Multi-Level Image Segmentation for Urban Land-Cover Classifications

Nurhanisah Hashim and Juazer Rizal Abdul Hamid
Faculty of Architecture, Planning and Surveying, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

nurhanisahhashim@uitm.edu.my, juazer@uitm.edu.my

Abstract. Variation land-cover features, which include natural and man-made objects, lead to the advent of features that are spectrally similar. Object in urban area tend to have spectral similar response that can easily misclassified from one to another for example in the case of tree and grass as well as asphalt building roof and asphalt road. Object based classification approached instead of pixel based will improved the misclassification yet will increase the accuracy of land-cover classification. Using Worldview-2 multispectral satellite image as a primary data, while normalized Digital Surface Model (nDSM) derived from Light Detection and Ranging (LIDAR) data and indices image, the image segmentation process utilizing multiresolution segmentation algorithm and spectral difference was conducted. Before going through classification process, twelve segmentation levels were constructed to create image objects. Three classification algorithm including Support Vector Machine (SVM), BAYES and K-Nearest Neighbour (KNN) were choose to be tested to identify which algorithm gives the best classification result of the urban area target. The results from the study indicate statistically significant difference in classification accuracy between each algorithm: Based on Kappa statistics, user’s and producer’s accuracy, as well as visual examination and overall accuracy performance, BAYES with overall accuracy of 85.51% has depicted to have the best land-cover classification accuracy result.

1. Introduction
Mapping and identifying land-cover is important for environmental and global studies, planning activities and resource management. Remote sensing data such as satellite image can be useful for urban planners as updated information of urban land feature can be extracted from the image. Remote sensing data whether from airborne or space borne data are available at various spatial, temporal and spectral scales which will give an advantage for urban environment analysis since this environment is made up with a complex and a wide variety of surface composition. Hence having a wide range of remote sensing data as well as high accuracy data will consequent ly improve productivity of very accurate urban mapping.

The complexity of urban targets that include various type of features having high spatial and spectral heterogeneity makes the urban land-cover feature extraction complicated yet challenging. Features in urban area can be varied from man-made, water area to vegetation and forested area. Furthermore, it will be more challenging when it comes to map detailed urban targets known as intra-urban land-cover mapping which include various types of feature and material with different physical properties of impervious surface to be identified. Features in urban area have similar spectral coloration which is hard to discriminate between them by using only four traditional spectral bands [1]. A big range of spectral
bands will greatly improve the potential to discriminate spectrally similar features especially for mapping detailed urban target[2]. With very limited number of bands, similar urban features such as paved street and darkened asbestos will definitely hard to discriminate.

However, it is not necessary by only depends on high-resolution data itself. Selecting a suitable classification method that is convenient with high-resolution data is very important to assure providing accurate land-cover classification maps. Since the data that will be used in this research to extract various types of land cover are in high resolutions such as Worldview-2 satellite image and Light Detection and Ranging (LIDAR) data, it is not relevant or might not be efficient to use pixel based as classification method [3].

2. Literature Review

2.1 Pixel based and object-based classification method

The approached of traditional pixel classification method which is only based on pixel spectral value show unsatisfactory results when applied to high-resolution images [4]. In high-resolution image each pixel is related to the components of the image, thus resulting unnecessary classes detected when it is classified. In order to overcome this situation, object-based classification method had been developed which makes high-resolution data and image an ideal source for land cover classification. A lot of research had been conducted and proved that object based method had improved the classification result compared to pixel based method [5].

Before Object based classification method is developed, numerous study favored of pixel based classification method to extract various types of land-cover classification. Since the advent of high-resolution remote sensing image, pixel based classification method seems to be not very compatible to apply together with high-resolution image [6]. The pixel-based image classification will produce a large number of data redundancy since the method cannot meet high-resolution satellite image classification precision [7]. The drawback of pixel-based classification technique is, it only focuses on spectral information and not considering spatial information into account. A pixel is classified depending on its spectral value irrespective of the value of the neighboring pixels resulting a very sensitive to noise and often lack of spatial consistency [8]. As an alternative to pixel based classification, object based classification methods improve interpretability and accuracy in high-resolution image and have said to be well established method to work together with high-resolution image especially for urban feature classification [5].

