Exploiting contextual features in superpixels for land cover mapping using high resolution image time series

Phi-Week 2018

Dawa Derksen, Jordi Inglada, Julien Michel

November 12, 2018
Table of Contents

1 Introduction

2 Contextual features

3 Results
Context

- Large-scale land cover mapping
Context

- Each pixel is described by a time series of optical reflectances (R, G, B, IR, etc.) and spectral indices (NDVI, NDWI, etc.)
Introduction

Context

- Each pixel is described by a time series of optical reflectances (R, G, B, IR, etc.) and spectral indices (NDVI, NDWI, etc.)
- Series of Sentinel-2 images
  - 110km × 110km at 10m
  - 13 spectral features / date
  - 33 dates (year : 2016)
  - 489 features
  - Approximately 90Gb
Introduction

Context

- Each pixel is described by a time series of optical reflectances (R, G, B, IR, etc.) and spectral indices (NDVI, NDWI, etc.)
- Series of Sentinel-2 images
  - 110km × 110km at 10m
  - 13 spectral features / date
  - 33 dates (year: 2016)
  - 489 features
  - Approximately 90Gb
- This description is not sufficient to separate all of the desired land cover classes
Introduction

Context

- Continuous urban cover (red)
- Diffuse urban cover (orange)
- Industrial and commercial areas (mauve)
Context

- Continuous urban cover (red)
- Diffuse urban cover (orange)
- Industrial and commercial areas (mauve)
- The difference lies in the context of the pixel
Table of Contents

1 Introduction

2 Contextual features

3 Results
Context

- Popular approach nowadays
  - Deep Convolutional Neural Network (D-CNN)
    - Context implicit in the first layers
    - Has proven to be accurate on many problems
    - Heavy computational load

- Operational context, large data volumes
- Speed and efficiency are essential
- Can the same performance be achieved with alternative methods?
Proposed method

- Using features from superpixel neighborhoods at one or several scales
Superpixels: illustration

SLIC superpixel segmentation

Mean Shift segmentation
Auto-Context

- From an initial pixel-based classification, we can calculate the histogram of the classes in one or several neighborhoods.
Auto-Context

- From an initial pixel-based classification, we can calculate the histogram of the classes in one or several neighborhoods.

- This histogram is used as a feature for generating a new classification.
  - The process can be iterated several times.

- Compact feature (14 vs 330 dimensions)
- Adapted for use multi-scale application
Deep Convolutional Network

- U-net type architecture adapted to time series classification
- Combined with fully connected, pixel resolution MLP
- Weight sharing in the first layers
Table of Contents

1 Introduction

2 Contextual features

3 Results
Results

- Tests on 4 different 110x110 km areas

<table>
<thead>
<tr>
<th></th>
<th>Kappa</th>
<th>OA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>T31TDN</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF (pixel)</td>
<td>89,53 %</td>
<td>91,87 %</td>
</tr>
<tr>
<td>AC</td>
<td>89,42 %</td>
<td>92,19 %</td>
</tr>
<tr>
<td>MLP_Unet</td>
<td>89,82 %</td>
<td>92,49 %</td>
</tr>
<tr>
<td><strong>T30TXQ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF (pixel)</td>
<td>82,87 %</td>
<td>90,61 %</td>
</tr>
<tr>
<td>AC</td>
<td>86,90 %</td>
<td>93,31 %</td>
</tr>
<tr>
<td>MLP_Unet</td>
<td>87,74 %</td>
<td>93,77 %</td>
</tr>
<tr>
<td><strong>T31TGK</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF (pixel)</td>
<td>64,24 %</td>
<td>71,01 %</td>
</tr>
<tr>
<td>AC</td>
<td>66,66 %</td>
<td>73,10 %</td>
</tr>
<tr>
<td>MLP_Unet</td>
<td>67,20 %</td>
<td>73,67 %</td>
</tr>
<tr>
<td><strong>T31UDQ</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RF (pixel)</td>
<td>75,02 %</td>
<td>79,40 %</td>
</tr>
<tr>
<td>AC</td>
<td>84,70 %</td>
<td>88,70 %</td>
</tr>
<tr>
<td>MLP_Unet</td>
<td>86,13 %</td>
<td>89,78 %</td>
</tr>
</tbody>
</table>

**Table:** Computation times

<table>
<thead>
<tr>
<th>Method</th>
<th>Training time/CPU</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF</td>
<td>≈25h</td>
</tr>
<tr>
<td>Auto-Context</td>
<td>≈80h</td>
</tr>
<tr>
<td>MLP-Unet</td>
<td>≈3300h</td>
</tr>
</tbody>
</table>
Results

Image, first date, RGB channels
Results

Pixel based classification (Random Forest)
Results

Auto-Context histograms

- Annual summer crops
- Annual winter crop
- Broad-leaved forest
- Coniferous forest
- Natural grasslands
- Woody moorlands
- Continuous urban fabric
- Discontinuous urban fabric
- Industrial or commercial units
- Road surfaces
- Water bodies
- Intensive grassland
- Orchards
- Vineyards
Results

MLP-Unet

- Annual summer crops
- Annual winter crop
- Broad-leaved forest
- Coniferous forest
- Natural grasslands
- Woody moorlands
- Continuous urban fabric
- Discontinuous urban fabric
- Industrial or commercial units
- Road surfaces
- Water bodies
- Intensive grassland
- Orchards
- Vineyards
Results

Image

Auto-Context

MLP-Unet

RF

- Annual summer crops
- Annual winter crop
- Broad-leaved forest
- Coniferous forest
- Natural grasslands
- Woody moorlands
- Continuous urban fabric
- Discontinuous urban fabric
- Industrial or commercial units
- Road surfaces
- Water bodies
- Intensive grassland
- Orchards
- Vineyards
Conclusion and perspectives

- Deep Convolutional Neural Network
  - Strongest results on urban classes
  - Also deteriorates some other classes
- Superpixel + Auto-Context histograms
  - Similar overall performance to Deep Learning methods
  - Consistent improvement on all classes
  - Lower computational burden
- Can be an alternative to Deep Learning
- Should be validated further
  - Wider range of Sentinel-2 areas
  - Very High Resolution Pleiades images
Thank you for your attention