Road Detection in Urban Areas Using Random Forest Tree-Based Ensemble Classification

Safaa M. Bedawi[1,2(✉)] and Mohamed S. Kamel[2]

[1] National Authority for Remote Sensing and Space Sciences, Cairo, Egypt
[2] Center of Pattern Analysis and Machine Intelligence, Waterloo, ON, Canada
{sbedawi,mkamel}@pami.uwaterloo.ca

Abstract. The rapid growth in using remote sensing data highlights the need to have computationally efficient geospatial analysis available in order to semantically interpret and rapidly update current geospatial databases. Object identification and extraction in urban areas is a challenging problem and it becomes even more so when very high-resolution data, such as aerial images, are used. In this paper, we use Random Forest Classifier tree based ensemble to enhance the extracting accuracy for roads from very dense urban areas from aerial images. Both the spatial and the spectral features of the data are used for pre-classification and classification. Comparisons are made between the RF ensemble and other ensembles of statistic classifiers and neural networks.

The proposed method is tested to aerial and satellite imagery of an urban area. The result shows that the RF ensemble enhances the overall classification accuracy for roads by 8 %. Also, it demonstrates that the approach is viable for large datasets due to its faster computational time performance in comparison to other ensembles.

Keywords: Random forest classifier · Ensemble of classifiers · Remote sensing · Very high resolution · Aerial images · Road extraction

1 Introduction

Objects extracted from very high resolution Remote Sensing (RS) imagery [1] have numerous applications in urban planning, forest monitoring, disaster management, and climate modeling. Urban land-cover/land-use maps are still generated by human experts, which makes the process both expensive and time consuming. Human experts tend to favor higher spatial resolution to higher spectral ones as higher spatial resolution increases the visibility of terrestrial features. This is the case especially with urban objects through reducing per-pixel spectral heterogeneity and thereby improving land cover identification. This explains why aerial imagery has traditionally been the primary source used for urban planning. Recent developments in sensor technology demonstrate a shift from aerial imagery to satellite based images for urban applications, as a new high spatial resolution multispectral satellite has recently been launched (e.g., GeoEye and WorldView). However, increase in resolution has also lead to augmentation of manual costs. This has also lowered accuracy, particularly in urban image classification, as urban areas are dense objects that become visible with the use of very high resolution. This visibility leads to displaying complex urban features [2], which

may not be the case for other non-man made land covers and land uses such as forests, wetland, desert landscape, and agriculture.

Various classifiers have been used in extracting land-cover/land-use from RS imagery. Typical methods include multivariate regression models, spectral mixture models, machine learning models and integration with geographical information systems [3] among others. It is desirable to use spectral-spatial data in order to extract as much information as possible concerning the area being classified. The superiority of one technique over the others cannot be claimed [4]. In contrast to standard classifiers, which are based solely on the decision of a single classifier, the ensemble approach combines several different classifier outputs. In doing so the overall accuracy usually increases. Random Forest classifiers (RF) are one example of such a classifier system [5]. Ensembles of Multiple Classifiers/Multiple Classifier Systems have proved to be the most remarkable applications for over two decades in RS applications [6–10, 12].

In this paper, the RF Tree Based ensemble is used for the classification of urban data when using aerial images. Motivated by its relatively low computation requirement, robustness to outliers and because of reported good results with other RS in literature, we choose the RF Tree Based Ensemble. To the best of our knowledge, few researchers have exploited the use of RF in very high-resolution aerial images for dense urban areas [10, 11], especially when there is no height information available. In our experiment we use both the spatial and spectral features when performing classification. We compare the performance of RF ensemble with three types of ensembles of neural network and three ensemble based ones on statistical classifiers.

The paper is organized as follows. Section 2 briefly introduces the Random Forest Classifier while Sect. 3 describes ensembles of multiple classifiers. In Sect. 4, we present the results and finally, our conclusion is drawn in Sect. 5.

2   Random Forests (RF)

Random Forest [13] is a tree-based ensemble machine- learning technique that is increasingly used in RS image classification. A Random Forest Classifier consists of a number of decision trees whose predictions are typically combined using majority voting. The goal of the training procedure is to reduce the variance of the ensemble by attempting to produce de-correlated trees. This is achieved by learning each tree on a random subset of the dataset and by using a random subset of the input variables. We selected each trained sample from the original training sample by the bootstrapped method.

Gini Index is used as a based for construction of RF classifier. This targets locating the biggest homogeneous subclass within the training set to differentiate the rest of the train sample [14].

We can reduce the computational complexity and reduce the correlation between trees by limiting the number used in split. This makes it possible for RF to handle the complexities found in very high resolution RS imagery for urban areas.

3 Ensemble of Multiple Classifiers

The concept of ensemble of multiple classifiers can be described concisely as: The final classification decision is taken by the fusion of the output of multiple learning machines based on a certain decision fusion scheme [4]. Multiple classifiers are commonly structured in 2 schemes: parallel and serial connection. The parallel combination is typically used in remote sensing applications.

