

Research Article

New Fusion Algorithm for Improving Secondary Forest Cover Mapping

Jwan Al-doski, Shattri B. Mansor, Zailani Khuzaimah

Department of Civil Engineering, Faculty of Engineering, Universiti Putra Malaysia 43400, Serdang, Selangor, Malaysia

Correspondence should be addressed to Jwan Al-doski, Aldoski-Jwan@hotmail.com

Publication Date: 22 April 2020

DOI: <https://doi.org/10.23953/cloud.ijarsg.456>

Copyright © 2020 Jwan Al-doski, Shattri B. Mansor, Zailani Khuzaimah. This is an open access article distributed under the **Creative Commons Attribution License**, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract Secondary forests play a central role in recovering earlier lost carbon and biodiversity via deforestation and degradation, yet little data is applicable to the magnitude of numerous succession phases. Such information is considered a priority in tropical regions with elevated past and current disturbance rates; however, regrowth in the area is rapid. Focusing on Kuala Krai district, Kelantan state, Malaysia, this paper offer a new fusion algorithm by using the clustering method (fuzzy k-means (FKM)) and Vector Supporting Machines (SVM) procedures. The methodology scheme applied was split into two phases, a clustering map firstly was acquired using FKM from the Sentinel-2A MSI (10 m) image; at the same time, the initial image used to extract Green Normalized Vegetation Index (GNDVI) layer. Using SVM classifier, the classification map was created. Second, SVM and FKM fusion as a hybrid classifier were tested, verified and compared to MLC-parametric and SVM-nonparametric classification algorithms. The study results reveal the effectiveness of the GNDVI layer and FKM segmentation map to enhance SVM classification through applying the Sentinel-2A MS image by approximately 8 % and 14 %, respectively, as opposed to SVM and MLC. Thus this study is inspiring as it is extremely difficult to generate a reliably map land cover in heterogeneous areas, especially in tropical areas, and yet this job is crucial for conservation projects, climate change mitigation strategies, and expansion plans and regional development policies.

Keywords *Secondary Forest Mapping; Vector Supporting Machines; Fuzzy K-Means; Sentinel-2A MS Image*

1. Introduction

Whereas tropical deforestation and degradation, always correlated with forest fragmentation and biodiversity loss, proceed at high levels, certain tropical regions are undergoing secondary forest expansion that performs a minimizing roles in the disappearance of species (Hansen et al., 2013; Haro-Carrión and Southworth, 2018). These complicated forest dynamics cover regions and area changes, and related environmental consequences become even more complicated when agricultural land modifications, including forest plantations, and other procedures of land change are integrated (Haro-Carrión and Southworth, 2018; Lambin, Geist and Lepers, 2003). Detailed forest cover regions and their changes using remote sensing methodologies are still required to strengthen our understanding of the true effects of human operations in certain tropical areas (Lambin, Geist and

Lepers, 2003; Frédéric Achard et al., 2010). In humid tropics like Malaysia, mapping forest cover area and changes utilizing remote sensing technologies have been successfully predicted (Maxwell, Warner, and Fang, 2018). Distinct satellite images were applied and examined using various techniques of image classification to map secondary forest land cover in the tropics (Frédéric Achard et al., 2010; Gibbs et al., 2010). The most prevalent parametric land algorithms involve maximum likelihood and linear discriminant analysis (Lu and Weng, 2007; Yonezawa, 2007). In regard to classify satellite data (single or multi-date images) of heterogeneous land cover spectral signatures, non-parametric algorithms such as Artificial Neural Networks (ANN), Decision Tree Classifiers (DTC) and Vector Supporting Machines (SVM) have demonstrated enhanced efficiency over more traditional classification methods (Lu and Weng, 2007; Immitzer, Atzberger, and Koukal, 2012).

However, Machine learning algorithms have been commonly used over the previous centuries as classification algorithms and maybe some evaluations of their comparative results compared to other classifiers have been carried out in the tropical region (Carreiras, Pereira, and Shimabukuro, 2006). Among these algorithms, SVMs have proved their classification accuracy. Precisely, elevated accuracies in land cover mapping and outperforming other algorithms have been demonstrated by SVMs (Schulz, Hänsch, and Sörgel, 2011). The accomplishment of SVM is linked to SVM classifier's inherent characteristics, which can manage unplaced issues, and to the dimensionality curse, this offers strong, scarce solutions and defines non-linear limits between classifications of land cover decisions (Hughes, 1968). It aims to distinguish secondary forest class from other land cover classifications by detecting a plane in a multidimensional characteristic room that optimizes its separation, rather than using stats to define such classes (Awad and Khanna, 2015). SVM classifiers do not need decent training sets, but only samples of training. Foody and Mathur (2006) stated use of small training sets of intentionally chosen mixed pixels with support vectors, as this technique does not lose classification accuracy and would save considerable time (Foody and Mathur, 2006).

