Forest type mapping using object-based classification method in Kapilvastu district, Nepal

A. K. Chaudhary[1*], A. K. Acharya[1] and S. Khanal[1]

In the recent years, object-based image analysis (OBIA) approach has emerged with an attempt to overcome limitations inherited in conventional pixel-based approaches. OBIA was performed using Landsat 8 image to map the forest types in Kapilvastu district of Nepal. Systematic sampling design was adopted to establish sample points in the field, and 70% samples were used for classification and 30% samples for accuracy assessment. Landsat image was pre-processed, and the slope and aspect derived from the ASTER DEM were used as additional predictors for classification. Segmentation was done using eCognition v8.0 with the scale parameter of 20, ratios of 0.1 and 0.9 for shape and color, respectively. Classification and Regression Tree (CART) and nearest neighbor classifier (k-NN) methods were used for object-based classification. The major forest types observed in the district were KS (*Acacia catechu*/ *Dalbergia sissoo*), Sal (*Shorea robusta*) and Tropical Mixed Hardwood. The k-NN classification technique showed higher overall accuracy than the CART method. The classification approach used in this study can also be applied to classify forest types in other districts. Improvement in classification accuracy can be potentially obtained through inclusion of sufficient samples from all classes.

Key words: Landsat, machine learning algorithm, object-based classification

Remote sensing provides a useful source of data from which land-cover information can be extracted for assessing and monitoring vegetation changes. In the past several decades, air-photo interpretation has played an important role in detailed vegetation mapping (Sandmann and Lertzman, 2003), while applications of medium spatial resolution satellite imagery such as Landsat Thematic Mapper (TM) and SPOT high-resolution visible (HRV) alone have often proven insufficient or inadequate for differentiating species-level vegetation in detailed vegetation studies (Harvey and Hill, 2001).

Recently, object-based image analysis (OBIA) approach has been widely utilized for remote sensing studies as an alternative and also comparatively better classification approach to the conventional pixel-based image classification techniques. Successful launch of very high-resolution (VHR) commercial imaging satellites in the late 1990s are helpful for resource inventory and monitoring (Ehlers *et al.*, 2003; Ehlers, 2004). The VHR imagery is anticipated to be an alternative option to aerial photographs for characterization of forest structure and dynamics through automatic image classification technique. In the recent years, Ikonos imagery has been frequently used for vegetation mapping using pixel-based image classification methods (Wang *et al.*, 2004a; Wulder *et al.*, 2004; Metzler and Sader, 2005; Souza and Roberts, 2005). Pixel-based method, however, has constraints with VHR image classification because of decrease in classification accuracy due to high-spectral variability within classes (Yu *et al.*, 2006; Lu and Weng, 2007). It also ignores the context and the spectral values of adjacent pixels (Townshend *et al.*, 2000; Brandtberg and Warner, 2006). Various image classification techniques have been developed such as object-based, textural, and contextual image classifications in order to reduce the limitations associated with VHR images (Guo *et al.*, 2007; Lu and Weng, 2007).

The geographic object-based image analysis (GEOBIA) technique emerged since the late 1990s, to overcome human interpreters' ability to identify and delineate features of interest (Benz *et al.*, 2004; Meinel and Neubert, 2004). The GEOBIA technique could be useful to solve the problems of high-spectral variability within the

[1] Department of Forest Research and Survey, Kathmandu, Nepal
[*] E-mail: chaudharyashok1@gmail.com

same land-cover classes in VHR imagery (Yu *et al.*, 2006; Lu and Weng, 2007). To overcome the high-resolution problem and salt-and-pepper effect, it is useful to analyze groups of contiguous pixels as objects instead of using the conventional pixel-based classification unit. This will reduce the local spectral variation caused by crown textures, gaps, and shadows. In addition, with spectrally homogeneous segments of images, both spectral values and spatial properties, such as size and shape, can be explicitly utilized as features for further classification. The basic idea of this process is to group the spatially adjacent pixels into spectrally homogenous objects first, and then conduct classification on objects as the minimum processing units.

