International Journal of Applied Earth Observation and Geoinformation 49 (2016) 87–98

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

A systematic comparison of different object-based classification techniques using high spatial resolution imagery in agricultural environments

Manchun Li[a], Lei Ma[a,b,*], Thomas Blaschke[b], Liang Cheng[a], Dirk Tiede[b]

[a] Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, 210023 Nanjing, China
[b] Department of Geoinformatics—Z_GIS, University of Salzburg, Hellbrunner Str. 34, A-5020 Salzburg, Austria

ARTICLE INFO

Article history:
Received 27 November 2015
Received in revised form 31 January 2016
Accepted 31 January 2016
Available online 11 February 2016

Keywords:
GEOBIA
OBIA
Random Forest
Segmentation scale
Training set size
Feature selection
Mixed object
Classification
High spatial resolution

ABSTRACT

Geographic Object-Based Image Analysis (GEOBIA) is becoming more prevalent in remote sensing classification, especially for high-resolution imagery. Many supervised classification approaches are applied to objects rather than pixels, and several studies have been conducted to evaluate the performance of such supervised classification techniques in GEOBIA. However, these studies did not systematically investigate all relevant factors affecting the classification (segmentation scale, training set size, feature selection and mixed objects). In this study, statistical methods and visual inspection were used to compare these factors systematically in two agricultural case studies in China. The results indicate that Random Forest (RF) and Support Vector Machines (SVM) are highly suitable for GEOBIA classifications in agricultural areas and confirm the expected general tendency, namely that the overall accuracies decline with increasing segmentation scale. All other investigated methods except for RF and SVM are more prone to obtain a lower accuracy due to the broken objects at fine scales. In contrast to some previous studies, the RF classifiers yielded the best results and the k-nearest neighbor classifier were the worst results, in most cases. Likewise, the RF and Decision Tree classifiers are the most robust with or without feature selection. The results of training sample analyses indicated that the RF and adaboost. M1 possess a superior generalization capability, except when dealing with small training sample sizes. Furthermore, the classification accuracies were directly related to the homogeneity/heterogeneity of the segmented objects for all classifiers. Finally, it was suggested that RF should be considered in most cases for agricultural mapping.

1. Introduction

Recent advances in the airborne and spaceborne remote sensing technology and image segmentation techniques offer new opportunities for remote sensing agricultural mapping (Wulder and Coops, 2014; Ma et al., 2014; Zhang et al., 2015), whilst the OBIA/GEOBIA ((Geographic) Object-based Image Analysis) paradigm in the field of remote sensing classification is already widely accepted (Liu et al., 2006; Blaschke et al., 2014). Plenty of classification approaches are documented within the GEOBIA framework, especially the implementation of expert rule-sets, which make use of the extremely extended feature space spanned by the use of the available object-specific features at several segmentation scales (context features/neighborhood relation, scaled-hierarchy relations, form features etc.) (Benz et al., 2004; Blaschke, 2010; Tiede et al., 2010; Strasser and Lang, 2015). Nevertheless, supervised classification algorithms based on objects rather than pixels as classification units are still very important. According to previous studies, the comparison of classification approaches within a GEOBIA framework can be divided into two general topics: 1) a comparison of GEOBIA and traditional per-pixel image analysis; and 2) a comparison of different classification techniques within GEOBIA only. Although there is general agreement regarding the former (Yan et al., 2006; Duro et al., 2012), the selection of a suitable classifier is still a problem for any per-pixel and GEOBIA method due to the diversity of data sources, the training set size and feature and, especially, the selection of segmentation parameters (e.g. scale/size of objects) and spectrally mixed objects bringing some uncertainty into the comparison of methods (Yu et al., 2008). In the following sub-section we

* Corresponding author at: Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, 210023 Nanjing, China.
E-mail address: maleinju@gmail.com (L. Ma).