The other very well-developed classification algorithm is Fuzzy k-means (FKM) clustering algorithm is a technique of multivariate units in different research of vegetation, soil and forestry (Tapia, Stein, and Bijker, 2005; He, T. et al., 2014]. The algorithm FKM was used mainly to resolve the class overlap concern, but its viability can be reduced if the data sets are massive (He, T. et al., 2014; Shaikh and Patil, 2017]. The fundamental issue for increasing the classification accuracy of SFC mapping is the appropriate selection of algorithms as suggested by some researchers. For example, (Nguyen and Pham, 2016] incorporated a NDVI and DEM with Landsat 8 image spectral bands to reduce the influence of shadows on image classification, differentiate between natural and planted forests, and generate a forest inventory support LCM from Hoa Binh Province. A precision classification was conducted on a multi-source dataset (bands 1–7, and 9, NDVI, and DEM) compared to spectral picture outcomes. Generally speaking, general precision increased by 5.23% (from 84.51 to 89.74%) The multisource classification with SVM was used by (Watanachaturaporn, Arora and Varshney, 2008) distinct textural measures are a significant source of ancillary data and their advantages for classifying land cover mapping were outlined in research using various techniques and classifiers.

We implemented a combination approach to deal with Secondary Forest Cover (SFC) mapping in Sentinel-2A MSI image in order to use certain the advantages of SVM and FKM clustering. The SVM classifier has been shown to produce a spectral classification map, while the FKM clustering algorithm has been implemented to get a segmentation map ensemble. SVM and FKM algorithm fusion proposes to minimize class sorting problems by achieving the feature vector and realizing the optimal nonlinear classification boundaries with SVM.

2. Develop New Fusion Algorithm

A parametric MLC, a non-parametric SVM as well as a hybrid as integration of unsupervised, SVM and FKM fusion classifiers were used to compare different classifiers. Since extensive definitions have been

stated by many researchers (Awad and Khanna, 2015; Deilmay, Ahmad and Zabihi, 2014; Zhang, Ren and Jiang, 2015), we will not explain how the algorithms MLC and SVM operate here.

2.1. Gaussian- RBF Based SVM Classifier

After defining the Sentinel-2A MSI data set which is used for SFC mapping, for just the supervised classification stage, a robust classifier should be selected (He, T. et al., 2014). The linear SVM classifier is selected because of its inherent robustness to high-dimensional information sets and unplaced issues. The initial suggested SVM algorithm by (Vapnik and Lerner, 1963). The SVM's fundamental concept is to map multidimensional information into such a higher-dimensional space where there is a hyperplane that can be used to linearly distinguish the initial information, maximizing the margin between distinct classifications (Awad and Khanna, 2015; Zhang H., et al., 2015) and (Boser, Guyon and Vapnik, 1992) suggested a technique by introducing the kernel trick to the maximum-margin hyperplane to generate nonlinear classifiers. The classifier seeks to establish a separation rule linear type among instances caused by a higher-dimensional sample space mapping feature (Nisbet, Elder and Miner, 2009)

A linear separation in a certain space corresponds in the initial input space to a nonlinear separation. An instance is shown in Figure 1. The kernel key is the foundation of such an algorithm: since mapped tests tend only in the form of dot items in the SVM formulation, such procedures can be replaced by valid kernel functions $K(.,.)$, going back specifically to the internal product value in that space, Eq. (1). The solution is provided by the maximum margin width hyperplane, which ensures the highest generalization capability on earlier unknown data. It is necessary to optimize the dual optimized formulation.

$$\max_{\alpha} \sum_{i=1}^N \alpha_i - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j \omega_i \omega_j k(x_i, x_j), \quad s.t. 0 \leq \alpha_i \leq c \text{ and } \sum_{i=1}^N \alpha_i \omega_i = 0, \quad (1)$$

Where a user-identified parameter is C, that regulates the trade-off between the difficulty of the system and the training error, the coefficients outlining are α_i the key of the optimization and (binary case) $\omega_i \in \{+1; -1\}$ are sample-related class labels x_i . When the answer to Eq. (2) is discovered, the unknown sample label x' is provided sign of decision function, i.e. its location on the separating hyperplane

$$\omega' = \text{sign}[\sum_{i=1}^N \alpha_i \omega_i k(x_i, x') + b] \quad (2)$$

Experiments are performed in this study using a kernel Gaussian -RBF:

$$k(x_i, x_j) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$$

Where σ is the user-identified bandwidth of the Gaussian function. Typically The Gaussian RBF is used in several environmental and LC mapping applications for its computational complexity. The one-against-all system is implemented to fix multi-class issues