The object-based classification procedure includes image segmentation, training sample selection, classification feature selection, tuning parameter setting and, finally, algorithm execution. The accuracy of image classification is influenced by segmentation quality (Dorren *et al.*, 2003; Meinel and Neubert, 2004; Addink *et al.*, 2007). Dorren *et al.* (2003) stated the importance of image object-size in forest classification and mapping. There are no specific guidelines to take optimal segmentation size and it is a matter of trial-and-error methods which influence segmentation quality (Definiens, 2004; Meinel and Neubert, 2004). Kim *et al.* (2008) emphasized spatial autocorrelation analysis to determine optimal segmentation size for forest stands. In the recent years, pixel-based classification with texture information has been employed to improve the accuracy of forest/vegetation mapping (Ferro and Warner, 2002). Therefore, this study was carried out to map forest types using object-based classification technique and recommend the appropriate classification technique for other districts.

Materials and methods

Study area

Kapilvastu district is situated in Lumbini Zone of Western Development Region of Nepal. Geographically, it extends from 27°25’ N to 27°84’ N latitude and from 82°75’ E to 83°14’ E longitude (Fig. 1). It spreads ranging from 93 to 1,491 m above sea level. The district enjoys tropical and sub-tropical climate. Kapilvastu district covers 1,738.00 km^2 land representing with forest cover area of 63,438.42 ha.

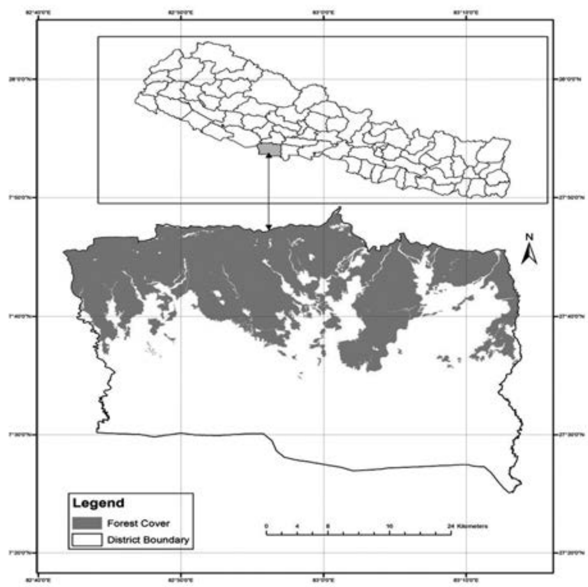

Fig. 1: Map of study area

Preliminary work

The field crew members were trained on the collection of Global Positioning System (GPS) location of sample points, basal area calculation, crown cover measurement and forest type signature collection in the field through various trainings. Consultation was done with District Forest Office (DFO), Sector Forest Offices and Ilaka staffs for delineation of major forest types.

Data

Multi-spectral satellite imagery of Landsat 8 was obtained from United States Geological Survey (USGS). The characteristics of the image are presented in Table 1. The ASTER DEM of 30 m resolution was obtained from USGS (2015) and terrain parameters (slope and aspect) were calculated which were used in Classification and Regression Tree (CART) analysis.

Table 1: Characteristics of Landsat 8 Image

Satellite	Sensor	Path-Row	Date	Band
Landsat 8	OLI and TIRS	142–41	13 Feb, 2014	2–7
Landsat 8	OLI and TIRS	143–41	19 Jan, 2014	2–7

Sampling design

Systematic sampling design was employed to establish sample points in the field. A forest mask for the study area was taken from the recent Forest Resource Assessment (FRA) of the Terai (DFRS, 2014). Since, FRA forest cover was done on physiographic region scale, there were some minor discrepancies. Those were manually edited using high resolution Google Earth Image. A systematic grid at the interval of 500 m was generated within the district and all of the generated sampled points in regular grid (n= 213) were visited in the field using GPS and dominant species in the plot identified based on the species basal area.

