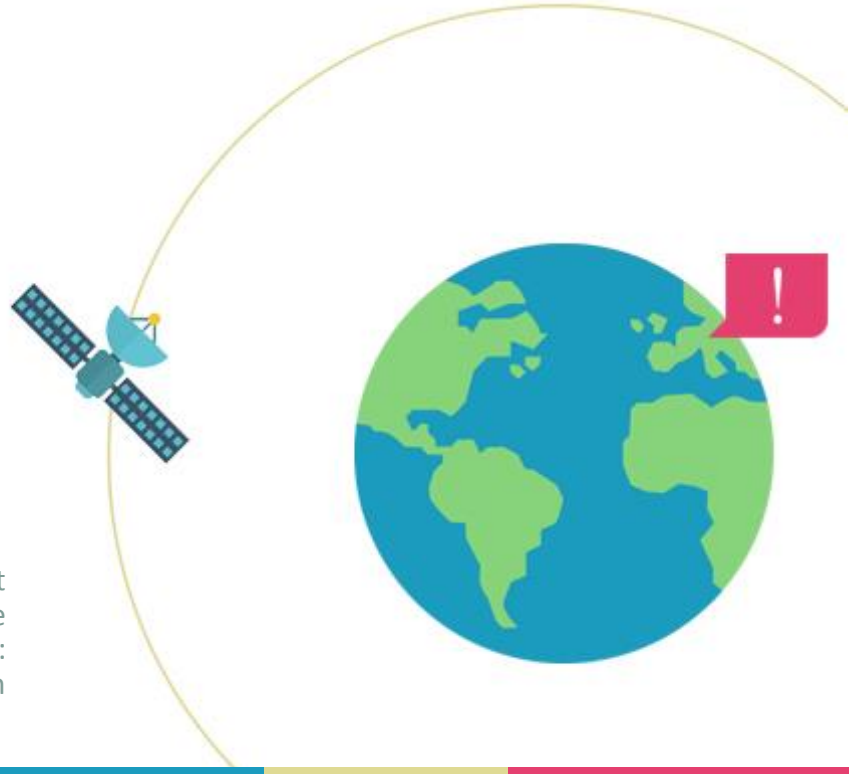
EOPEN Developer Platform

- Deep-learning analysis available



Thanks!

Any questions?



This research was supported by the International Research & Development Program of the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (Grant number: 2018K1A3A7A03089842) and partially supported by the European Commission under contract H2020-776019