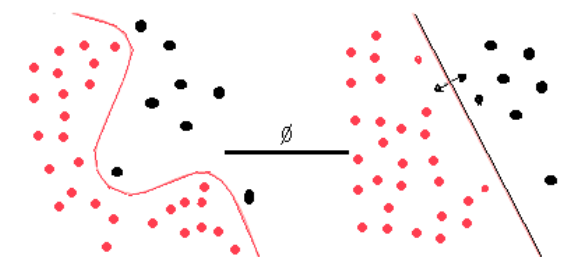


Figure 1: SVM- Kernel Machine

2.2. FKM Clustering for Forest Cover Segmentation

A fuzzy segmentation is introduced to pre-classify land cover for mapping. There is plenty of motivation for this decision. First, no set objects can be acknowledged as the concept for ground coverings was intrinsically vague, and therefore no valid, quantitative statuses are available. Second, many units between limits are intersected (He, T. et al., 2014). In an FKM cluster, the degree to which an object belongs to all contestant classes requires to retain a record (He, T. et al., 2014). Exactly, a true figure is classified in the range for all items [0, 1] considered a membership value [referred to as $\mu(Xc)$] is enrolled in all c classes where a value of $\mu(Xc)=0$ clarifies that the object does not reside to the class or set, Xc , and $\mu(Xc)= 1$ Specifies that it belongs completely to the set or class, Xc , and thus can be considered a replica of the set. Values among $\mu (Xc) = 0$ and $\mu (Xc) = 1$ stipulate the comparative power of the degree to which product has group Xc 's classic elements. Consequently, the outcome of FKM clustering is a register of the degree to which the object belongs to each class becoming considered each object getting evaluated. The FKM clustering algorithm is used for four sentinel 2A (10 m) image pixel values. This method offers a range of units recognized by the largest subscription value class, based on the degree of fuzziness provided by the parameter fuzziness φ and the number of land cover classes (k). In this research, sighted N data, φ , and the highest partition coefficient F will be used for k [Eq. (3)] and H (entropy parameter) [Eq. (4)]

$$F = \frac{F' - \frac{1}{k}}{1 - \frac{1}{k}}, \text{ s. t. } F' = \frac{1}{N} \sum_{i=1}^N \sum_{c=1}^k (m_{ic})^2 \quad (3)$$

$$H = \frac{H' - 1 + F}{\log k - 1 + F}, \text{ s. t. } H' = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^k m_{ic} \log(m_{ic}) \quad (4)$$

m_{ic} is pixel i to category c , $c = 1, \dots, k$ Both F' and H' depends entirely on the number of categories k . In the fuzzy method, categorization for a range of class numbers and parameters was reiterated for the (k) and φ . Sentinel 2A image (10 m), we tried k 2 to 15 and reached the greatest accuracy when $k= 4$ (Figure 2). As shown in the observation of different writers, the φ was set to 2.0.

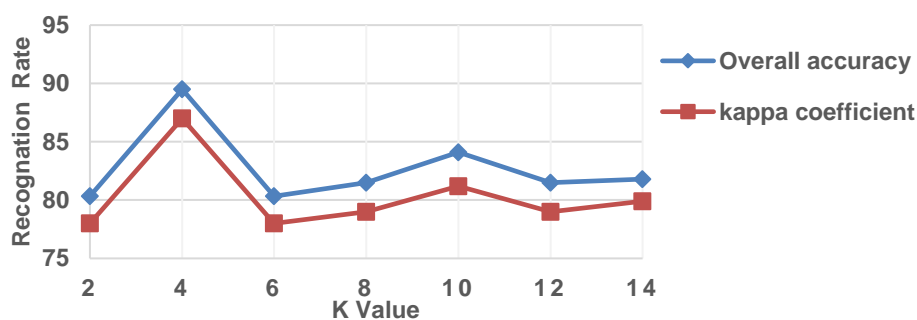


Figure 2: Choose the optimal k for FKM

2.3. Green Normalized Difference Vegetation Index (GNDVI)

In addition to the choice of image classifiers, it is acknowledged that the use of ancillary data is essential for image classification performance. Ancillary data were effectively used to enhance the classification of images, in particular by including topographical measurements (elevation and slope), NDVI, GNDVI and texture measurements in the image classification manner in addition to spectral information to separate characteristics with comparable spectral features (Coburn and Roberts, 2004) GNDVI has become a standard land cover remotely sensed product (Xue and Su, B., 2017; Dzieszko, M., Dzieszko, P., and Królewicz, 2012) for discrimination and interpretation of mapped vegetation units it has been commonly used see Figures 5 (b) and 6 (b). GNDVI has been calculated from (Gitelson, Kaufman and Merzlyak, 1996).

$$\text{GNDVI} = \text{NIR} - \text{G} / \text{NIR} + \text{G} \quad (5)$$

Where NIR and G were used for near-infrared and green data bands.

2.4. Classification Architectures of SVM and FKM Fusion

An appropriate technique must be identified to take benefit of the FKM and SVM algorithms. The architectures of the classification are provided: (i) clustering of FKM and (ii) classification of SVM. Figure 3 shows the main scheme. Using FKM clustering algorithm, the Sentinel-2A MSI image is classified and clustering maps are produced. At the same time, the initial image extracts GNDVI layer. The original image is added to both the clustering map and the GNDVI layer. The SVM classifier is then used for classification. SFC map is finally acquired.

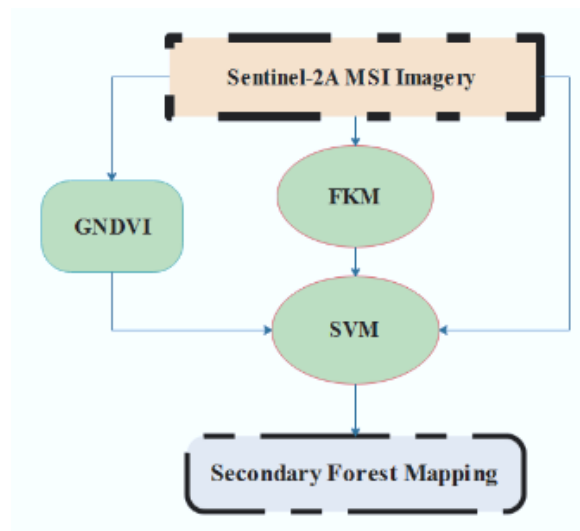


Figure 3: The main scheme for a new fusion technique (SVM and FKM algorithms)

3. Material

3.1. Study Area

The test regions selected is the district Kuala Krai, Kelantan state, Malaysia. The co-ordinates are between longitudes 102.280105°E and 102.229727°E and latitudes 6.202459°N and 5.882407°N (Figure 4). The annual average rainfall is over 6,000 mm. The temperature is 27.5 C0 per month. The rainy, equatorial climate suits oil palm plantations, making oil palm one of the region's largest cash crops. Test research elevation progressively raises from 400 -900 m above sea level so our research region is comparatively flat.