Point sampling

Horizontal point sampling was used to estimate basal area for forest type mapping purpose. In this sampling a series of sampling points were selected systematically distributed over the entire area to be inventoried. Trees around this point were viewed through any angle-gauge at breast height and all trees forming an angle bigger than the critical angle of instruments were counted. The basal area per hectare was calculated by multiplying Basal Area Factor (BAF) of instrument with number of tally trees to identify the forest types in the field. The particular species representing basal area greater than 60% corresponds to the same forest type as the species. Based on the dominance of species basal area, the three major forest types namely Khair/Sissoo (KS-*Acacia catechu*/ *Dalbergia sissoo*), Sal (S-*Shorea robusta*) and Tropical Mixed Hardwood (TMH) (DFRS/FRA, 2014) were found in the district. The sample plot distribution according to their categories is shown in Figure 2. The total sample points (n=213) were divided into training data sets (70%) for forest type classification and 30% sample points for evaluating classification accuracy.

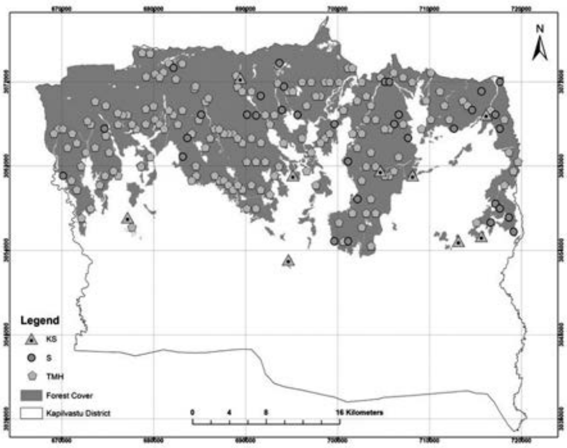

Fig. 2: Field sample points

Image analysis and mapping

The Landsat 8 images acquired were pre-processed (layer stacking, image enhancement and mosaic king). Before image segmentation and classification, non-forest areas such as agriculture, grassland, built up, river etc. were masked out since the main concern of this proposed study was to focus on forest types. Spectral bands (Band 2–Band 7) of Landsat 8 along with slope and aspect derived from ASTER DEM were used in the segmentation. Segmentation was done using eCognition Version 8.0 with a scale parameter of 20. The values of 0.1 and 0.9 were chosen for the ratios of shape and color, respectively. Spectral signatures of individual forest types were extracted from the different bands of the masked image by using training data and then classification was performed by standard nearest neighbor classifier (k-NN) and Classification and Regression Tree (CART) method which takes into consideration of spectral parameters and ancillary data (Definiens, 2004). The CART algorithm is one of the most commonly used decision trees that works as a binary recursive partitioning procedure by splitting the training sample set into subsets based on an attribute value (set) and then by repeating this process on each derived subset. The tree-growing process stops when no further splits are possible for subsets. The maximum depth of the tree is the key tuning parameter in the CART, determining the complexity of the model. In general, a larger depth can build a relatively more complex tree with potentially higher overall classification accuracy. Therefore, in this study the tree depth was set at 10. The k-NN algorithm uses an instance-based learning approach and does classification by assigning class based on the class attributes of its K-nearest neighbors. The CART technology is recently applied in ecology; it provides a low-cost, high quality alternative to approximate the human learning process and make accurate generalizations concerning the relationships of input variables and the value of the target feature, without such difficulties (Maniezzo *et al.*, 1993).

Accuracy

The overall accuracy was calculated for summary measures (Gong *et al.*, 1992), which can be used to compare individual class difference between distinct classifications (Coburn and Roberts, 2004).

Results and discussion

Forest types

The forests of Kapilvastu district were classified into three major forest types namely Khair/Sissoo (KS), Sal (S) and Tropical Mixed Hardwood (TMH). Forest types classification results based on CART and k-NN nearest neighborhood classification methods are presented in Fig. 3 and 4.

