URBAN LAND COVER CLASSIFICATION WITH VERY HIGH RESOLUTION SATELLITE IMAGERY BY MACHINE LEARNING

Chenghua Shi, Xiaojiing Huang, Chenguang Hou, Soo Chin Liew
Centre for Remote Imaging Sensing and Processing (CRISP), National University of Singapore.
10 Lower Kent Ridge Road, Blk S17 Level 2, Singapore 119076
Email: {crssc, crshxj, crshc, scliew}@nus.edu.sg

KEY WORDS: Land cover classification, Machine learning, Convolutional Neural Network, Urban remote sensing, Very high resolution

ABSTRACT: Land cover classification is an important topic of satellite image analysis. In recent years, machine learning image classification has been increasingly used in land mapping and feature extraction showing advantages in many cases. In this paper we used a machine learning method for urban land cover classification using WorldView-2 imagery with 8 multispectral bands at 2-m resolution and a 0.5-m panchromatic band. The WorldView-2 imagery was acquired on 26 Jan 2020 covering a region of interest (ROI) in Singapore from 103°48'22.921"E to 103°51'48.901"E and 1°24'20.486" N and 1°21'6.606"N was chosen in this study. After radiometric correction and spectral-preserving pan-sharpening, the pan-sharpened multispectral reflectance was used for classification. Eight classes (Tree, Grass, Cloud, Water, Building, Bare soil, Shadow and Road/Paved) were defined. A thousand samples, each with 11 spectral features wee extracted from the imagery. The features were 8 spectral bands from the pan-sharpened multispectral image plus NDVI = (NIR1 - Red)/(NIR1 + Red), REVI = (RedEdge – Red)/(RedEdge + Red) and NYVI = (NIR2 - Yellow)/(NIR2 + Yellow). Eighty percent of the samples were used to train a Convolutional Neural Network (CNN) model with depth 5 and kernel size 5 x 5 pixels. The remaining 20% were used for validation and an accuracy of 92.1% was achieved. The trained CNN was applied to classify the whole ROI and the result was compared to a ground truth map. Two regions were selected for accuracy assessment. One was dominated with man-made objects (Buildings and Road/Paved), the other mainly contains natural objects (Tree, Grass, Water and Bare soil). The overall accuracy was 94%. This high accuracy strongly suggests that the machine learning method is a good approach for urban land cover classification using very high resolution satellite imagery

1. INTRODUCTION

Land Cover and land use classification has been an important and challenging task of imagery analysis in earth science (Bhaskaran et al., 2010). Land Cover and land use classification results are the basis for many environmental and Socio-economic applications. Numerous methods have been developed over the past decades for satellite image classification (Platt et al., 2008). The traditional pixel-based methods can be coarsely grouped into unsupervised and supervised classification approaches that classify images pixel by pixel only according to spectral features, ignoring the correlation of surrounding pixels (Gao 2009, Regniers, 2016). But the object-oriented methods classify images not only pixel's spectrum values but also consider the relationship of pixels. Through image segmentation, intra-object (spectrum, shapes and texture) and inter-object information were used for classification. (Platt et al., 2008, Su et al., 2008), With the increasing spatial resolution of satellite imagery numerous land cover classification methods have been developed with improved level of details and accuracy (Mathieu et al., 2007). In recent years, machine learning image classification has been increasingly used in land mapping and feature extraction showing advantages in many cases （Rodriguez-Galianoa et al., 2012, Gaetano et al, 2018, Talukdar et al., 2020).

However many of these approaches cannot classify the targets with high level of accuracy in the very complex urban area. For example, it is difficult to distinguish Building from Road, Bare soil and Tree is quite hard to be separated from Grass. With 2.0 meter resolution of 8 bands ranging from visible to near infrared and 0.5 meter resolution of panchromatic band, the pan-sharpened multi-spectral Worldview-2 imagery enables us a very detailed level of land cover classification with high accuracy in urban area. By use of Machine Learning methods this problem can be solved quite well.

2. MATERIAL AND METHHOD

2.1 Datasets and Preparation

The Worldview-2 imagery acquired on 26 Jan 2020 covering a region of interest (ROI) in Singapore from 103°48'22.921"E to 103°51'48.901"E and 1°24'20.486" N and 1°21'6.606"N was selected as study area shown as Figure1. The size of image is 12720 × 11920 pixels equivalent to 6.36 × 5.96 km2. The study area covers urban area with complex man-made objects as well as trees and grass.