Figure 4: Location of Kuala Krai, Malaysia

3.2. Sentinel-2A Data and Pre-processing

Sentinel-2A is launched in June 2015; Sentinels Data can be downloaded from (<https://scihub.copernicus.eu/>) for free. The Sentinel-2A MS data is distinguished by a 13 spectral bands with a ground spatial resolution between 10 and 60 m (Table 1). Sentinel-2A MS imagery acquired on 3 July 2017 at Level L1C geocoded (solar azimuth 1270, solar elevation 660) with standard product of TOA reflectance and then was pre-processed. Firstly, the image atmospherically corrected with sen2cor software version 2.3 (Telespazio VEGA Deutschland GmbH, Darmstadt, Germany) to generate and BOA reflectance corrected Level-2A bands on the Anaconda Python platform using the Sen2Cor processor (version 2.3) (Jérôme et al., 2016). The format of the output product is a compilation of TIFF images with three distinct resolutions (10, 20 and 60 m) reproduced bands. In this research, we used 10-meter bands to obtain SFC map. The Sentinel-2 MSI image was then geometrically corrected (Lima et al., 2019; Baillarin et al., 2012) using 15 GCPs from main features like roads and DEM to attain enhanced geodetic accuracy (Habib et al., 2017). The first-order polynomial function and the nearest protocol was enacted to correct systematic variations between neighboring images in a few instances. The total transformation RMSE equal 0.07 which was less than a pixel was attained. Then the Sentinel-2 MS image re-projected to UTM coordinating scheme, WGS 1984 datum and 47 south zone using the nearest neighbours resampling method. The data were spatially subset with ENVI 5.1 software to the study area.

Table 1: Sentinel-2 MSI spectral bands information

Sentinel-2 Bands	Central Wavelength (μm)	Resolution (m)	Bandwidth (nm)
Band 1—Coastal aerosol	0.443	60	20
Band 2—Blue	0.490	10	65
Band 3—Green	0.560	10	35
Band 4—Red	0.665	10	30
Band 5—Vegetation Red Edge	0.705	20	15
Band 6—Vegetation Red Edge	0.740	20	15
Band 7—Vegetation Red Edge	0.783	20	20
Band 8—NIR	0.842	10	115
Band 8A—Narrow NIR	0.865	20	20
Band 9—Water Vapor	0.945	60	20
Band 10—SWIR—Cirrus	1.380	60	30
Band 11—SWIR	1.610	20	90
Band 12—SWIR 20	2.190	20	180

To decrease complexity calculation and enhance classification accuracy, two sample images were clipped after topographical correction by DEM, see (Figure. 5(a)) for sample 1 image, and (Figure 6(a)) for sample 2. By visual inspection, image interpretation in both sample images emphasized a total of five LC classes of interested areas. Finally, sample 1 image labeled 3367 pixels (Figure 5(b)) and sample 2 image labeled 2919 pixels (Figure 6(b)). Water bodies, primary forests, secondary forests, urban areas and other classes were the sort of land cover. Careful attention has been given to scatter training areas across each image to ensure that they are indicative of the whole image and to collect as many samples of practice for each land cover class (Table 2) as proposed criteria have been met to determine the appropriate minimum sample size. The Jeffries – Matusita transformed divergence index was used to evaluate the separability of sample information. This study separability confirmed, it was rather high for water bodies, urban regions, and other regions, but for primary forest and secondary forest classes, it is much smaller. All of these pixels were used to train and validate classifiers for SVM supervision.

Table 2: Size of land cover samples (# pixels) collected for develop new fusion approach

No	Land Cover Classes	Sample pixels	Sample 1	Sample 2
1	Water Bodies	398	242	
2	Primary Forest	568	460	
3	Secondary Forest	784	860	
4	Urban Area	825	774	
5	Other	792	583	
Total	3367		2919	

3.3. Experimental Setup

The MLC, SVM, SVM and FKM fusion classifiers were used to explore different types of algorithms. All algorithms have been initiated in Windows 10 using ENVI+IDL 5.1. The combination of both algorithms was mainly split into two phases in this research. First, by means of Eq (5), the GNDVI layer was calculated utilizing Sentinel-2A MS red and near-infrared bands; the clustering FKM algorithm was applied to generate a division map with 10 m imagery across all four bands. The GNDVI layer and division map were then piled onto the original MS image of the Sentinel-2A. Second, in order to calculate and generate the SFC map, the SVM classifier was lastly set up. A 3 x 3 pixel majority filter was introduced after map manufacturing to minimize salt and pepper noise to improve accuracy. Reference data collection for precision assessment was based on stratified random sample selection, with sample units being collected at a minimum of 1 m to avoid future effects of spatial autocorrelation.

