ICGTD 2020 International Conference on Green Technology and Design

itenas

Investigation of Classification Algorithm for Identification of Oil Palm Plantation Using Multiscatter and Multiresolution SAR Data

Soni Darmawan
Faculty of Civil Engineering and Planning
Institut Teknologi Nasional Bandung
Bandung Indonesia
soni_darmawan@itenas.ac.id

Rika Hernawati
Faculty of Civil Engineering and Planning
Institut Teknologi Nasional Bandung
Bandung Indonesia
rikah@itenas.ac.id

Anwar Ersyad Ananta
Faculty of Civil Engineering and Planning
Institut Teknologi Nasional Bandung
Bandung Indonesia
ersyadananta@gmail.com

***Abstract*— Oil palms are plants with green leaves, it is very difficult to differentiate between oil palms and forests only by color grouping. It is therefore important to recognize and distinguish species of oil palm trees specifically from other tree and forest species for proper planting and management, one of them is by usage of the classification algorithm. The aim of this research is to identify the best classification algorithm for the determination of oil palm plantations using radar satellite imagery. The methodology including primary and secondary data collection, primary data are using Synthetic Aperture Radar (SAR) images including Sentinel-1A, ALOS PALSAR-2 and TerraSAR-X, while secondary data is the oil palm plantation training field. The processing of SAR data was involved radiometric and geometric correction and scattering calibration. The classification technique were using the algorithm of parallelepiped, support vector machine (SVM), and maximum likelihood. From three algorithm of classification. the results showed that the support vector machine (SVM) algorithm was the best accuracy with 86.18 % in Sentinel-1A, 91.17 % in ALOS PALSAR-2, and 77.57 % in TerraSAR-X.**

Keywords— oil palm, classification, Support Vector Machine, Parallelepiped and Maximum Likelihood

I. INTRODUCTION

Oil palm (*Elaeis Guineensis Jack*) is a tropical plant that is widely grown in Asia, Africa and South America. This plant thrives in tropical areas. Oil palm trees have warm temperatures, sunshine, and high rainfall to maximize production [1]. The spread of oil palm in Indonesia dominates the plantation area starting from Aceh, the East Coast of Sumatra, Java, Kalimantan and Sulawesi. Oil palm can generate huge profits so that many old forests and plantations are converted into oil palm plantations [2]. Indonesia is one of the largest exporters and producers of palm oil in the world, reaching 38.17 million tons in 2017, so as the largest producer, good management of oil palm plantation is very important [3], [4]. Apart from production, Indonesia is also at the top in the area of oil palm plantations with an area of 8.9 million ha in the form of productive crops (TM) [5]. The bright prospect of palm oil commodity in the world vegetable oil trade has prompted the Indonesian government to spur the development of oil palm plantation areas in every province in Indonesia [5]. Nowadays, remote sensing technology has been used to carry out observations on oil palm plantations, both for observing the number of trees up to the age of the oil palm. Remote sensing is the science and art of obtaining information about an object, area, or phenomenon through analysis of data obtained with a tool without direct contact with the object, area or phenomenon being studied [6]. In remote sensing, there are 2 types of sensors used, active sensors and passive sensors. An active sensor is a remote sensing system that makes use of the waves transmitted by the sensor which are then recaptured by the sensor after being reflected by an object on the earth's surface. Passive sensors are remote sensing systems that record reflected waves with an energy source that comes from the sun or the object itself [4], [7]–[10]. In this study, a radar image (Radio Detection and Ranging) which has an active sensor will be used.

Previous research on oil palm had been conducted, by [4] were researched on the used of remote sensing imagery ALOS PALSAR 2 to investigate the age and yield of oil palm fresh fruit bunches. Also [11] was conducted research on the used of PALSAR remote sensing imagery to mapped oil palm plantations in Cameroon using a 50 meter resolution image mosaic. Other studies is [12], conducted research on the use of remote sensing imagery. ALOS integrated image data to detect smallholder oil palm plantations in South Sumatra, Indonesia. [13] was conducted research on the use of remote sensing images in the form of SAR data applications for discrimination of oil palm trees [14]. [15] was conducted a study on the use of remote sensing imagery to estimate aboveground biomass and oil palm plantations using ALOS PALSAR data. And [16] was conducted research on the use of remote sensing imagery to map oil palm expansion using SAR to study the impact on the CO2 cycle. However, there is still a lack of research on the algorithm for oil palm area classification using multi-scatter and SAR resolution images, especially in Indonesia, so we try to investigate the classification algorithm for oil palm plantation identification using multi-scatter and multi-resolution SAR images in this study.