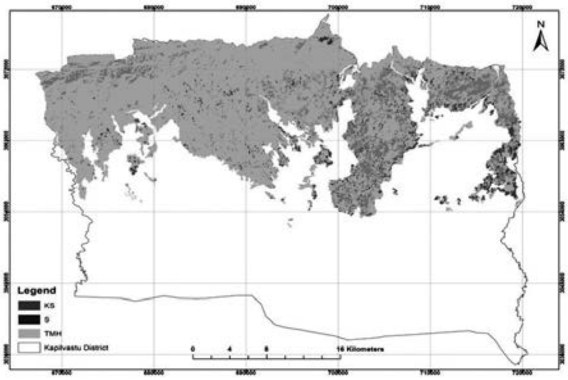

Fig. 3: Forest type classification using CART method

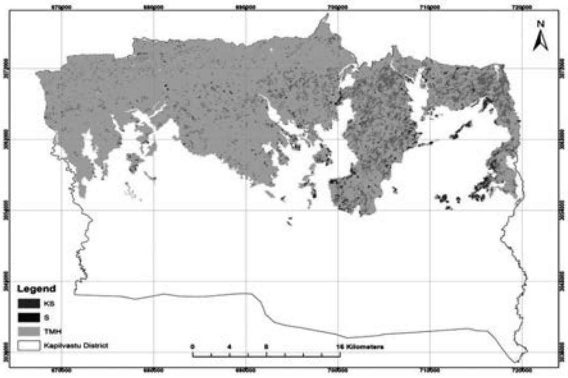

Fig. 4: Forest type classification using k-NN method

Blaschke (2003) observed that the optimal size of segmentation is critical and challenging task in GEOBIA. Therefore, as the optimum segmentation size increases the classification accuracy, it is likely that the classification results can be further improved by evaluating and selecting the optimal size.

Qian *et al.* (2015) evaluated and compared the performance of four machine-learning classifiers namely Support Vector Machine (SVM), Normal Bayes (NB), CART and k-NN using an object-based classification procedure, and found that CART method was superior to the k-NN classification. The results from this study were, however, in contrast which might be due to the less number of field samples as well as limited number of predictor variables (forest types).

The areas of different forest types were calculated using the CART and the k-NN classification methods. The areas of KS (2,865.4 ha) and S (11,427.7 ha) calculated using the CART classification method were found to be higher as compared to those (KS 1,944.9 ha and S 9,365.1 ha) obtained using the k-NN classification method. On the contrary, the area of the TMH (52,128.4 ha) computed using the k-NN classification was higher than the one (49,145.3 ha) worked out using CART method (Table 2).

Table 2: Area of three forest types calculated using CART and k-NN methods

S.N.	Forest type	CART	k-NN
		Area (ha)	Area (ha)
1.	KS	2,865.42	1,944.90
2.	S	11,427.70	9,365.13
3.	TMH	49,145.30	52,128.39

Accuracy assessment

The accuracy assessments of the two classification methods were accomplished to assess the qualities of the classified map products. The overall accuracy (69.7%) using the CART classification method was found to be slightly lower than the one (72.7%) obtained using the k-NN classification method whereas the user's accuracies for Sal (20.0%) and TMH (84.3%) were recorded higher in the CART classification than those (S: 14.3% and TMH: 80.7%) recorded in the k-NN classification (Table 3). On the contrary, the user's accuracy for KS (50%) using the k-NN classification method stood higher as compared to the one (20%) obtained using the CART method. On the other hand, the producer's accuracy for TMH (82.7%) based on the CART method was found to be lower than the one (88.5%) based on the k-NN method whereas the producer's accuracy for 'S' based on the CART method was found to be exactly two times more (18.2%) than the one (9.1) based on the k-NN method. Both the classification methods gave the same result of 33.3% producer's accuracy for KS. The classification accuracy as reported by Czaplewski and Patterson (2003) was only 40%