Figure 1. The ROI of true color worldview-2 image that was acquired on the 26th of Jan, 2019.

The data from the satellite images were preprocessed before using it for classification work. The data was corrected radiometrically by converting the digital numbers of the imagery to radiance values. This was performed with the following equation,

$$L_{\lambda} = KN + s \tag{1}$$

Where N is the digital number from the satellite image, K is the calibration constant, which can be obtained from the metafile, and s is the offset.

The radiance values was then converted to reflectance values with atmospheric correction from

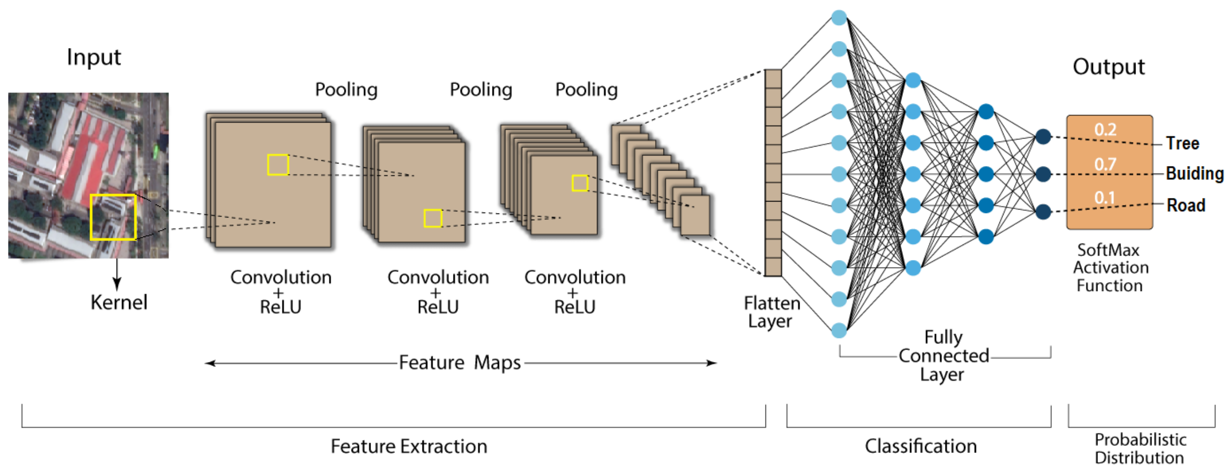
$$\rho_{TOA}(\lambda) = \frac{\pi L_{\lambda} d^2}{F_{\lambda s} \cos\theta_s} \tag{2}$$

Where d is the distance of the sun from earth, $F_{\lambda}s$ is the solar irradiance and Θ is the cosine of the sun angle (Gueymard, 1995).

The 2 meter resolution multiple bands reflectance image was then pan-sharpened into pan-sharpened multi-spectral image by use of 0.5 meter grayscale panchromatic image.

2.2 Creation of CNN model

The structure of our Convolutional Neural Network (CNN) model is illustrated in Figure 2.



904 samples from 8 classes (Tree, Grass, Cloud, Water, Building, Bare soil, Shadow and Road/paved) were chosen in the study area to train the Convolutional Neural Network (CNN) model. The input Kernel size is 5X5 pixels, 8 Wordview2 bands (coastal blue, blue, green, yellow, red, red edge, near infrared 1 and near infrared 2 plus NDVI [defined as (nir - red) / (nir + red)], REDVI defined as [(redEdge – red) / (redEdge + red)] and NYDVI defined as [(nir2 - yellow) / (nir2 + yellow)] total 11 layers are input into the CNN model. The Network Depth is 5 and Average Pooling was used. Sample statistics are shown in Table1. 80% of samples were for training the model and the rest 20% were for validation. The output of the CNN model is classified image, its accuracy can reach 92.1%.

Table1 statistics of samples

Class	Tree	Grass	Cloud	Water	Building	Bare Soil	Shadow	Road/paved	Total
polygons	366	128	12	54	112	102	92	38	904
pixels	2310019	911672	216849	1369380	3033080	2174454	480074	2393600	12889128

3. Results

The trained CNN was applied to classify the whole ROI image. The classification results image is shown in Figure 3.