II. METODOLOGY

This research methodology is carried out in several stages. First is preparing primary and secondary data such as ALOS PALSAR-2, Sentinel-1A, and TerraSAR-X imagery. Second is preprocessing including geometric and radiometric corrections for each image, take training samples. Third is classifying using the Parallelepiped, Maximum Likelihood, and Support Vector Machine algorithm. And the last step is to find out an accurate assessment of the classification for each algorithm and also to produce land cover mapping for oil palm plantations. The analysis methods can be found in Figure 1.

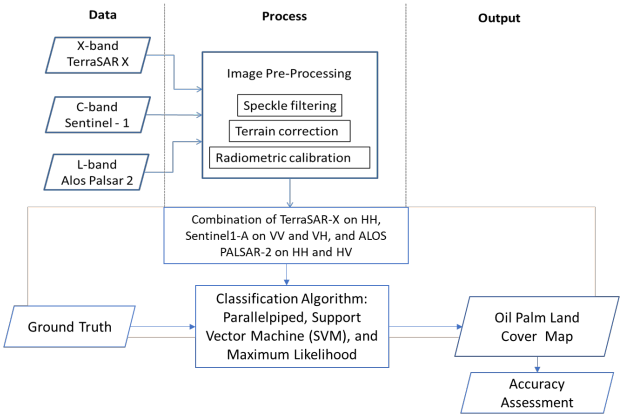

Fig 1. Methodology

A. Study Area

This research area is located in the District of Asahan, North Sumatra. The location is in 2.98 East Longitude; 99.67 North Latitude and 2.92 East Longitude; 99.75 North Latitude (Figure 2).

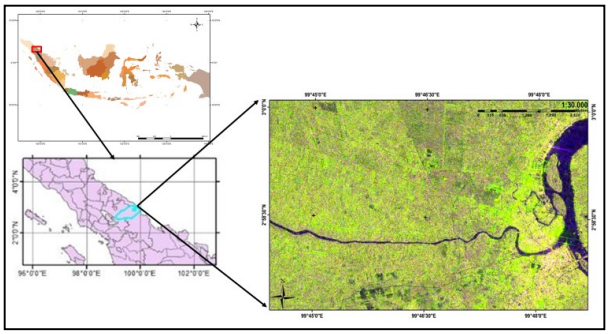

Fig 2. Study Area of Asahan, North Sumatra

B. Data Collection

In this study, we used SAR data include Terra SAR-X with X-band, Sentinel-1A with C-band, and ALOS PALSAR-2 with L-band (Figure 2) and the acqusition in table 1. TerraSAR-X and ALOS PALSAR-2 were obtained from LAPAN (National Institute of Aeronautics and Space) and Sentinel-1A was obtained from ESA (European Space Agency) Copernicus.

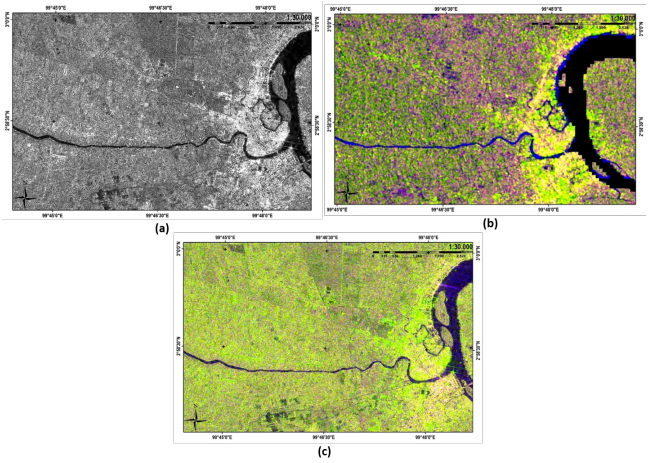

Fig 2. Data Collection (a) TerraSAR-X, (b) Sentinel-1A, and (c) ALOS PALSAR.