Table 3: Accuracy assessment of CART and k-NN methods

	Ground-truth field samples											
	CART						**k-NN**					
Classes	**KS**	**S**	**TMH**	**Total**	**User's accur. (%)**	**Error of com. (%)**	**KS**	**S**	**TMH**	**Total**	**User's accur. (%)**	**Error of com. (%)**
KS	1	3	1	5	20	80	1	1		2	50	50
S		2	8	10	20	80		1	6	7	14.28	85.72
TMH	2	6	43	51	84.31	15.69	2	9	46	57	80.70	19.30
Total	3	11	52	66			3	11	52	66		
Producer's accur. (%)	33.33	18.2	82.69				33.33	9.09	88.46			
Error of omiss. (%)	66.67	81.8	17.31				66.67	90.91	11.54			
Overall accur. (%)				**69.69**						**72.72**		

or less for thematic information extraction at the species-level based on the Landsat TM and SPOT HRV Images. The results of this study however higher accuracy although there were only three classes.

Conclusion

The proposed methods offer a reasonably accurate forest type classification approach. Out of the two classification algorithms, the k-NN classification technique showed higher overall accuracy than the CART method. The approach combining image segmentation and machine learning method can be applied for mapping the forest types in other Terai districts and potentially in other areas as well. More detailed classification can be potentially obtained through inclusion of adequate number of samples in more classes and also inclusion of smaller patches of forests by adjusting the sampling approach so that they are included in the training and test samples.

References

Addink, E. A., de Jong, S. M. and Pebesma, E. J. 2007. The importance of scale in object-based mapping of vegetation parameters with hyperspectral imagery. *Photogrammetric Engineering and Remote Sensing* **72 (8):** 905–912.

Benz, U. C., Hofmann, P., Willhauck, G., Lingenfelder, I. and Heynen, M. 2004. Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS-ready information. *ISPRS Journal of Photogrammetry and Remote Sensing* **58:** 239–258.

Blaschke, T. 2003. Object-based contextual image classification built on image segmentation, *Proceedings of the 2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data*, 27–28 October, Washington D.C., USA, 113–119.

Brandtberg, T. and Warner, T. 2006. High resolution remote sensing. In *Computer Applications in Sustainable Forest Management* (eds.) G. Shao and K. M. Reynolds, Springer-Verlag, Dordrecht, Netherlands, 19–41.

Coburn, C. A. and Roberts, A. C. B. 2004. A multiscale texture analysis procedure for improved forest stand classification. *International Journal of Remote Sensing* **25 (2):** 4287–4308.

Czaplewski, R. L. and Patterson, P. L. 2003. Classification accuracy for stratification with remotely sensed data. *Forest Science* **49 (3):** 402–408.

Definiens, 2004. **eCognition User Guide 4,** Definiens AG, Germany.

DFRS. 2015. **State of Nepal's Forests.** Department of Forest Research and Survey (DFRS), Kathmandu, Nepal.

DFRS/FRA. 2014. **Standard Guidelines for Forest Cover and Forest Types Mapping.**

Technical Report No. 3. Department of Forest Research and Survey, Forest Resource Assessment Nepal Project, Kathmandu, Nepal.

Dorren, L. K. A., Maier, B. and Seijmonsbergen, A. C. 2003. Improved Landsat-based forest mapping in steep mountainous terrain using object-based classification. *Forest Ecology and Management* **183:** 31–46.

Ehlers, M., Gaehler, M. and Janowsky, R. 2003. Automated analysis of ultra high resolution remote sensing data for biotope type mapping: New possibilities and challenges. *ISPRS Journal of Photogrammetry and Remote Sensing* **57:** 315–326.

Ehlers, M. 2004. Remote sensing for GIS applications: New sensors and analysis methods. In *Remote Sensing for Environmental Monitoring, GIS Applications, and Geology III* (eds) Ehlers, M., Kaufmann, J. J. and Michel, U. Proceedings of SPIE, Bellingham, Washington, USA, 1–13.