2.2 Image Segmentation: Multiresolution Algorithm

The multi-resolution segmentation had been widely used among the researchers when creating the image object in an object-based classification. The multiresolution segmentation belongs to the category of region-based techniques, implemented a region growing technique in order to merge the objects from a very fine size (sub objects) which is in lower segmentation level to a large object (super object) at the higher level while minimize the heterogeneity between objects by control the weight of the object size in the merging process. Since a different value of weight will create a different size of objects; thus, giving a large value of weight resulting a larger objects created from several objects caused of merging process [3].

For an effective multi scale analysis, segmentation of an image can be created through hierarchical approaches at different object level. This approach will give an advantage for object-based method whereas various sizes of objects can be created at each hierarchy level as set up by the analyst since various size of objects often exist in the scene especially in urban area [9]. This multi scale analysis is possible with e-Cognition software where object on different size can be represent with different scale by different object level [10]. Study done by [11] employed four different scale levels which is 10 (level 1), 25 (level 2), 50 (level 3) and 100 (level 4) to segment different object in a different classes. Grass, trees/shrubs, pools and unmanaged soil were segmented at scale level 1, Buildings at scale level 2, lakes/ponds at scale level 3 and scale level 4 segmenting other impervious surface and produced high
accuracy ranging from 80% to 99% for each class, hence conclude that it is such a relevant approach for segmenting different objects at different scale level.

2.3 High resolution remote sensing data for urban feature extraction
Worldview-2 high-resolution satellite image space borne sensor launched in 2010 and comes with additional four new spectral bands, which make this satellite image one of the promising sources to extract land-cover features. While Worldview-2 image gives high accuracy in horizontal direction, LIDAR give high-resolution and accuracy in elevation. LIDAR had been widely used to generate Digital Terrain Model (DTM) and Digital Surface Model (DSM). Airborne LIDAR data are traditionally used to classify ground and non-ground points, and some methods have been developed to detect vegetation, or building features from LIDAR point clouds. Other than that, LIDAR also are capable to record the intensity of the reflected energy [12] [13].

Integrating Worldview-2 satellite image with airborne LIDAR data can lead to high accuracy of land-cover classification result. Since Worldview-2 image offers four new bands, which gives higher spectral resolution from any other imagery as well as 0.5 meter for spatial resolution, it is believed that the higher spectral and spatial resolution offer by this image have a potential to overcome the spectral variations issues associated with other high resolution multispectral image when dealing with spectrally and spatially similar objects that usually appear in urban environment. A lot of research has been conducted on integrating satellite image or aerial image with LIDAR data for feature extraction and land-cover classification purpose [14] [15] ; thus, came out with very high classification result compared to only used satellite image alone. This is because the elevation attributes from LIDAR data can reduce the spectral similarity problem of urban features by offering its vertical spatial variation. Furthermore, it has been shown that heterogeneous spectra of building can be separated using elevation information derived from LIDAR hence will improve the classification result.

3. Methodology
This section will be divided into four (4) sub-section that will explain on the study area, data acquisition, pre-processing and processing based on the methodology of the study. General workflow as shown in Figure 1.

![Figure 1. General Workflow](image)

3.1 Study area
The study area is located at the area of the city of Shah Alam, Selangor with high spatial heterogeneity of different land-cover classes. The coordinate of the lower left is N101.50487814, E3.07205192 and the coordinate of the upper right is N101.51529563, E3.08940841. The total coverage of the study area is 1.0 x 1.7 Km square. The study area as shown in Figure. 2.

3.2 Datasets
Due to the complexity of the urban features in the study area, datasets to be acquired should also be of high-resolution data, which supposedly have the capability to facilitate detection and identification
better and more accurately. In this aspect, two datasets were used comprising of Worldview-2 and LIDAR representing the city of Shah Alam. The main data used in this study is Worldview-2 satellite image. This sensor provides a very high spatial resolution, which can give an advantage for extracting objects such as buildings, water body, shrubs, tree canopy and many more. Literature have shown that by integrating LIDAR elevation data as an ancillary data will give better accuracy for classification result. In order to further explore this aspect, LIDAR elevation data, which is normalized Digital Surface Model (nDSM) was considered as some part of the classification.

![Study Area](image)

**Figure 2. Study Area**

3.3 **Pre-processing**
There are total of two steps for pre-processing which include radiometric correction and pan-sharpening image. For radiometric correction, the conversion of DN to at surface reflectance is necessary in order to get full used of the image and the information on the band reflectance. The pan-sharpening process was undergone by using the Gram-Schmidt method to fused the Worldview 2 multispectral image with the Panchromatic band to derive 0.5-meter spatial resolution imagery.