The performance of an ensemble is highly correlated with individual classifiers and their combination scheme. For this reason, it is imperative to make a decision about how to choose classifiers from a classifier ensemble and how to combine them [15]. In classifier ensemble approaches, two approaches have been commonly applied in literature: (1) the static selection, where the best classifier (or a subset of classifiers) for all samples is selected from the individual classifiers pool. (2) Dynamic selection, where for each unclassified pixel is a specific classifier (or a subset of classifiers) that appears to be more suitable to be selected [16].

This study focuses on the Static Classifier Selection. In this method, a classifier ensemble is addressed that use a variant of the base classifier that is known to be a weak base classifier where the classifier is not tuned to performs its best. We distributed the feature space randomly among the ensemble. As a combination scheme we used majority voting.

4 Experiment Setup and Outcomes

In this part, we investigate the ability of RF Tree Based Ensemble to extract land-use classes in dense urban areas. Its average performance is also compared to other classifier based ensemble such as three ensembles of neural networks: FFNN based classifiers, radial basis neural network base classifiers and three ensembles of statically based classifiers: Linear Classifier, K-nearest Neighbour Classifier and Parzen Window Classifiers.

4.1 Data Set

One important point of using machine learning for very high resolution aerial/satellite image analysis is the size of the data used in the analysis. In literature, most studies rely on ground truth data that were manually labeled for both training and testing purposes [11, 17]. However, this is not only time consuming but also results in small datasets in aerial image analysis. Usually, very high resolution datasets cover a fairly small area of a city, ranging from 1 km^2 to 10 km^2 [11]. Good results on a small dataset do not necessarily indicate good performance regarding a whole urban area, specifically if that area differs from the scene observed while training. Consequently, acquiring labeled data that are highly accurate is essential for both evaluating present approaches and training new algorithms.

In our experiments, hand-labeling data is not necessary as the ground truth information is provided by the city. The wealth of correctly labeled data for roads makes it

an excellent land-use/land-cover where one can apply machine-learning algorithm for road extraction. In our experiment we detect roads from a large dataset for the city of Kitchener-Waterloo (K-W) and the city of Toronto Ontario, Canada. The Geospatial Centre of the University of Waterloo [18] had made the dataset available for this research. We used three datasets: two aerial datasets for the city of KW and one QuickBird satellite for the city of Toronto. The ortho-rectified aerial mosaic images for the KW dataset are 12 cm in pixel resolution and were taken by a digital color airborne camera with 8-bit radiometric resolution as well as infrared (CIR) mosaic images. We divided the ortho-mosaic into 280 images to be input into the classifiers while the ortho-rectified aerial mosaic images for the Toronto greater area dataset 19 is available in RGB bands only and was taken in April 2007. The QuickBird satellite dataset [20] is of 60 cm resolution and was taken in 2006. The main land-cover/land-uses of interests in our study are roads, buildings and green areas such as parks.

4.2   Experiment Setup

The data is segmented first as in [21] where both the spatial and spectral features were used in the clustering based segmentation process.

We used standard MATLAB classifiers that were trained with 50 % of the input data, validated over 20 % of the input data tested over 30 % of the data. The divided datasets have the same classes' distribution as the originally input data set in each of the three dataset used. The input features of the ensemble are the colour (RGB, Lab and HIS) and texture (Gray-level Co-occurrence Matrix) of the segmented parts. Using the 3 multispectral bands of the image for a window of 5 by 5 pixel size, the input feature vector is 261 dimensional image features.

For the RF tree based ensemble we investigated the effect of the number of individual trees. We conducted an experiment were the number of trees was varied from 10–100 trees and used the default values in Matlab for the rest of the variables. We found that 30 trees give the best performance in our case.

We are comparing our results to those of neural network and statically based ensembles. Each ensemble has 9 base classifiers and each classifier in the ensemble was fed with an input feature vector of 29 sub-features. All classifiers were trained/validated separately applying the training/validation sets. The classification results were averaged over forty runs. As we targeted a set of weak classifiers, no parameter optimization was done for the ensemble.

4.3   Experiment Results

The training and test accuracies for the different approaches are demonstrated in Table 1. The results are averaged over the three datasets. The table clearly indicates the advantages of the RF tree based ensemble. The accuracy increased up to 89 % for road class, which is 14 %, enhanced over the best ensemble method and 8 % enhanced over the average ensemble performance. The computation time of RF- tree is almost 1/3 less

Table 1. Comparison of the averaged classification accuracies of road using: Random forest tree based ensemble, and ensembles of Linear Classifiers, KNN Classifiers, Parzen Window Classifiers and Neural Networks Classifiers, applied on the three datasets training, validation and test sets images.