From the pictures themselves, the information was field-truthed by specialist understanding. For the aggregate and accuracy assessment obtained by each class, a confusion matrix was developed, and this is the most common method for remote sensing classification accuracy assessment.

4. Results and Discussions

Figures 5 and 6 demonstrate the MLC, SVM and FKM and SVM classifier fusion classification maps. In Figure 5, all classification techniques recognized secondary class forest as the land cover class taking over half of the entire area of the region, accompanied by primary class forest. Entirely techniques recognized classes of urban areas and water bodies as smallest region of the land cover class. Other class, on the contrary, represented the biggest percentage in Figure 6.

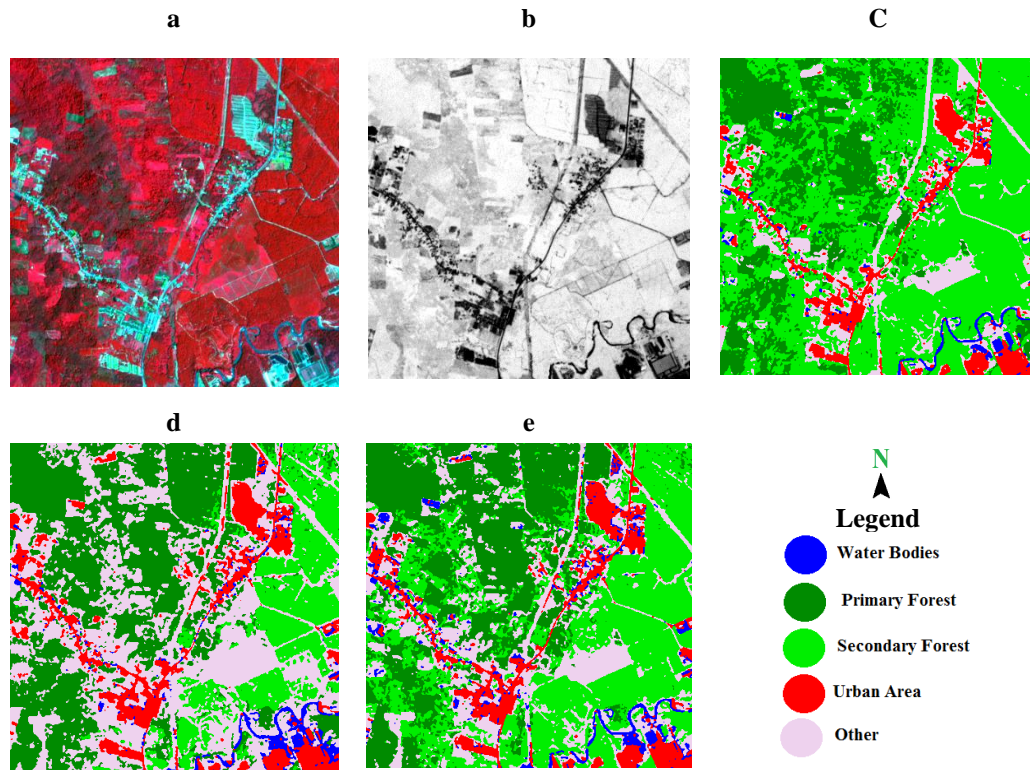


Figure 5: Multispectral high resolution Sentinel-2A MS image acquired over Kota Bharu city, Kelantan state, Malaysia (sample 1 image). (a) RGB image structure (b) GNDVI (c) MLC map (d) SVM map (e) SVM and FKM Fusion map.