Ferro, C. J. S. and Warner, T. A. 2002. Scale and texture in digital image classification. *Photogrammetric Engineering and Remote Sensing* **68 (1):** 51–63.

Gong, P., Marceau, D. J. and Howarth, P. J. 1992. A comparision of spatial feature extraction algorithms for land-use classification with SPOT HRV data. *Remote Sensing* **40:** 137–151.

Guo, Q., Kelly, M., Gong, P. and Liu, D. 2007. An object-based classification approach in mapping tree mortality using high spatial resolution imagery. *GIS science and Remote Sensing* **44 (1):** 24–47.

Harvey, K. R. and Hill, G. J. E. 2001. Vegetation mapping of a tropical freshwater swamp in the Northern Territory, Australia: A comparison of aerial photography, Landsat TM and SPOT satellite imagery. *International Journal of Remote Sensing* **22 (15):** 2911–2925.

Kim, M., Madden, M. and Warner, T. A. 2008. Estimation of optimal image object size for the segmentation of forest stands with multispectral Ikonos imagery, *Object-based Image Analysis - Spatial Concepts for Knowledge- driven Remote Sensing Applications* (eds) Blaschke, T., Lang, S. and Hay, G. J. Springer-Verlag, Berlin, 291–307.

Lu, D. and Weng, Q. 2007. Survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing* **28 (5):** 823–870.

Maniezzo, V., Morpurgo, R. and Mussi, S. 1993. D-KAT: A Deep Knowledge Acquisition Tool. *Expert Systems* **10 (3):** 157–166.

Meinel, G. and Neubert, M. 2004. A comparison of segmentation programs for high-resolution remote sensing data. Commission VI in Proceeding of XXth International Society for Photogrammetry and Remote Sensing (ISPRS) Congress, 12–23 July, Istanbul, Turkey, unpaginated (CD-ROM).

Metzler, J. W. and Sader, S. A. 2005. Model development and comparison to predict softwood and hardwood per cent cover using high and medium spatial resolution imagery. *International Journal of Remote Sensing* **26 (17):** 3749–3761.

Qian, Y., Zhou, W., Yan, J., Li, W. and Han, L. 2015. Comparing machine learning classifiers for object-based land cover classification using very high resolution imagery. *Remote Sensing* **7:** 153–168.

Sandmann, H. and Lertzman, K. P. 2003. Combining high-resolution aerial photography with gradient-directed transects to guide field sampling and forest mapping in mountainous terrain. *Forest Science* **49 (3):** 429–443.

Souza, C. M. and Roberts, D. 2005. Mapping forest degradation in the Amazon region with IKONOS images. *International Journal of Remote Sensing* **26 (3):** 425–429.

Townshend, J. R. G., Huang, C., Kalluri, S. N. V., Defries, R. S., Liang, S. and Yang, K. 2000. Beware of per-pixel characterization of land-cover. *International Journal of Remote Sensing* **21 (4):** 839–843.

USGS. 2015. NASA EOSDIS Land Processes DAAC, USGS Earth Resources Observation and Science (EROS) Center, Sioux Falls, South Dakota. http://www.gdem.aster.ersdac.or.jp/ accessed on 1 May, 2015.

Wang, L., Sousa, W. P. and Gong, P. 2004a. Integration of object-based and pixel-based classification for mapping mangroves with Ikonos imagery. *International Journal of Remote Sensing* **25 (24):** 5655–5668.

Wulder, M. A., White, J. C., Niemann, K. O. and Nelson, T. 2004. Comparison of airborne and satellite high spatial resolution data for the identification of individual trees with local maxima filtering. *International Journal of Remote Sensing* **25 (11):** 2225–2232.

Yu, Q., Gong, P. N., Clinton, G., Biging, M. K. and Shirokauer, D. 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing* **72 (7):** 799–811.