Research on Urban Construction Land Monitoring Methods Based on Sentinel-1A SAR Data and Fully Convolutional Networks: Postprint

Authors: Lu Wenlu, Liu Zhifeng

Date: 2020-06-12T00:00:00+00:00

Abstract

Sentinel-1A synthetic aperture radar data is not affected by weather conditions such as clouds and fog and possesses rich texture information, providing a new data source for extracting urban construction land information. This paper develops an urban construction land monitoring method based on Sentinel-1A synthetic aperture radar data and fully convolutional networks. The main advantage of this method lies in its ability to effectively composite Sentinel-1A synthetic aperture radar data under different polarization modes and comprehensively integrate multi-scale features. Application in Ganzhou District, Zhangye City, Gansu Province demonstrates that the extraction results of this method achieve an overall accuracy of 92.50% and a Kappa coefficient of 0.85. Compared with the existing method "KTH-Pavia Urban Extractor", the Kappa coefficient improves by 37.10% and the overall accuracy improves by 11.50%. Therefore, this method exhibits good application potential.

Full Text

Urban Construction Land Extraction from Sentinel-1A SAR Data Using a Fully Convolutional Network

Lu Wenlu[1,2], Liu Zhifeng[1,2], He Chunyang[1,2], Xia Pei[1,2]

[1]Center for Human-Environment System Sustainability (CHESS), State Key Laboratory of Earth Surface Processes and Resource Ecology (ESPRE), Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
[2]School of Natural Resources, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China

Abstract: Timely and accurate extraction of urban construction land is essential for assessing the impacts of urban expansion on the environment. Optical remote sensing is susceptible to weather conditions and contains insufficient information on the texture of urban features, and is thus not conducive to urban construction land extraction. Synthetic aperture radar (SAR) is unique in its potential to be used in all weather conditions and at any time. Furthermore, its backscattering and polarimetric information is sensitive to the dielectric and geometric properties of the urban land surface, providing a new means for rapid and accurate extraction of urban construction land. Among existing SAR data, the Sentinel-1A SAR data is widely used due to its free access. Nonetheless, existing methods are poorly adapted to fully utilize the Sentinel-1A SAR data, thereby limiting its application. The fully convolutional network is a deep learning network developed on the basis of the convolutional neural network, which adopts "pixel to pixel" image recognition and has become an effective method for extracting urban construction land. This paper aims to develop a method for detecting urban construction land based on Sentinel-1A SAR data, and thus a fully convolutional network. Firstly, a fully convolutional network consisting of five Inception modules was developed with reference to GoogleNet. Each Inception module includes three convolutional layers with convolutional kernel sizes of 1×1, 3×3, and 5×5 respectively, three corresponding activation layers, and a concatenation layer. The Ganzhou District in Zhangye City, Gansu Province, China, where urban construction land and bare land show similar spectrum features, was used as a case study area to verify our fully convolutional network. The extraction results show that the area of construction land in the Ganzhou District was 4,006 hectares in 2018, accounting for about 20% of the total area. The patch area of urban construction land was generally found to be between 0.01 and 1 hectare. Urban construction land was mainly distributed in the west and northeast, and the circle of 2-4 km from the city center exhibited the most concentrated urban construction land. Our accuracy assessment shows an overall accuracy of 92.50% and a Kappa coefficient of 0.85. By comparing our results to extraction results using the "KTH-Pavia" method (the most widely used method for extracting urban construction land from SAR data), it was found that extraction results based on the fully convolutional network are closer to the real urban construction land in spatial patterns. Furthermore, the overall accuracy and Kappa coefficient were, respectively, 11% and 37% higher than the "KTH-Pavia" method. The principal reason for this higher accuracy is that the fully convolutional network can better integrate multi-polarization and multi-scale texture information from SAR data. Furthermore, the fully convolutional network contains multiple convolutional structures and supports multi-source data inputs. The method developed in this study has both greater accuracy than existing methods and is applicable to urban construction land extraction based on different SAR data in different regions. It therefore has potential for widespread application.