3.4 **Image Segmentation**
The image segmentation process was done by utilizing several segmentation algorithm approaches, which is multiresolution and spectral difference to create object features. The aim of the segmentation process is to create an object in order to extract several landcover classes including natural and man-made features. Before deciding on how many levels to be used, some experiment need to be done to specify the best parameter settings of multi resolution segmentation algorithm for several feature classes. Six classes had been decided to go through experimental process in order to find out the best parameter setting and to explore the relation of the parameters given by the multi resolution algorithm. The six classes including small single building, sport field/court, swimming pool, building with rectangular shape, large building, and water area. These six classes represent the whole features of the study area in terms of object shape and size; thus, the segmentation parameter setting which give the best segmentation result will be applied to create several segmentation levels later on.

Twelve segmentation levels were decided to form with each level having different parameter settings. At this stage, eight temporary classes will be created and assign the segmented object to its class based on hierarchical classification approach. The eight classes including large building, medium building, small building, swimming pool, lake, sports field, transportation area and vegetation. The layer that will
be used in order to construct segmentation layer including all 8 multispectral bands, panchromatic band, several spectral indices [16], and LIDAR nDSM. Other than multi resolution algorithm, spectral difference algorithm also adopted in segmentation process for this study. Spectral difference was used to refine the existing segmentation result created earlier by multi resolution algorithm at the same image by merging spectrally similar image object produced by multi resolution segmentation.

3.5 Object based classifier for land-cover classification
Three classifiers namely Support Vector Machine (SVM), K-Nearest Neighbour (KNN) and BAYES were tested for object-based classification. This approach was applied in this study to investigate the capability of indices layer as well as four new additional bands of Worldview-2 image for urban features extraction and classification. Hence the performance of above-mentioned classifier will be determined based on classification result by each classifier. The three classifiers used in this study were chosen based on great performance given by each classifier. Literature have shown that Support Vector Machine has been the newest and robust classifier in remote sensing community [17]. SVM also have been most favorable classifier among the researcher for land-cover classification purpose.

4. Result and Analysis

4.1 Construction of Segmentation Level for Feature Extraction and Classification.
The top bottom technique was adopted with hierarchical approached for this study. Thus, the segmentation and extraction process were start with bigger size objects since it was found that the bigger objects are easier to segment compared to small objects. Starting at Level 1, this level was created to segment and extract some of the large building as appeared in the study area. With the scale parameter of 30 and adding the LIDAR nDSM as the additional layer, the large building was well segmented. In order to assign the large building object created to its class, the mean nDSM, mean Normalized difference NIR-1 and Blue Index (NDNB) and number of pixel value were used as a membership function. All the three conditions are the threshold value that have to be set up in order to separate large building objects with other objects.

Level 2, Level 3 and Level 5 were created mainly for segment and extract medium building objects. Since the medium buildings has the smaller size compared to large building; thus, the scale parameter assigned to these levels are smaller than the first level which is 25, 20 and 15 respectively while the layer used remain the same. For the extraction of medium buildings at these levels, the mean nDSM and mean NDNB were still used to set up the threshold value. Moreover, the mean rectangular fit and mean length/width also included as a threshold function to separate the medium building objects. Since most of the medium building object in these two levels consists of apartment houses which is nearly rectangular in shape; thus, the threshold value set up for the mean rectangular fit has to be greater than 0.9 (1 is the maximum value) while mean length/width have to be greater than 2 and smaller than 4. Figure 3 shows the segmentation of medium building in the study area.

Segmentation at Level 7 mainly focus on creating the objects of sports field and lake. With the scale parameter set up to 12, these two features were well segmented by adding the Normalized Difference Green and Red edge (NDGR) and NDNB layers. In order to assign these objects to its class, the mean Normalized Digital Water Index (NDWI) and mean NDNB were defined to separate both lake and sports field. Segmentation on Level 8 and Level 9 were done mainly for creating and separate small building features while Level 10 for pool object. These three levels which is Level 8, Level 9 and Level 10 were given the same scale parameters which is 10 but different additional layers were used; thus, creating different size and shape of objects at each level. At this stage, separating the small single building objects quite complicated as small single buildings was made up with various types of roof that leads to wide range of threshold value have to be set up. These wide range of values will include other features such as transportation area and vegetation. Here, several conditions of membership function were considered. The mean NDNB and mean nDSM value were used for separating the small buildings with the
transportation area objects while mean Normalized Difference Vegetation Index (NDVI) value to distinguish small buildings from vegetation objects.