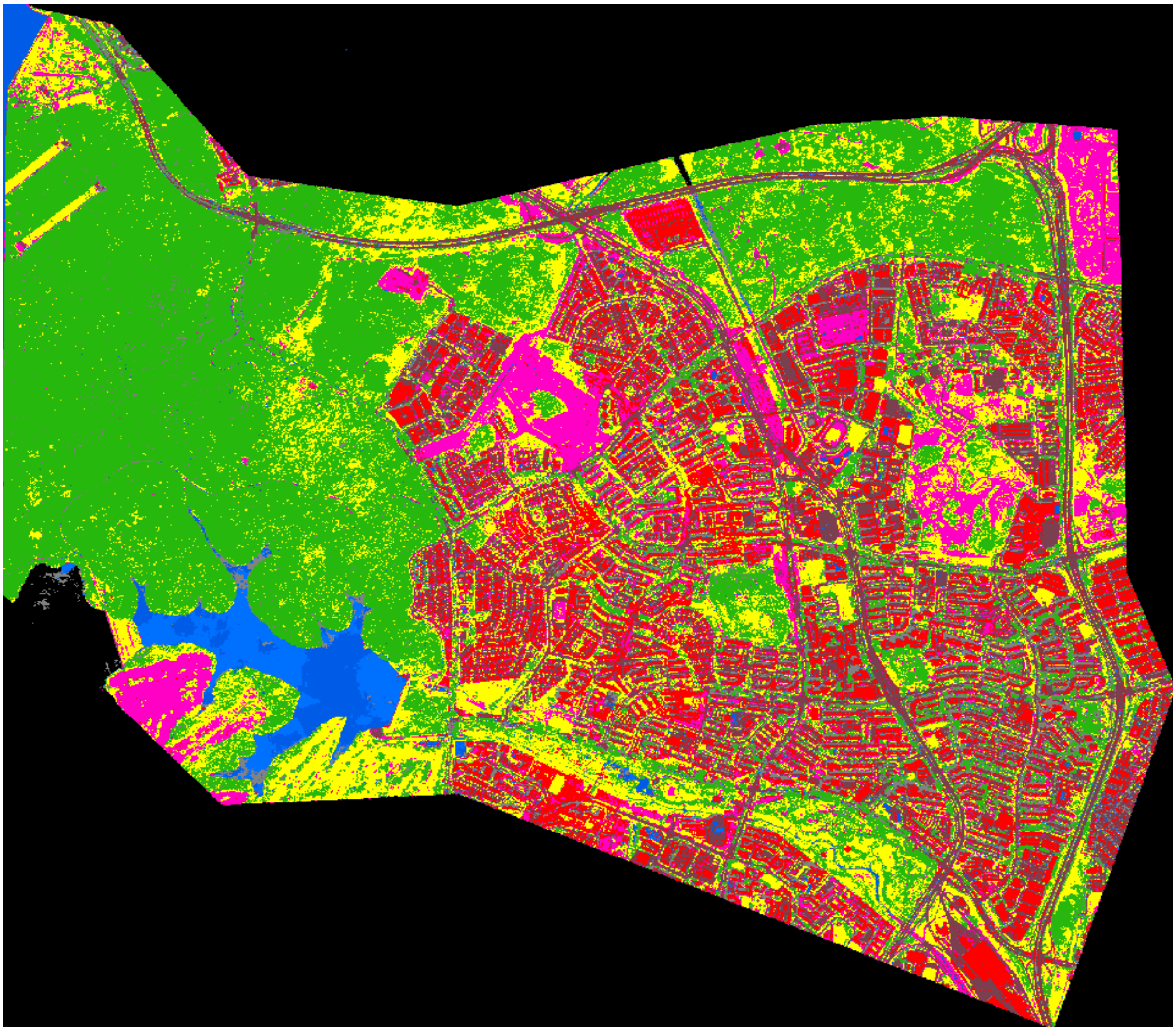


Figure 3. The classification results of ROI image. The classification color code is: Green = Tree, Yellow = Grass, White = Cloud, Blue = water, Red = Building, Magenta = Bare soil, Purple = Road/paved, Gray = Shadow and Black = No Data.

The result was compared to ground truth map that was created by visually inspection. Two regions were selected for accuracy assessment. One was dominated with man-made objects e. g. Buildings and Road/Paved (Figure 4), the other mainly contains mainly natural objects e. g. Tree, Grass, Water and Bare soil (Figure 5).