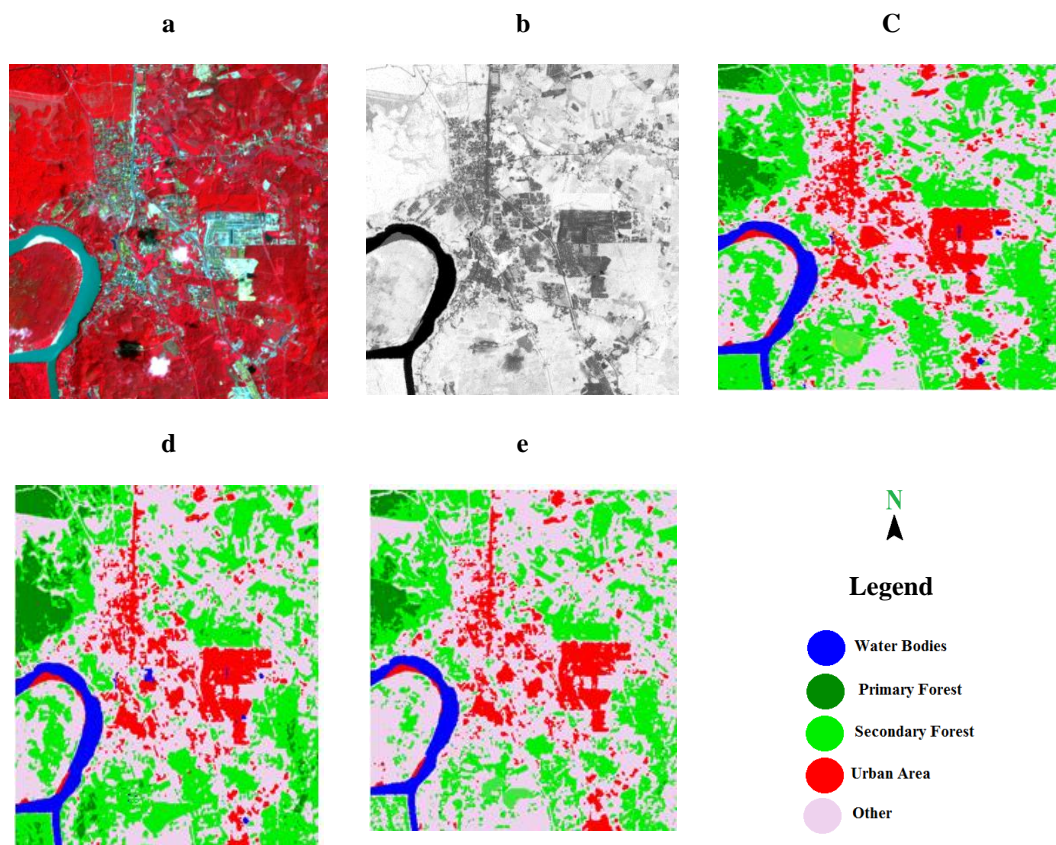


Figure 6: Multispectral high resolution Sentinel-2A MS image acquired over Kota Bharu city, Kelantan state, Malaysia (sample 2 image). (a) RGB image structure (b) GNDVI (c) MLC map (d) SVM map (e) SVM and FKM Fusion map

To evaluate the separation efficiency of classes, confusion matrices were generated per each classification algorithm. In the sample 1 image, each algorithm with OV evaluated 78.05%, 86%, and 93.8%, respectively, of MLC, SVM, and SVM and FKM fusion (Table 3). Whereas for sample 2 image, the MLC, SVM, FKM and SVM fusion classifiers OA was 85.2%, 90.4%, and 98.6% (Table 4). From both tables, the MLC technique developed accuracy for the individual classes of the smallest producer and user. The most details were on the MLC map, while the least details were on the SVM classification map. It is due to the subsequent translation of the SVM algorithm into a convex issue of optimization that can ensure the ideal global. MLC is focused, however, on addressing the resident issue and guaranteeing the local optimal. Tables 3 and 4 show that in SFC maps from SVM classifier is more profitable than MLC. It also coincides with the fact that in SFC maps, the SVM algorithm is better than MLC which was denoted from (Zhang, Ren and Jiang, 2015; Mondal et al., 2012). Amongst other classifiers, the fusion of the SVM and FKM algorithms had the largest OA. In this research, the greatest OA accuracy produced by FKM and SVM fusion suggests that our strategy is beneficial in carrying out SFC maps. The outcome was significantly affected by the samples of training although there were some shadows in urban area and samples of water bodies training. In splitting secondary forest and primary forest classes, the suggested technique was less efficient. It might be due to the Sentinel-2A MS image date.

Table 3: Sample 1 image land cover accuracy % of MLC, SVM and fusion classifiers

Land Cover Classes	MLC		SVM		SVM and FKM	
	PA%	UA%	PA%	UA%	PA%	UA%
Water Bodies	73.68	100	97.22	100	94	98.95
Primary Forest	72.50	94.57	95.33	96.62	94	94
Secondary Forest	65	71.23	67	83.87	97	92.38
Urban Area	86.59	87.57	86.9	94.51	93	92.08
Other	90.41	58.67	91.2	65.35	91	91.92
OA%	78.05		86		93.8	
K Coefficient	0.72		0.82		0.92	

PA, Producer accuracy; UA, User accuracy

Table 4: Sample 2 image land cover accuracy % of MLC, SVM and fusion classifiers

Land Cover Classes	MLC		SVM		SVM and FKM	
	PA%	UA%	PA%	UA%	PA%	UA%
Water Bodies	81	100	92	98.92	98	100
Primary Forest	65	100	98	72.59	96	98.97
Secondary Forest	100	64.52	63	96.92	99	100
Urban Area	100	100	100	100	100	100
Other	100	84.3	100	92.52	100	94.34
OA%	85.2		90.4		98.6	
K Coefficient	0.81		0.88		0.98	

PA, Producer accuracy; UA, User accuracy

The FKM and SVM algorithms fusion for sample 1 image confusion matrix is shown in Table 5. Despite the extremely accurate general outcomes achieved by the fusion of SVM and FKM algorithms, it was significantly less efficient in acknowledging secondary forest, urban area, and others. Approximately 3% of the secondary forest was confused as the primary forest, whilst also 6% of the others class was wrongly marked as the class of the urban region. Table 6 also demonstrates the sample area 2 mage confusion matrix of FKM and SVM algorithms fusion. In highlighting primary forest, others, and water bodies, the algorithm was less efficient. Approximately 4% of primary forest was confused as others and 2% of others were wrongly marked as the class of water bodies.