TABLE 1. SAR Data Acquisition

Sensor	Terra SAR-X	Sentinel-1A	ALOS PALSAR-2
Polarization	Single Polarization (HH)	Dual Polarization (VV, VH)	Dual Polarization (HH, HV)
Mode		Interferometric Wide Swath	
Wavelenght Range	X-band	C-band	L-band
Pixel Spacing	3m	5-20m	10m
Processing Level		Level 1	
Product Type		Ground Range Detected Image	
Date of Acquisition	August 2017	July 2018	January 2015

C. Pre-Processing

First geometric and radiometric corrections, filtering using the Lee algorithm and identifying the training sample extracted from the land cover chart of the Plantation Agency of Indonesia.

D. Composite

Identification and classification of land use using SAR was performed using RGB image with combination of TerraSAR-X on HH (figure 3a), Sentinel1-A on VV and VH (figure 3b), and ALOS PALSAR-2 on HH and HV. Each polarization were conversion from digital number (DN) to backscatter coefficient.

The DN values for TerraSAR-X can be converted to gamma naught values in decibel unit (dB) using the following equation (1) [17]:

$$\beta^{\circ}_{dB} = 10 \times \log_{10}(K_{\gamma} DN^2) \quad (1)$$

The Sentinel-1A product uses radiometric calibration look-up table (LUT) to do the calibration. The important conversion of amplitude to DN and from DN to sigma nought were done automatically on SNAP and once the sigma nought values was obtained, the computation of backscatter ($\sigma°_{dB}$) can be performed in equation (2) [3]:

$$\sigma°_{dB} = 10.\log_{10}[\gamma_i] \quad (2)$$

where γ_i is the gamma-calibrated scattering coefficient of Sentinel-1A data.

The conversion of HH (DN_{HH}) and HV (DN_{HV}) backscatter intensities into NRCS (that is $\sigma°_{HH}$ and $\sigma°_{HV}$) [18], [19] was based on Shimada et al. [19] shown in equation (3) and (4):

$$\sigma°_{HH}\ (\mathrm{dB}) = 10\ \mathrm{x}\ \log_{10}(DN_{HH}^2) - CF \quad (3)$$

$$\sigma°_{HV}\ (\mathrm{dB}) = 10\ \mathrm{x}\ \log_{10}(DN_{HV}^2) - CF \quad (4)$$

where, σ° is scattering coefficient and CF is the calibration factor. The CF is dependent on the processing date, in this study CF is equal to -83.0 both for HH and HV [18].

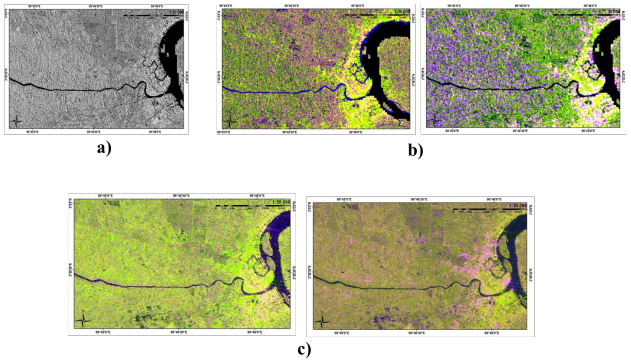

Fig 3. The combination of polarization in SAR data: a).HH polarization onTerraSAR-X; b). Combination of VV and VH polarization of Sentinel-1A; c). Combination of HH and HV polarization of ALOS PALSAR-2

E. Training sample

In order to classify, The supervised classifier uses the ground truth ROI image as a training set to identify the class and the total accuracy and kappa coefficient is determined from the parallelepiped, Support vector machine, and maximum likelihood classified image and ground truth ROI [20]. The size of the training sample was 3x3 pixels (figure 3).