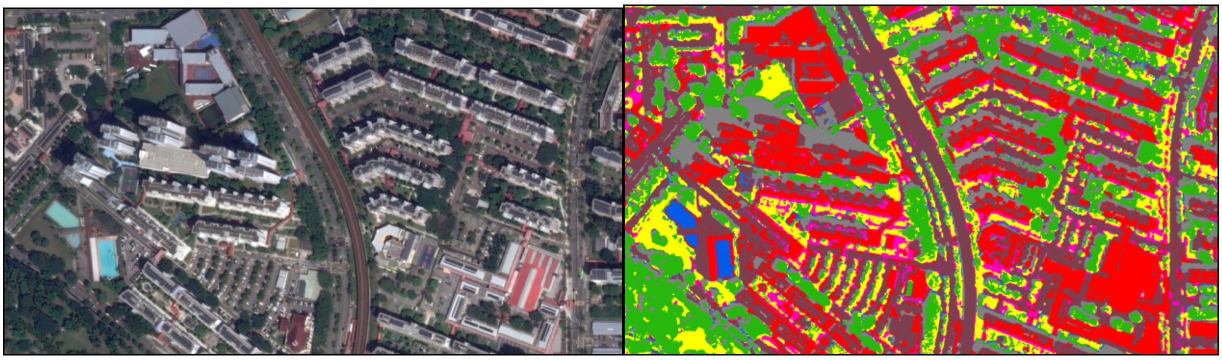


Figure 4. The first region. Left is true color image. Right is classification. The classification color code is: Green = Tree, Yellow = Grass, White = Cloud, Blue = water, Red = Building, Magenta = Bare soil, Purple = Road/paved and Gray = Shadow.

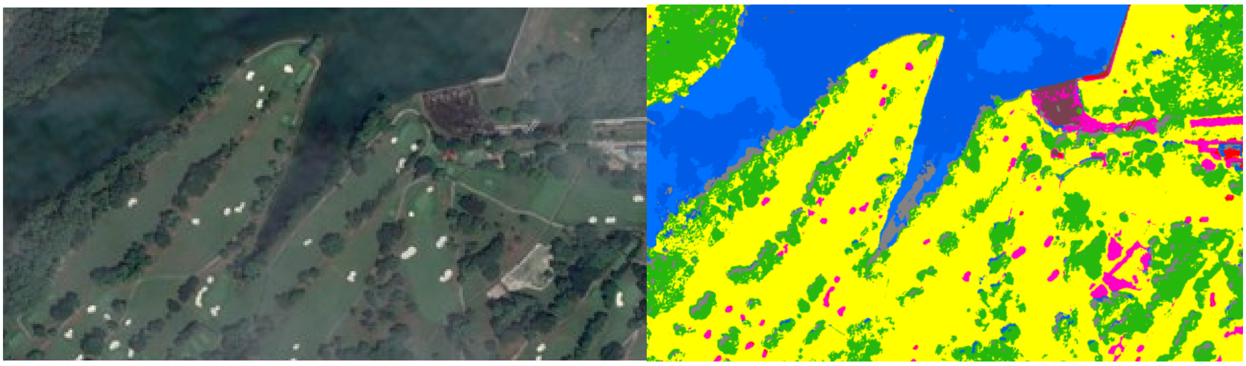


Figure 5. The second region. Left is true color image. Right is classification. The classification color code is: Green = Tree, Yellow = Grass, White = Cloud, Blue = water, Red = Building, Magenta = Bare soil, Purple = Road/paved and Gray = Shadow.

A confusion matrix calculated from those two regions is shown in the Table 2. Accuracy was 94.75 %. Tree, Grass and water were well classified, Producer's accuracy were more than 93%, Building had the lowest Producer's accuracy (71.1%): the most errors were they were misclassified as Road/paved (25.0%); the Bare-Soil had the second lowest Producer's accuracy (80.4%): the most errors were misclassified as Grass (9.6%) and Road/paved (5.7%); Cloud and Shadow also had very high Producer's accuracy but these might not represent the real situation because their small coverages in those two regions. As for User's accuracy, highest values were in Tree, Grass Water and Building, Cloud and Shadow cannot represent the real situation; Bare-Soil and Road/paved had lowest value. The Overall Accuracy was 94.5%, Kappa Coefficient = 0.9064.

