Postprint: A Study on Sandy Land Information Extraction Methods Based on CART Decision Tree

Authors: Zhang Xiwei, Wang Lei, Wang Xiyuan, Wang Xiyuan

Date: 2019-09-11T00:00:00+00:00

Abstract

To investigate methods for sand information extraction, an object-oriented approach based on CART decision trees was employed to extract sand information in the Shapotou District of Zhongwei City. Initially, multi-scale segmentation and spectral difference segmentation were performed on the study area to obtain an object layer, followed by the selection of appropriate extraction features and training sample points. Finally, the selected extraction features and sample points were input to generate a CART rule tree, which was then used to classify land features and extract sand information. The results demonstrate that the object-oriented CART decision tree method for sand information extraction exhibits a high degree of automation and accuracy; the overall classification accuracy of the constructed CART decision tree can reach 77%, which is 1.12 times that of the nearest neighbor classification and 1.57 times that of the support vector machine classification. Additionally, NDBI (Normalized Difference Bareness Index), GSI (Grain Size Index), and the mean value of SWIR 2 (Band 7) can effectively distinguish three easily confused land features—sand, gobi, and bare rock/gravel land—and constitute three important feature indices in the sand extraction process.

Full Text

Preamble

DOI: 10.12118/j.issn.1000-6060.2019.05.19
Journal: ARID LAND GEOGRAPHY

Authors:
ZHANG Xi-wei[1], WANG Lei[2], WANG Xi-yuan[1,3]
[1] School of Physics and Electrical-Electronic Engineering, Ningxia University, Yinchuan 750021, China

[2] Breeding Base for State Key Laboratory of Land Degradation and Ecological Restoration in Northwest China, Yinchuan 750021, China
[3] Ningxia Hui Autonomous Region Desert Information Intelligence Key Laboratory, Ningxia University, Yinchuan 750021, China

Abstract: This paper employed an object-oriented method combined with the CART decision tree algorithm to extract sand information with a high degree of automation and comprehensive feature extraction. The main research process comprises: (1) selecting the study area and preprocessing the corresponding imagery; (2) utilizing multi-scale segmentation and spectral difference segmentation to obtain the object layer; (3) selecting rich extraction features and training sample objects; (4) training features and sample objects to generate the CART rule tree; (5) applying all objects to the rule tree to obtain classification results; and (6) comparing these results with Nearest Neighbor and Support Vector Machine classification outcomes. Compared with existing CART decision tree research on sand information extraction, the overall classification accuracy reached 77%, which is 1.12 times that of the Nearest Neighbor classification result and 1.57 times that of the Support Vector Machine classification result. Additionally, the Normalized Difference Bare Index (NDBI), Granularity Size Index (GSI), and the seventh band (SWIR2) successfully distinguished three easily confused objects: sand, Gobi, and bare rock, which represent three important characteristic indices in the sand extraction process. The experiment demonstrates that this method is a feasible approach for actual desertification monitoring.

Keywords: object-oriented; multi-scale; spectral differences; CART decision tree; sand extraction

2. Study Area and Data

The study area is Shapotou District, located between 104°17 ∼106°10 E and 36°06 ∼37°50 N, covering an area of 5922.4 km^2. The region features typical arid and semi-arid landscape characteristics, making it representative for desertification monitoring research.

2.1 Data Preprocessing

The primary data source was a Landsat-8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) image acquired on August 9, 2016 (Path/Row: 130/034, Scene ID: LC81300342016222LGN00). The Landsat-8 satellite, launched by NASA in 2013, carries two key sensors: OLI with nine spectral bands at 30 m spatial resolution (15 m for panchromatic) and TIRS with two thermal bands. The image underwent systematic radiometric calibration and atmospheric correction using ENVI software' s FLAASH module to ensure data quality and spectral fidelity.

[Figure 1: see original paper] Preprocessing data of Shapotou District

2.2 CART Decision Tree Methodology

The Classification and Regression Tree (CART) algorithm is a binary recursive partitioning method that generates a rule-based decision tree through training samples. The object-oriented approach integrates spectral, shape, and contextual features for classification.

The feature extraction process employed multiple indices including the Normalized Difference Bare Index (NDBI) and Granularity Size Index (GSI). The CART decision tree was constructed using training samples representing key land cover types: sand, Gobi, bare rock, vegetation, water, and impervious surfaces. The algorithm automatically selected optimal splitting variables and thresholds to maximize class separability.

The segmentation process utilized multi-scale and spectral difference algorithms to generate homogeneous objects. The scale parameter was optimized to ensure objects captured meaningful landscape features while preserving spectral heterogeneity. The rule tree was then applied to all image objects to produce the final classification map.

[Figure 2: see original paper] CART decision tree classification results

The overall classification accuracy achieved 77%, substantially outperforming both Nearest Neighbor (1.12× higher) and Support Vector Machine (1.57× higher) methods. The NDBI, GSI, and SWIR2 band proved particularly effective for discriminating spectrally similar surfaces (sand, Gobi, and bare rock), demonstrating the method' s robustness for operational desertification monitoring applications.

References

[1] HUANG Huiping. Scale issues in object-oriented image analysis [D]. Beijing: Institute of Remote Sensing Applications, Chinese Academy of Sciences, 2003.

[2] LI Xiaoqin, ZHANG Zhende, ZHANG Peimin. Remote sensing survey and monitoring of sandy land information based on GIS [J]. Arid Environmental Monitoring, 1998, (4): 213-217, 252.

[3] QIAO Pinglin, ZHANG Jixian, LIN Zongjian. An artificial neural network method for desertification information extraction [J]. Acta Geodaetica et Cartographica Sinica, 2004, 33(1): 58-62.

[4] ZHANG Xiaojuan, YANG Yingjian, GAI Liya, et al. Research on vegetation classification method based on CART decision tree algorithm and maximum likelihood ratio [J]. Remote Sensing Information, 2010, (2): 88-92.

[5] QI Le, YUE Cairong. Remote sensing image classification based on CART decision tree method [J]. Forest Inventory and Planning, 2011, 36(2): 62-66.

[6] SU Xiaoyu, GAN Fuping, WAN Lifei, et al. The object-oriented classification of high resolution images with spectrum analysis [J]. Journal of Graphics, 2012, 3(1): 73-79.

[12] QI Le. Remote sensing image classification based on CART decision tree method [D]. 7t: 7t&+# 2011 [$WENSi\ Remotesensingim\ nitoringofdesertificationinGolmudarea[J\ \ Remot$ 228)(cid : 143) wi(cid : 246) (cid : 143)(cid : 247) G»(cid : 229)(cid : 230)(cid : 231)[23(cid : 243)$N5OP$(cid : 190)$nology$ 2011 (cid : 150)(cid : 192)‵(cid : 160)¡[J ¤, 2004, 33(1): 58-62.

[13] WANG Zhibo, GAO Zhihai, WANG Bengyu, et al. The study of extracting sandy land information from remote sensing image based on object-oriented method [J]. Remote Sensing Technology and Application, 2012, 27(5): 770-777.

[14] XU Hanqing, et al. Object-oriented classification of ZY-3 satellite imagery [J]. Journal of Nanjing Normal University (Natural Science), 2006, 29(3): 118-121.

[15] Rouse JW, Haas RH, Schell JA, et al. Monitoring vegetation systems in the Great Plains with ERTS [J]. NASA Special Publication, 1974, 351: 309.

[20] OUYANG P, et al. Sand information extraction method based on CART decision tree [J]. Journal of Desert Research, 2012, 32(1): 73-79.

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