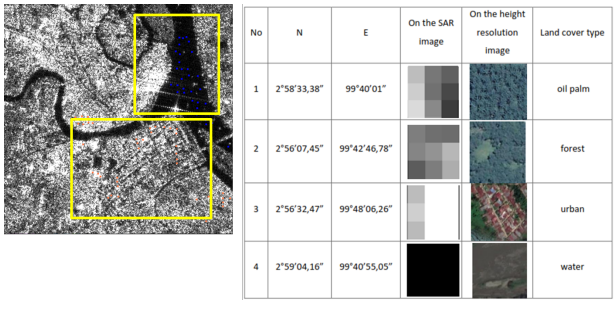

Fig 3. Training sample

F. Classification

In general , the purpose of the classification of images is to classify various spectral classes identifying various types of land cover. Supervised classification procedures include the assignment and identification of image pixels to spectral classes of interest on the basis of their spectral characteristics [21]. A variety of algorithms has been suggested for the supervised classification of remote sensing data. In this study, we use three well-known classifiers, Parallepiped, SVM (Support Vector Machine), and Maximum Likelihood.

1) Parallepiped

The parallelepiped classifier is one of the simplest supervised classification methods and only examines histograms of the individual class spectral components [21].

2) Support Vector Machine (SVM)

The versatility of the SVM classifier includes an ability to handle minimal training data, which is a considerable benefit for land-cover mapping in remote sensing [22]. SVM attempts to map training vectors into a high-dimensional feature space where linear separation could be implemented with the aid of a kernel function [23]. SVM consists of several hyperparameters: kernel type, gamma and penalty value. These hyperparameters can be tuned and modified to improve SVM efficiency in the image classification [24]. Support vectors taken from training data are employed as a measure of separability of classes [25]. Optimal separability is, therefore, shown by wide distances between support vectors.

3) Maximum Likelihood

The maximum likelihood can be considered the most common supervised approach in the classification of images and is based on the statistical distribution of the data [26]. The accuracy of this approach depends on the correct estimate of the mean vector and the covariance structure of the spectral classes [10]. This approach also demands a reasonable number of training data for the correct calculation of the covariance matrix to perform as an effective classifier [26].

G. Accuracy

Based on the sample data (RoI) to analyze the classification of each class, we assessed the classification results using standard accuracy assessment metrics, including error matrix, overall accuracy, and user and producer accuracy using the equation (5)–(8) [27].

$$\text{Kappa} = \frac{N\sum_{i=1}^{r} Xii - \sum_{i=1}^{r} Xi+(X+i)}{N^2 - \sum_{i=1}^{r} Xi+(X+i)} \times 100\% \quad (5)$$

$$\text{Overall Accuracy} = \frac{\sum_{i=1}^{r} Xii}{N} \times 100\% \quad (6)$$

Where N is the number of all pixels used for observation, r is the number of rows in the error matrix (number of classes), Xii is diagonal values of the contingency matrix of row i and column i, $X + i$ is column pixel number i, and Xi + row pixel number i.

III. RESULT AND DISCUSSION

The spatial distribution of classification algorithm from TerraSAR-X, Sentinel-1A, and ALOS PALSAR-2 can be seen in figure 4. Each classification algorithm were categorized into four land cover classes, such as oil palm, forest, water and built-up areas.