Ensemble of classifier	Training accuracy	Validation accuracy	Test accuracy
Linear classifier	76.155	75.021	74.814
KNN	100	81.514	80.831
Parzen Window	81.121	81.051	80.571
FFNN	80.112	79.552	79.588
RBN	81.522	82.017	87.536
PNN	81.561	82.451	82.019
RF tree base ensemble	90.612	89.691	89.301

(a) A sample of a test scene from the KW aerial image dataset

(b) The extracted roads are overlaid the original test scene

Fig. 1. Road classification and extraction using a RF tree based ensemble for KW aerial dataset.

than the neural network compared ensemble approaches. Qualitative result is shown in Fig. 1 for KW aerial dataset.

5 Conclusion

Road classification in dense urban areas from aerial data has been investigated. Experimental results indicate that the RF tree based ensemble yielded excellent accuracies: 89 % for classification of complex dense urban scenes, and it outperformed the highest accuracies for the other compared ensemble by 14 %. These results are obtained using a large dataset which are expected to get close results when applied to other urban datasets.

In addition, RF computational time is normally 55 % less than that of other ensemble methods used in our experiments. This should encourage the use of RF classifiers for large datasets of very high-resolution images and when updating geo-spatial databases.

References

1. Mayer, H.: Object extraction in photogrammetric computer vision. ISPRS J. Photogram. Remote Sens. **63**(2), 213–222 (2008)
2. Campbell, J.: Introduction to Remote Sensing, 4th edn. The Guilford Press, New York (2007)
3. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. Int. J. Remote Sens. **28**, 823–870 (2007)
4. Kuncheva, L.I.: Combining Pattern Classifiers: Methods and Algorithms. Wiley, Chichester (2004)
5. Bo, Y.C., Wang, J.F.: Combining multiple classifiers for thematic classification of remotely sensed data. J. Remote Sens. **5**, 555–564 (2005)
6. Benediktsson, J.A., Chanussot, J., Fauvel, M.: Multiple classifier systems in remote sensing: from basics to recent developments. In: Haindl, M., Kittler, J., Roli, F. (eds.) MCS 2007. LNCS, vol. 4472, pp. 501–512. Springer, Heidelberg (2007)
7. Del Frate, F., Pacifici, F., Schiavon, G., Solimini, C.: Use of neural networks for automatic classification from high-resolution images. IEEE Trans. Geosci. Remote Sens. **45**, 800–809 (2007)
8. Du, P., Zhang, W., Sun, H.: Multiple classifier combination for hyperspectral remote sensing image classification. In: Benediktsson, J.A., Kittler, J., Roli, F. (eds.) MCS 2009. LNCS, vol. 5519, pp. 52–61. Springer, Heidelberg (2009)
9. Giacinto, G., Roli, F.: Design of effective neural network ensembles for image classification processes. Image Vis. Comput. J. **19**(9/10), 699–707 (2001)
10. Baltsavias, E.P.: Object extraction and revision by image analysis using existing geodata and knowledge: current status and steps towards operational systems. ISPRS J. Photogram. Remote Sens. **58**(3–4), 129–151 (2004)
11. Kluckner, S., Mauthner, T., Roth, P.M., Bischof, H.: Semantic classification in aerial imagery by integrating appearance and height information. In: Zha, H., Taniguchi, R.-i., Maybank, S. (eds.) ACCV 2009, Part II. LNCS, vol. 5995, pp. 477–488. Springer, Heidelberg (2010)
12. Yu-Chang, T., Kun-Shan, C.: An adaptive thresholding multiple classifiers system for remote sensing image classification. Photogram. Eng. Remote Sens. **75**, 679–687 (2009)
13. Breiman, L.: Random forest. Mach. Learn. **45**(1), 5–32 (2001)
14. Waske, B., Benediktsson, J.A., Arnason, K., Sveinsson, J.R.: Mapping of hyperspectral aviris data using machine learning algorithms. Can. J. Remote Sens. **35**(S1), 106–116 (2009)
15. Kang, H.J., Doermann, D.: Selection of classifiers for the construction of multiple classifier systems. In: Proceedings of the Eight International Conference on Document Analysis and Recognition, pp. 263–268 (2005)
16. Smits, P.C.: Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection. IEEE Trans. Geosci. Remote Sens. **40**(4), 801–813 (2002)
17. Nguyen, T., Kluckner, S., Bischof, H., Leberl, F.: Aerial photo building classification by stacking appearance and elevation measurements. In: Proceedings ISPRS, 100 Years ISPRS-Advancing Remote Sensing Science on CDROM (2010)

18. Tri-Cities and Surrounding Communities Orthomosaics [computer file]. Waterloo, Ontario: The Regional Municipality of Waterloo (2014)
19. Greater Toronto Area Orthoimagery 2007 [computer file]. Ontario Ministry of Natural Resources (2007)
20. QuickBird Satellite Imagery {computer file}, Digital Globe, Longmont, Colorado, USA (2006)
21. Bedawi, S.M., Kamel, M.S.: Segmentation of very high resolution remote sensing imagery of urban areas using particle swarm optimization algorithm. In: Campilho, A., Kamel, M. (eds.) ICIAR 2010. LNCS, vol. 6111, pp. 81–88. Springer, Heidelberg (2010)