![Figure 3. Segmentation of Medium Building a) and b) with Scale Parameter Weight to 25, c) and d) with Scale Parameter Weight to 20.](image)

Level 11 and Level 12 were created in order to segment and separate vegetation and transportation area features. Since vegetation features (grass and tree) always appear near the transportation area, for example, beside the road or between the road; thus, using a larger scale will merge these features together especially grass and transportation area since they are having similar nDSM value. After trying out several multi resolution segmentation process, value 6 and 4 were decided as a scale parameter for Level 11 and Level 12 respectively in order to have a good segmented object between these two features. Using small scale value will create smaller objects which merging process needs to be done later on. At this stage, the spectral difference segmentation algorithm was decided to use to merge and refine segmented object, which is transportation area and vegetation object created earlier. The spectral difference segmentation algorithm will merge the neighboring object-based on maximum spectral difference value defined by the user. The same process was used for transportation area. However, this time, mean NDNB value were used as a threshold condition to separate the transportation area, which actually consists of two types based on the brightness of the features. Figure 4 shows the object segmentation at Level 10 and Level 12 which almost 60 percent of the image had been segmented at Level 10 while fully segmented at Level 12.

### 4.2. Land cover classification through Object based image analysis

The accuracy assessment result shows that each sets of experiment gave different classification results when using the same training and testing sample with different classifier. The summarization of overall accuracy and Kappa Index Assessment (KIA) were presented in Table 1.
Table 1 shows the SVM classifier gives the highest overall accuracy and Kappa statistic results followed by BAYES classifier, and K-NN classifier to be the lowest accuracy. Examining the individual class accuracies, the vegetation has been identified with highest user’s and producer’s accuracy. From the results, it can be said that class with wide area coverage leads to higher user’s and producer’s accuracy such as vegetation and grass with minimum to zero omission and commission error produced, while class with small area coverage such as brick road and metallic roof leads to lower accuracy result. This situation occurs especially when dealing with area that covers with various types of class that will produce confusion due to overlapping spectral reflectance information between classes. Other than that, as an object-based image classification approach, the classifier is highly dependent on the quality of the segmentation result, which may lead to confusion especially among spectrally similar object such as road and roof material. However, this problem could be minimized by adding indices layer, texture analysis image, spectral transformation image and others.

### Table 1. Accuracy Assessment Result for Land Cover Classification using SVM, BAYES and KNN classifier

| Classifier/ Classification Accuracy | SVM  | BAYES | KNN  |
|-----------------------------------|------|-------|------|
| Overall Accuracy                  | 87.93| 85.51 | 79.65|
| Kia                               | 86.24| 83.44 | 76.96|

Besides evaluation from statistic value produced by the error matrix, evaluation based on visual interpretation of the final land-cover maps also can be done in order to determine the best land-cover classification result. Based on overall accuracy and Kappa statistic, classification using SVM produced highest classification accuracy with 87.93% and 86.24% respectively. However, depending only on overall accuracy does not reveal on how well the individual classes were classified. Thus, it is important to determine the best classification result based on overall accuracy, individual class accuracy as well as by visually examining the result may lead to more realistic result. Based on fairly detailed evaluation, it determines that classification using BAYES produced generally the best result although having a slight lower overall accuracy than the SVM classification.

5. Conclusion

Literature have shown utilizing the ancillary data can improve the accuracy assessment of land-cover classification as well as separate spectral similar object that usually appears in urban area. However, only limited study proved this by adapting an object-based image analysis method upon supervised classification approached. The rule-set have been a favorable approach among the researcher but still, the confusion between the urban class occur such as trees with grass, bare land with house roof, as well as road with house roof. In that matter; thus, this study utilized the rule-set method in order to create the image object through a segmentation process and classify them initially into eight classes including large building, medium building, small building, swimming pool, lake, sports field, transportation area and vegetation. Here, the capability of indices, which created through spectral separability of Worldview-2 bands were tested through spectral separability investigation in order to segment the image object and assigned to its particular class; thus, 12 segmentation level was created through multi resolution segmentation algorithm. Three supervised classifiers including SVM, BAYES and KNN were tested through object-based image analysis method to identify which classifier and what combination of bands/layer will minimize the confusion between spectral similar urban target as well as improved the classification accuracy.
Figure 4  a) Image Segmentation at Level 10, b) Image Segmentation at Level 12

Figure 5 Urban Landcover Classification using BAYESIAN Classifier
References

[1] E. B. Wetherley, D. A. Roberts, and J. P. Mcfadden, “Mapping spectrally similar urban materials at sub-pixel scales,” vol. 195, pp. 170–183, 2017.