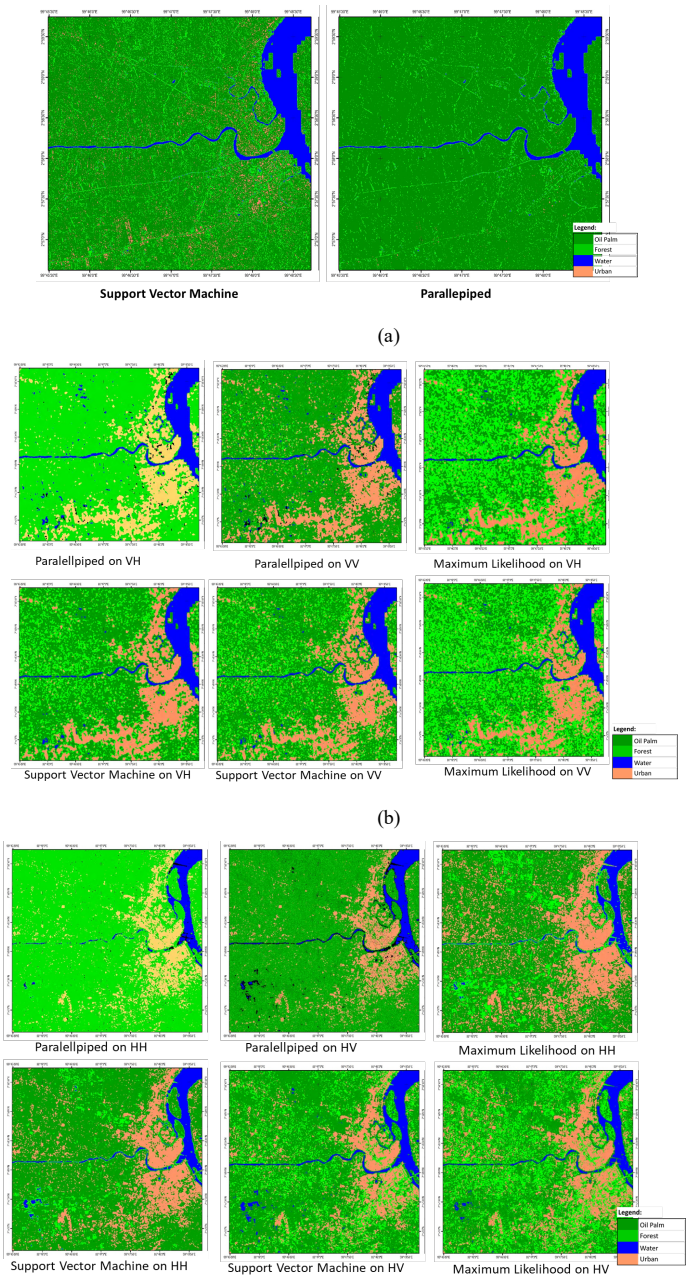

Fig. 4. Spatial distribution of classification algorithm on a). TerraSAR-X; b). Sentinel-1A, and c). ALOS PALSAR-2

For average of overall accuracy and kappa coefficient can be seen on the table 2 and table 3. On the table 2 and table the best algorithm with the highest accuracy value is the Support Vector Machine, the accuracy of 91.17% on ALOS PALSAR, 86.18% on Sentinel 1, and 77.57% on TerraSAR X. According to [28], The most important characteristic is the ability of SVM to generalize well from the limited amount and / or quality of training data. SVM can produce comparable accuracy using a much smaller training sample size. This is in line with the concept of 'support vector' which relies only on a few data points to define the classifier hyper plane. SVM handles quadratic problems so that it always reaches global minimums. An added advantage is that there is no need to repeat classifier training using a different initialization or random architecture. Additionally, as non-parametric, SVM does not assume a known statistical distribution of the data to be classified [28], [29].

TABLE 2. Average of Overall accuracy

Algorithm	TerraSAR X	Sentinel 1	Alos Palsar
Parallelepiped	69.15%	81.04%	78.19%
Maximum Likelihood	-	84.33%	89.61%1
Support Vector Machine	77.57%	86.20	91.19%

TABLE 3. Average of Kappa coefficient

Algorithm	TerraSAR X	Sentinel 1	Alos Palsar
Parallelepiped	0.48	0.70	0.66
Maximum Likelihood	-	0.77	0.85
Support Vector Machine	0.65	0.79	0.87

Based on table 2 and 3 also can be seen that the ALOS PALSAR imagery has high accuracy compared to other images, the Parallelepiped algorithm with an accuracy of 78.19%, the Maximum Likelihood algorithm with an accuracy of 89.62%, and the Support Vector Machine algorithm with an accuracy of 91.18%. according to [30], the scattering value for HH polarization using L-band increases the attenuation of waves propagating through the soil surface and ridges, along with the increase in leaves, leaves and stems. L-band frequency provides the best diagnosis of oil palm canopy structure to increase oil palm growth and also the scattering coefficient at 1 GHz (L-band) will be more sensitive to the structure and moisture content of oil palm canopy relative to higher frequencies [8]. The higher the probability of the signal penetrating the vegetation canopy [4]. Longer wavelengths (e.g. L-band 15-20 cm) penetrate the canopy and convey information on leaves, branches, stems and soil conditions; the longer the wavelength the greater the penetration [4].