http://dx.doi.org/10.1016/j.jag.2016.01.011
0303-2434/

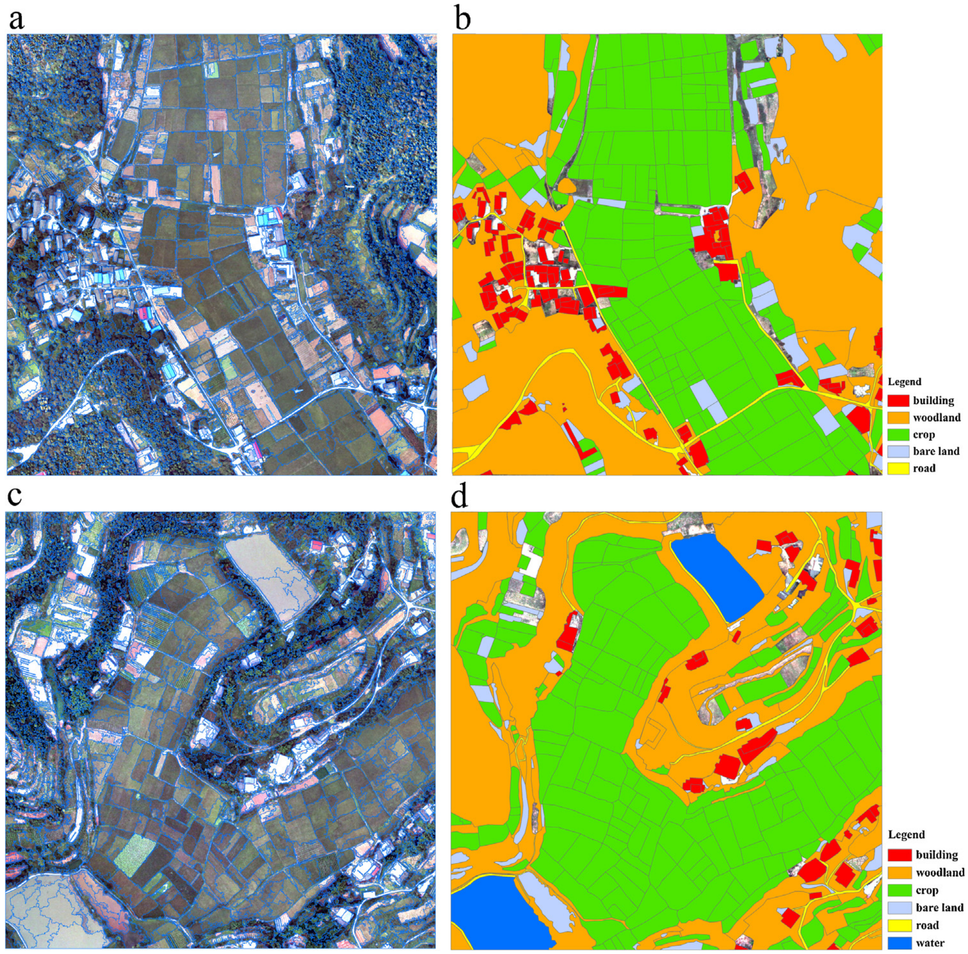

Fig. 1. Study sites. The segmented layers and the corresponding reference layers. (a) The segmented layer for area 1 at a scale of 100; (b) the manual interpretation layer for area 1; (c) the segmented layer for area 2 at a scale of 100; and (d) the manual interpretation layer for area 2.

conduct a brief literature review and succinctly identify the benefits and the main difficulties when comparing classification techniques.

Despite the expert rule-set classifications in GEOBIA, researchers already try using statistical and machine-learning classification techniques, including Linear Discriminant Analysis (LDA) (Pu and Landry, 2012), Random Forest (RF) (Stumpf and Kerle, 2011), Decision Tree (DT) (Mallinis et al., 2008), K-Nearest Neighbor (KNN) (Luque et al., 2013), naiveBayes (Dronova et al., 2012), Support Vector Machines (SVM) (Heumann, 2011). Since around 2010, the Adaboost technique, as another ensemble classifier, has received more attention in remote sensing classification due to the high accuracy (Chan and Paelinckx, 2008). So far, only a few GEOBIA applications have used ensemble classifications beyond RF. Thus, Adaboost as another example of ensemble classifier was used with GEOBIA other than RF.

In per-pixel analysis, the classification accuracy is usually accredited to the classification technique (Rogan et al., 2008). Chan and Paelinckx (2008) also evaluated Random Forest and Adaboost tree-based ensemble classifications using airborne hyperspectral imagery and yielded almost the same accuracy results as established per-pixel classifiers. Brenning (2009) compared eleven classification algorithms in automatic rock glacier detection using terrain analysis and multispectral remote sensing, and found that mapping results of PLDA (Penalized Linear Discriminant Analysis) are significantly better than all other classifiers, including both SVM and RF. For the purpose of land-cover classification, Shao and Lunetta (2012) compared the support vector machine, neural network, and CART (classification and Regression Trees) algorithms using Moderate Resolution Imaging Spectroradiometer time-series data, and found that SVM was the superior algorithm. Xu et al. (2014) compared seven classification techniques for marine oil spill identification using RADARSAT-1 imagery and showed that the classification was able to benefit from ensemble techniques (bundling and bagging). These few examples demonstrate the importance of selecting optimal classifiers for remote sensing classification or prediction.

We hypothesize that for GEOBIA it is not sufficient to analyze the choice of the classifier only, because the resulting accuracies also depend on the segmentation scale, on the selection of features, and on the existence of spectrally mixed objects (Ma et al., 2015). It therefore seems to be impossible to generically advise on the selection of a specific classification technique for a specific application case. For instance, Laliberte et al. (2006) and Mallinis et al. (2008) found that the overall classification accuracies of the classification tree was better than that of the K-NN algorithm. In contrast, Tehrany et al. (2014) suggested that K-NN generally performed better for land-cover mapping, while they compared with DT and SVM using SPOT 5 imagery. In addition, previous researches (Duro et al., 2012; Ghosh and Joshi, 2014) demonstrated a superior capability of producing higher classification accuracies by SVM or RF, but Dronova et al. (2012) concluded that RF always performed worse, while they examined six families of statistical machine-learning classifiers. We assumed that these inconsistent results are related to the unsystematic studies of comparison, while, as mentioned before, the vast majority of comparisons analyzed only a single factor (i.e., scale) or relatively few classification algorithms.