Table 2. The confusion matrix of results

Class	Tree	Grass	Cloud	Water	Building	Bare-Soil	Shadow	Road / paved	ProdACC (%)	UserAcc (%)
Tree	97.1	5.2	0.1	0.0	0.1	1.3	1.3	2.8	97.1	99.1
Grass	1.4	93.0	0.2	0.0	0.1	9.6	0.0	1.3	93.0	89.1
Cloud	0.0	0.0	99.1	0.0	0.0	0.0	0.0	0.0	99.1	100.0
Water	0.0	0.0	0.2	99.4	0.0	0.0	1.4	0.5	99.4	99.4
Building	0.0	0.2	0.0	0.1	71.1	2.6	1.0	2.9	71.1	97.3
Bare-Soil	0.1	1.4	0.0	0.0	1.8	80.4	0.3	1.9	80.4	58.0
Shadow	1.4	0.1	0.0	0.1	2.0	0.4	95.3	3.3	95.3	37.7
Road/ paved	0.0	0.1	0.5	0.4	25.0	5.7	0.6	87.3	87.3	59.0
Overall Accuracy = 94.5%										
Kappa Coefficient = 0.9064										

4. Discussion and Conclusion

The Convolutional Neural Network (CNN) model of Deep Machine Learning method has been used for the classification of very high spatial resolution satellite images. The method was applied to a world-view 2 image and an accuracy of 94.5 % was achieved. With this method, Water was very well classified because its specific features. Vegetation (including Tree and Grass) and non-vegetation (Including Building, Bare-soil Road/paved) were separated well. Inside Vegetation classes, Tree and Grass were also separated to good degree. But among the Vegetation classes, Building, Road/paved and Bare Soils were a little misclassified through the lowest Producer's accuracy rate still above 70%. Because the spectral value of these classes are quite close to each other, these Producer's accuracy rates are acceptable.

The classification results can be improve with adding more samples and remove bad samples. Further improvement may be made by modifying the parameter of the CNN model, for example, when shared weight for classes changed, the model's accuracy also vary. Some auxiliary data from other resources can be added to the input layers to increase the accuracy.

5. REFERENCES

Bhaskaran, S., Paramananda, S., Ramnarayan, M., 2010. Per-pixel and object-oriented classification methods for mapping urban features using IKONOS satellite data. Applied Geography (30) 650-665

Gao J. 2009. Digital analysis of remotely sensed imagery. New York (NY): The McGraw-Hill Companies.

Goetz, S. J., Wright, R. K. Smith, , A. J. , Zinecker, E. and Schaub, E., "IKONOS imagery for resource management: Tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region," Remote sensing of environment, vol. 88, no. 1, pp. 195-208, 2003.

Gueymard, C. A., 1995, SMARTS: a simple model of the atmospheric radiative transfer of sunshine: algorithms and performance assessment. *Technical Report No. FSEC-PF-270-95. Cocoa, FL: Florida Solar Energy Center.*

Krizhevsky A, Sutskever I, Hinton G (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. Curran Associates, Inc., pp 1097–1105

Mallinis, G., Koutsias, N., Tsakiri-Strati, M., Karteris, M., 2008. Object-based classification using Quickbird imagery for delineating forest vegetation polygons in a Mediterranean test site. ISPRS Journal of Photogrammetry and Remote Sensing (63) 237-250.

Mathieu, R., Freeman, C., Aryal, J., 2007. Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery. Landscape and Urban Planning, (81) 179-192.

Platt, R.V., Rapoza, L., 2008. An evaluation of an object-oriented paradigm for land use/land cover classification. The Professional Geographer (60), 87-100.

Gaetano, R., Ienco, D., Ose K. and Cresson, R., 2018, A Two-Branch CNN Architecture for Land Cover Classification of PAN and MS Imagery, Remote Sensing (11), 1746

Regniers, O.; Bombrun, L.; Lafon, V.; Germain, C. Supervised Classification of Very High Resolution Optical Images Using Wavelet-Based Textural Features. IEEE Trans. Geosci. Remote Sens. 2016, 54, 3722–3735

Rodriguez-Galianoa, V. F., Ghimire, B., Rogan J., Chica-Olmo, J. and Rigol-Sanchezc, J. P. 2012, An assessment of the effectiveness of a random forest classifier for land-cover classification, ISPRS Journal of Photogrammetry and Remote Sensing, (67) 93-104

Su,W., Li, J.et al., 2008, Textural and local spatial statistics for the object-oriented classification of urban areas using high resolution imagery. International Journal of Remote Sensing (29) 3105-3117

Talukdar, S., Singha, P., Mahato, P., Shahfahad, Swades Pal, Liou, Y. andAtiqur Rahman, 2020, Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review, Remote Sensing. 2020, 12(7), 1135;