Keywords: Sentinel-1A; Sentinel data; synthetic aperture radar; urban construction land; fully convolutional network; deep learning

1. Introduction

Rapid urbanization has led to significant changes in land use patterns, making timely and accurate monitoring of urban construction land essential for sustainable development planning and environmental impact assessment [1]. Traditional optical remote sensing methods are limited by weather conditions and often lack sufficient texture information for distinguishing urban features, particularly in arid and semi-arid regions where bare soil spectrally resembles urban surfaces [2-3]. Synthetic aperture radar (SAR) offers unique advantages due to its all-weather, day-and-night imaging capability and sensitivity to the dielectric and geometric properties of surface materials [4-8].

Previous studies have utilized various SAR systems for urban mapping, including ERS-1/2, ENVISAT, TerraSAR-X, and Radarsat-2 [18-25]. However, these platforms have limitations in data availability, spatial resolution, or temporal coverage. The European Space Agency's Sentinel-1A satellite, launched in April 2014, provides free, open-access C-band SAR data with a 12-day revisit cycle and high spatial resolution [28-29], making it particularly suitable for operational urban monitoring applications [27].

Despite these advantages, existing methods for urban construction land extraction from Sentinel-1A data remain suboptimal. Conventional approaches struggle to fully exploit the rich multi-polarization and texture information contained in SAR imagery. Deep learning methods, particularly fully convolutional networks (FCNs), have demonstrated remarkable success in semantic segmentation tasks by enabling pixel-to-pixel classification and multi-scale feature integration [38]. This study develops a novel FCN architecture based on Inception modules to leverage the full potential of Sentinel-1A data for urban construction land extraction.

2. Data and Study Area

2.1. Study Area

The Ganzhou District of Zhangye City in Gansu Province, northwestern China, was selected as the study area. This region represents a typical oasis city in an arid environment where urban construction land and bare land exhibit similar spectral characteristics in optical imagery, making it an ideal test site for evaluating SAR-based classification methods.

2.2. Data Sources

SAR Data: Sentinel-1A Interferometric Wide Swath (IW) mode Ground Range Detected (GRD) products were acquired for 2018. The data includes both VV and VH polarizations with a spatial resolution of 10 meters. The imagery was pre-processed using the Sentinel Application Platform (SNAP) for radiometric calibration, spearmint filtering, and terrain correction.

Validation Data: High-resolution Google Earth imagery from Digital Globe (spatial resolution 7.4 m, acquired March 2018) was used as reference data for training and validation. A total of 1,534 sample patches of size 32×32 pixels were manually delineated, with 1,226 used for training and 308 for testing.

3. Methods

3.1. Network Architecture

We developed a fully convolutional network consisting of five Inception modules, inspired by GoogleNet [50]. Each Inception module comprises three parallel convolutional layers with kernel sizes of 1×1, 3×3, and 5×5, followed by corresponding activation layers and a concatenation layer that merges multi-scale features. This architecture enables the network to capture texture information at different scales simultaneously, which is crucial for distinguishing urban structures in SAR imagery.

The network accepts multi-channel input including VV polarization, VH polarization, and their ratio, allowing effective integration of multi-polarization information. The final layer uses a 1×1 convolution to produce pixel-wise classification maps.

3.2. Training and Implementation

The network was implemented using the Caffe deep learning framework [51]. Training was performed using stochastic gradient descent with a base learning rate of 8×10 , batch size of 8, and momentum of 0.99. The learning rate policy was set to "step" with a step size of 5,000 iterations and gamma of 0.99. The SoftmaxWithLoss function was used as the loss criterion:

$$l(y, z) = \log \sum e^{z_j} - z_y$$

where z represents the network output and y denotes the true label.

Data augmentation including rotation and flipping was applied to increase training sample diversity. The model was trained for 100 epochs on an NVIDIA GPU.

4. Results

4.1. Urban Construction Land Extraction

The extraction results for 2018 indicate that urban construction land in Ganzhou District covered approximately 4,006 hectares, representing about 20% of the total district area. The patches were predominantly small to medium-sized, with individual construction land parcels generally ranging from 0.01 to 1 hectare in area.

Spatially, urban construction land was primarily distributed in the western and northeastern parts of the district. The zone located 2-4 km from the city center showed the highest concentration of construction activity, reflecting recent urban expansion patterns.

4.2. Accuracy Assessment

Accuracy was assessed using the independent test set of 308 samples. The fully convolutional network achieved an overall accuracy of 92.50% and a Kappa coefficient of 0.85 . Producer' s accuracy for urban construction land reached 89.72%, while user' s accuracy was 96.00%, indicating both high completeness and correctness in extraction.

