EMOBOT IA CHALLENGE

Introduction

Detecting change through multi-image, multi-date remote sensing is essential to developing and understanding of global conditions. This challenge uses features obtained from satellite images using computer vision for further processing using machine learning.

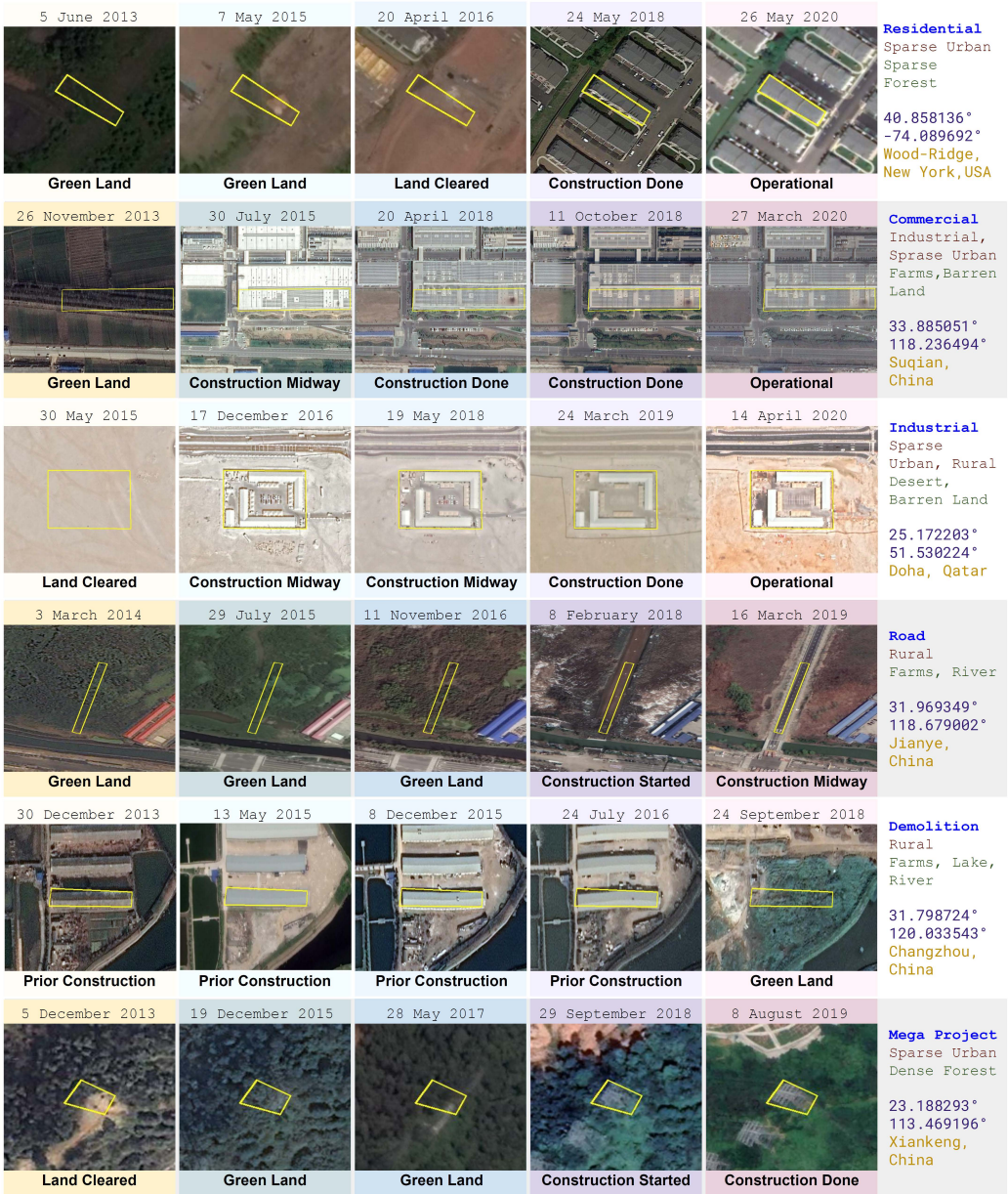

Figure 1. Samples from "test.geojson" showing different **change type**, **change status** on `different dates`, neighborhood label(s), and geography label(s). Latitude-longitude of the change polygon is shown along with city name. First row shows construction of a residential property in suburban area of New York, USA. Second row shows a commercial building in an industrial region which used to be farm lands of a fast growing second tier city in China. Third row shows an industrial construction in desert of Doha, Qatar which went from rural barren desert to a sparse urban area in a time period of 5 years. Fourth row shows construction of a road crossing a river in farm lands of rural China. Fifth row shows special case of urban change, demolition of a farm storage in the fast growing city Changzhou in China. Last row shows construction of a power grid unit which comes under mega project type.

Challenge Description

The goal of this challenge is to classify a given geographical feature into one of six classes. These geographical features are :

1. an irregular polygon
2. categorical values showing what was the status of the polygon on five different dates (e.g. the polygon was under construction on day 1 and construction was completed on next four dates)
3. what are the neighborhood urban features (e.g. the polygon is in a dense urban and industrial region)
4. what are the neighborhood geographic features (e.g. the polygon is near a river and a hill)

The classes to be predicted are
'Demolition': 0
'Road': 1
'Residential': 2
'Commercial': 3
'Industrial': 4
'Mega Projects': 5

Feature Engineering

1. Urban and Geography types are multi valued categorical columns. One hot encoding should help in doing better fature engineering.
2. Irregular polygons can be processed in several ways to create features likel area of polygon, perimeter of polygon, etc. Any type of geometrical property of a polygon can be used.
3. Number of days between two consecutive dates could also be used.

Dataset Description

Files

- **train.geojson** - the training set
- **test.geojson** - the test set
- **skelton_code.py** - starter code

Columns

- `date1` - Date 1 when the polygon was observed (DD-MM-YYYY)
- `date2` - Date 2 when the polygon was observed (DD-MM-YYYY)
- `date3` - Date 3 when the polygon was observed (DD-MM-YYYY)
- `date4` - Date 4 when the polygon was observed (DD-MM-YYYY)
- `date5` - Date 5 when the polygon was observed (DD-MM-YYYY)
- `change_status_date1` - Status of polygon on date1
- `change_status_date2` - Status of polygon on date2
- `change_status_date3` - Status of polygon on date3
- `change_status_date4` - Status of polygon on date4
- `change_status_date5` - Status of polygon on date5
- `urban_types` - comma seperated multiple values showing neighborhood urban types
- `geogprahy_types` - comma seperated multiple values showing neighborhood geographic types
- `geometry` - vector representation of geographic polygons
- `change_type` - label to be classified

Evaluation

The evaluation metric for this competition is Mean F1-Score.
The F1 metric weights recall and precision equally, and a good retrieval algorithm will maximize both precision and recall simultaneously. Thus, moderately good performance on

both will be favored over extremely good performance on one and poor performance on the other.