Table 5: Confusion matrixes representing the highest overall classification using SVM and FKM fusion in sample

1

Land Cover Classes	Ground truth (%)					
	Water Bodies	Primary Forest	Secondary Forest	Urban Area	Other	Total
Water Bodies	94	0	0	1	0	19
Primary Forest	0	94	3	0	3	20
Secondary Forest	0	3	97	0	5	21
Urban Area	6	1	0	93	1	20.20
Other	0	2	0	6	91	19.80
Total	100	100	100	100	100	

The size of surface objects relative to a sensor's spatial resolution is strongly associated with the divergence of the image. There have been some errors between primary and secondary forest classes. Approximately 3% of primary forest category pixels were regarded as secondary forest classes and around 5% of secondary forests were incorrectly categorized as other classes (Table 6), only 2% of other classes were erroneously categorized as primary forests and 6% were confused as urban areas, which can also be caused by the urban area training study. The reasons are as follows for the bad decision that distinguishes urban, other, and water bodies. First, the urban area on the

Sentinel-2A MS image was smaller than the other land cover classes, and some trees, grasses, and bare surface were included in urban area class sample. Second, the urban area and other class had comparable houses. It can also be discovered that factories were constructed on hills and that the urban area was positioned close to leakage from the classification map resulting from the policy of local land use.

Table 6: Confusion matrixes representing the highest overall classification using SVM and FKM fusion in sample 2

Land Cover Classes	Ground truth (%)					Total
	Water Bodies	Primary Forest	Secondary Forest	Urban Area	Other	
Water Bodies	98	0	0	0	0	19.60
Primary Forest	0	96	1	0	0	19.40
Secondary Forest	0	0	99	0	0	19.80
Urban Area	0	0	0	100	0	20
Other	2	4	0	0	100	21.20
Total	100	100	100	100	100	

5. Summary

This paper suggested a combination of classification techniques for SVM and FKM. When dealing with RS images as data sets for land cover mapping, this technique can enhance effectiveness the accuracy. The effectiveness of the GNDVI layer and FKM segmentation map has been illustrated in this studies to improve SVM classification in the Sentinel-2A MS image by approximately 8% compared to SVM and 14% compared to MLC. Experiments on the issue of classification of the Sentinel-2A MSI image showed excellent outcomes and inspired future and in-depth studies on the classification of land cover. This is the first time that we have infused SVM and FKM algorithms to classify SFC. Attention to greater resolution images and combine more data is the most important asset. Our findings are encouraging since it is highly hard to reliably map land cover in heterogeneous fields, particularly in tropical regions, and yet this task is essential for preservation initiatives, climate change mitigation strategies, and the design of development plans and rural development strategies. Our current strategy to fusion offers the benefit of being simple to enforce, as both GNDVI measurement and SVM classifier existence are readily available and price-effective in common and affordable remote sensing software, as SVM classifiers can use bigger training databases without compromising the precision of classification. Importantly, the very accurate findings gained from this method indicate its excellent potential for land cover mapping in tropical regions.

Acknowledgements

Authors wish to gratefully acknowledge Prof. Shattri B. Mansor, Director of GIS Research Centre at Universiti Putra Malaysia, and Dr. Zailani Khuzaimah for the financial support, recommendations and guidance through this study. The authors wish to acknowledge the helpful comments and advise from both anonymous reviewers.

References

- Awad, M. and Khanna, R. Support Vector Machines for Classification in *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers* Berkeley, CA: Apress, 2015, pp.39-66.
- Baillarin, S. J., Meygret, C., Dechoz, B., Petrucci, S., Lacherade, Tremas, T., Isola, C., Martimort, P., Spoto, F. Sentinel-2 level 1 products and image processing performances, in: *2012 IEEE international geoscience and remote sensing symposium*, 2012, pp.7003-7006: IEEE.

Boser, B. E., Guyon, I. M., and Vapnik, V. N. A training algorithm for optimal margin classifiers, in: *Proceedings of the fifth annual workshop on Computational learning theory*, 1992, pp.144-152: ACM.

Carreiras, J., Pereira, J., and Shimabukuro, Y. E. 2006. Land-cover mapping in the Brazilian Amazon using SPOT-4 vegetation data and machine learning classification methods. *Photogrammetric Engineering & Remote Sensing*, 72(8), pp.897-910.

Coburn, C., and Roberts, A. C. 2004. A multiscale texture analysis procedure for improved forest stand classification. *International journal of remote sensing*, 25(20), pp.4287-4308.

Deilmai, B. R., Ahmad, B. B., and Zabihi, H. Comparison of two classification methods (MLC and SVM) to extract land use and land cover in Johor Malaysia in *IOP Conference Series: Earth and Environmental Science*, 2014, 20(1), p.012052: IOP Publishing.

Dzieszko, M., Dzieszko, P., and Królewicz, S. 2012. Digital aerial images land cover classification based on vegetation indices. *Quaestiones Geographicae*, 31(3), pp.5-23.

Foody, G.M. and Mathur, A. 2006. The use of small training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by a SVM. *Remote Sensing of Environment*, 103(2), pp.179-189.

Frédéric Achard, Hans-Jürgen Stibig, Hugh D. Eva, Erik J. Lindquist, Alexandre Bouvet, Olivier Arino and Philippe Mayaux. 2010. Estimating tropical deforestation from Earth observation data. *Carbon Management*, 1(2), pp.271-287.

Gibbs, H. K., Ruesch, A. S., Achard, F., Clayton, M. K., Holmgren, P., Ramankutty, N. and Foley, J. A. 2010. Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. *Proceedings of the National Academy of Sciences*, 107(38), pp.16732-16737.