[2] C. M. D. De Pinho, L. M. G. Fonseca, T. S. Korting, C. M. de Almeida, and H. J. H. Kux, “Land-cover classification of an intra-urban environment using high-resolution images and object-based image analysis,” Int. J. Remote Sens., vol. 33, no. 19, pp. 5973–5995, 2012, doi: 10.1080/01431161.2012.675451.

[3] T. Blaschke, “Object Based image analysis: a new paradigm in remote sensing?,” Proc. ASPRS 2013 Annu. Conf. ..., no. 2011, 2013, [Online]. Available: http://ispace.researchstudio.at/sites/ispace.researchstudio.at/files/287_blaschke_asprs_2013_full.pdf.

[4] M. a Aguilar, M. M. Saldana, and F. J. Aguilar, “GeoEye-1 and WorldView-2 pan-sharpened imagery for object-based classification in urban environments,” Int. J. Remote Sens., vol. 34, no. December 2014, pp. 2583–2606, 2013, doi: 10.1080/01431161.2012.747018.

[5] L. Ela and O. C. J, “Classification of impervious land-use features using object-based image analysis and data fusion,” vol. 75, no. January, pp. 103–116, 2019.

[6] Z. Chen and B. Gao, “An Object-Based Method for Urban Land Cover Classification Using Airborne Lidar Data,” pp. 1–12, 2014.

[7] T. Blaschke, “Object based image analysis for remote sensing,” ISPRS J. Photogramm. Remote Sens., vol. 65, no. 1, pp. 2–16, 2010, doi: 10.1016/j.isprsjprs.2009.06.004.

[8] A. Puissant, S. Rougier, and A. Stumpf, “Object-oriented mapping of urban trees using Random Forest classifiers,” Int. J. Appl. Earth Obs. Geoinf., vol. 26, pp. 235–245, 2014, doi: 10.1016/j.jag.2013.07.002.

[9] R. G. Lotte and S. Amaral, “Land-cover classification of an intra-urban environment using high-resolution images and geographic object-based image analysis: the case of APA Mananciais do Rio Paraiba do Su,” 2015.

[10] S. Yadav, I. Rizvi, and S. Kadam, “Urban Tree Canopy Detection Using Object-Based Image Analysis for Very High Resolution Satellite Images : A Literature Review,” pp. 4–9, 2015.

[11] S. W. Myint, P. Gober, A. Brazel, S. Grossman-Clarke, and Q. Weng, “Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery,” Remote Sens. Environ., vol. 115, no. 5, pp. 1145–1161, 2011, doi: 10.1016/j.rse.2010.12.017.

[12] J. Zhang, X. Lin, and X. Ning, “SVM-Based classification of segmented airborne LiDAR point clouds in urban areas,” Remote Sens., vol. 5, no. 8, pp. 3749–3775, 2013, doi: 10.3390/rs05083749.

[13] B. Yu, H. Liu, J. Wu, Y. Hu, and L. Zhang, “Automated derivation of urban building density information using airborne LiDAR data and object-based method,” Landsc. Urban Plan., vol. 98, no. 3–4, pp. 210–219, 2010, doi: 10.1016/j.landurbplan.2010.08.004.

[14] Y. Kim, “Object-Based Classification of an Urban Area Through a Combination of Aerial Image and Airborne Lidar Data,” pp. 1–6, 2011.

[15] Y. Chen, W. Su, J. Li, and Z. Sun, “Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas,” Adv. Sp. Res., vol. 43, no. 7, pp. 1101–1110, 2009, doi: 10.1016/j.asr.2008.11.008.

[16] N. Hashim, J. R. A. Hamid, N. M. Saraf, N. Naharudin, M. A. Halim, and M. H. Razali, “Spectral Information Extraction from Worldview-2 Image for Urban Features Identification,” ICSGRC 2019 - 2019 IEEE 10th Control Syst. Grad. Res. Colloquium, Proceeding, no. August, pp. 76–81, 2019, doi: 10.1109/ICSGRC.2019.8837079.

[17] R. Malik, R. Kheddam, and A. Belhadj-aiissa, “Object-Oriented SVM Classifier for ALSAT-2A High Spatial Resolution Imagery : A Case Study of Algiers Urban Area,” 2015.