TABLE:1 Accuracy assessment of urban construction land extraction results based on fully convolutional network

Class	Producer' s Accuracy	User' s Accuracy
Urban Construction Land	89.72% (10.28%)	96.00% (4.00%)
Non-Urban Land	95.70% (4.30%)	89.00% (11.00%)
Overall Accuracy	**92.50%**	
Kappa Coefficient	**0.85**	

4.3. Comparison with Existing Methods

We compared our results with those obtained using the KTH-Pavia method [36-37], a widely adopted approach for SAR-based urban extraction. The visual comparison [Figure 6: see original paper] reveals that the FCN-based results exhibit greater spatial detail and better capture the fragmented patterns of urban construction land. Quantitatively, the FCN method outperformed KTH-Pavia by 11% in overall accuracy and 37% in Kappa coefficient (0.85 vs. 0.62).

The KTH-Pavia method tended to overestimate urban areas in regions with high backscatter from natural surfaces, while underestimating small, isolated construction patches. In contrast, the FCN approach, through its multi-scale feature learning capability, successfully distinguished subtle textural differences between urban structures and natural features.

5. Discussion

The superior performance of the fully convolutional network can be attributed to several key factors. First, the Inception-based architecture effectively integrates multi-polarization information from VV and VH channels, capturing both surface roughness and volumetric scattering characteristics of urban materials. Second, the multi-scale convolution kernels (1×1, 3×3, 5×5) enable simultaneous extraction of fine-grained details and contextual neighborhood information, which is essential for resolving the heterogeneous nature of urban landscapes.

Furthermore, unlike traditional pixel-based methods that rely on handcrafted features, the deep learning approach automatically learns hierarchical representations optimized for the specific task. This end-to-end learning paradigm reduces subjective parameter tuning and improves generalization capability across different urban environments.

The method demonstrates particular advantages in arid region cities where spectral confusion between construction land and bare ground is severe. By leveraging SAR' s sensitivity to geometric structure rather than spectral reflectance, the FCN can reliably identify urban features independent of seasonal vegetation variations.

6. Conclusion

This study developed a fully convolutional network method for extracting urban construction land from Sentinel-1A SAR data. The approach achieved high accuracy (92.50% overall, Kappa 0.85) in a challenging arid environment, outperforming the conventional KTH-Pavia method by significant margins. The network' s ability to integrate multi-polarization and multi-scale texture information makes it well-suited for operational urban monitoring applications.

The method' s applicability to free, open-access Sentinel-1A data enhances its potential for large-scale, routine monitoring of urban expansion in various geographical settings. Future work will focus on extending the approach to multi-temporal analysis and integrating additional data sources such as Sentinel-2 optical imagery for improved classification robustness.

References

[1] Reference content not fully recoverable from corrupted text. [2-3] References to optical remote sensing limitations. [4-8] References to SAR applications in urban monitoring. [9-10] References to SAR texture analysis. [15] Reference to building extraction methods. [16-17] References to ChinaXiv collaborative journal. [18-19] Radarsat-2 applications. [20-21] ENVISAT applications. [22-23] ERS-1/2 applications. [24-25] TerraSAR-X applications. [26] Reference to SAR data characteristics. [27] Sentinel-1A mission overview. [28-29] Sentinel-1A technical specifications. [36-37] KTH-Pavia method references. [38] Long et al. Fully convolutional networks for semantic segmentation. [40] LeCun et al. Gradient-based learning applied to document recognition. [41] Ouyang et al. Sentinel-1A products' characteristics. [42] Pan et al. Residential area extraction in GF-1 images. [43] He et al. Detecting global urban expansion. [44] Fu et al. Classification for high-resolution remote sensing imagery. [45] Zhang et al. Urban land use classification. [46] Wang et al. Urban expansion in oasis cities. [47] Liang et al. Evolution characteristics of urban functional land. [48] Feng et al. Urban construction land growth guidance. [49] Reference to urban functional land evolution. [50] Szegedy et al. Going deeper with convolutions. [51] Jia et al. Caffe: Convolutional architecture for fast feature embedding. [52]

Janssen et al. Accuracy assessment of satellite-derived land cover data. [53] Kokisaitoh. Deep learning from scratch.

Note: Figure translations are in progress. See original paper for figures.