Gitelson, A. A., Kaufman, Y. J., and Merzlyak, M. N. 1996. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote sensing of Environment*, 58(3), pp.289-298.

Habib, A., Akdim, N., Labbassi, K., Khoshelham, K., and Menenti, M. 2017. Extraction and accuracy assessment of high-resolution DEM and derived orthoimages from ALOS-PRISM data over Sahel-Doukkala (Morocco). *Earth Science Informatics*, 10,(2), pp.197-217.

Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., Townshend, J. R. G. 2013. High-resolution global maps of 21st-century forest cover change. *Science*, 342 (6160), pp.850-853.

Haro-Carrión, X. and Southworth, J. 2018. Understanding Land Cover Change in a Fragmented Forest Landscape in a Biodiversity Hotspot of Coastal Ecuador. *Remote Sensing*, 10(12), p.1980.

He, T., Sun, Y.-J., Xu, J.-D., Wang, X.-J., and Hu, C.-R. 2014. Enhanced land use/cover classification using support vector machines and fuzzy k-means clustering algorithms. *Journal of Applied Remote Sensing*, 8(1), p.083636.

Hughes, G. 1968. On the mean accuracy of statistical pattern recognizers. *IEEE transactions on information theory*, 14(1), pp.55-63.

Immitzer, M., Atzberger, C. and Koukal, T. 2012. Tree species classification with random forest using very high spatial resolution 8-band WorldView-2 satellite data. *Remote Sensing*, 4(9), pp.2661-2693.

Jérôme Louis, Vincent Debaecker, Bringfried Pflug, Magdalena Main-Knorn, Jakub Bieniarz, Uwe Mueller-Wilm, Enrico Cadau, Ferran Gascon. Sentinel-2 sen2cor: L2a processor for users, in: *Proceedings of the Living Planet Symposium, Prague, Czech Republic*, 2016, pp.9-13.

Lambin, E. F., Geist, H. J. and Lepers, E. 2003. Dynamics of land-use and land-cover change in tropical regions. *Annual review of environment and resources*, 28(1), pp.205-241.

Lima, T. A., Beuchle, R., Langner, A., Grecchi, R. C., Griess, V. C., and Achard, F. 2019. Comparing Sentinel-2 MSI and Landsat 8 OLI Imagery for Monitoring Selective Logging in the Brazilian Amazon. *Remote Sensing*, 11(8), p.961.

Lu, D. and Weng, Q. 2007. A survey of image classification methods and techniques for improving classification performance. *International journal of Remote sensing*, 28(5), pp.823-870.

Maxwell, A. E., Warner, T. A. and Fang, F. 2018. Implementation of machine-learning classification in remote sensing: An applied review. *International Journal of Remote Sensing*, 39(9), pp.2784-2817.

Mondal, A., Kundu, S., Chandniha, S. K., Shukla, R., and Mishra, P. 2012. Comparison of support vector machine and maximum likelihood classification technique using satellite imagery. *International Journal of Remote Sensing and GIS*, 1(2), pp.116-123.

Nguyen, T. T. H., and Pham, T. T. T. 2016. Incorporating ancillary data into Landsat 8 image classification process: a case study in Hoa Binh, Vietnam. *Environmental Earth Sciences*, 75(5), p. 430.

Nisbet, R., Elder, J., and Miner, G. 2009. *Handbook of statistical analysis and data mining applications*. Academic Press.

Schulz, K., Hänsch, R., and Sörgel, U. Machine learning methods for remote sensing applications: an overview in *Earth Resources and Environmental Remote Sensing/GIS Applications IX*, 2018, 10790, p. 1079002: International Society for Optics and Photonics.

Shaikh, A. H. and Patil, M. E. 2017. Enhancing the Performance of K-Means Clustering by using Fuzzy Partitioning Matrix. *International Journal of Computer Applications*, 166(4).

Tapia, R., Stein, A., and Bijker, W. 2005. Optimization of sampling schemes for vegetation mapping using fuzzy classification. *Remote Sensing of Environment*, 99(4), pp.425-433.

Vapnik, V., and Lerner, A. Y. 1963. Recognition of patterns with help of generalized portraits. *Avtomat. i Telemekh*, 24(6), pp.774-780.

Watanachaturaporn, P., Arora, M. K., and Varshney, P. K. 2008. Multisource Classification Using Support Vector Machines. *Photogrammetric Engineering & Remote Sensing*, 74(2), pp.239-246.

Xue, J., and Su, B. 2017. Significant remote sensing vegetation indices: A review of developments and applications. *Journal of Sensors*, 2017.

Yonezawa, C. 2007. Maximum likelihood classification combined with spectral angle mapper algorithm for high resolution satellite imagery. *International Journal of Remote Sensing*, 28(16), pp.3729-3737.

Zhang, H., Weng, Q., Lin, H., and Zhang, Y. 2015. *Remote sensing of impervious surfaces in tropical and subtropical areas*. CRC Press.

Zhang, Y., Ren, J., and Jiang, J. 2015. Combining MLC and SVM classifiers for learning based decision making: Analysis and evaluations. *Computational intelligence and neuroscience*, 2015, p.44.