IV. CONCLUSION

The study was showed the investigation of three classification algorithm using Parallepiped, Support Vector Machin, and Maximum Likelihood based on multi-scatter SAR data. In this study, the best algorithm with the highest accuracy value is the Support Vector Machine, the accuracy of 91.17% on ALOS PALSAR, 86.18% on Sentinel 1, and 77.57% on TerraSAR X. The ALOS PALSAR imagery has high accuracy compared to other images, the Parallelepiped algorithm with an accuracy of 78.19%, the Maximum Likelihood algorithm with an accuracy of 89.62%, and the

Support Vector Machine algorithm with an accuracy of 91.18%.

REFERENCES

[1] R. H. V Corley, *Oil Palm Physiology: A Review*. Incorporated Society of Planters, 1973.

[2] R. H. V Corley and P. B. Tinker, *The Oil Palm. [electronic resource]*, Fifth Edit. Chichester, West Sussex: Wiley Blackwell, 2016.

[3] I. Carolita *et al.*, "Comparison of Optic Landsat-8 and SAR Sentinel-1 in Oil Palm Monitoring, Case Study : Asahan, North Sumatera, Indonesia," *IOP Conference Series: Earth and Environmental Science*, vol. 280, no. 1, p. 012015, Aug. 2019, doi: 10.1088/1755-1315/280/1/012015.

[4] S. Darmawan, W. Takeuchi, A. Haryati, R. A. M. Najib, and M. Na'Aim, "An investigation of age and yield of fresh fruit bunches of oil palm based on ALOS PALSAR 2," in *IOP Conference Series: Earth and Environmental Science*, 2016, vol. 37, no. 1, doi: 10.1088/1755-1315/500/1/012075.

[5] I. Carolita *et al.*, *Profil Pusat Pemanfaatan Penginderaan Jauh Lembaga Penerbangan dan Antariksa Nasional*, no. 8. 2019.

[6] T. M. Lillesand and R. W. Kiefer, "Remote sensing and image interpretation.," *Remote sensing and image interpretation.* 1979, doi: 10.2307/634969.

[7] M. Shimada, O. Isoguchi, T. Tadono, and K. Isono, "PALSAR radiometric and geometric calibration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 12, pp. 3915–3932, 2009, doi: 10.1109/TGRS.2009.2023909.

[8] S. Darmawan, D. K. Sari, W. Takeuchi, K. Wikantika, and R. Hernawati, "Development of aboveground mangrove forests' biomass dataset for Southeast Asia based on ALOS-PALSAR 25-m mosaic," *Journal of Applied Remote Sensing*, vol. 13, no. 4, 2019, doi: 10.1117/1.JRS.13.044519.

[9] S. Darmawan, W. Takeuchi, Y. Vetrita, K. Wikantika, and D. K. Sari, "Impact of topography and tidal height on ALOS palsar polarimetric measurements to estimate aboveground biomass of mangrove forest in Indonesia," *Journal of Sensors*, vol. 2015, 2015, doi: 10.1155/2015/641798.

[10] S. Darmawan, I. Carolita, and E. Ananta, "Identification of Oil Palm Plantation on Multiscatter and Resolution of SAR Data Using Variety of Classifications Algorithm (Case Study: Asahan District, North Sumatera Province)," in *IOP Conference Series: Earth and Environmental Science*, 2020, vol. 500, no. 1, doi: 10.1088/1755-1315/500/1/012075.

[11] L. Li, J. Dong, S. N. Tenku, and X. Xiao, "Mapping oil palm plantations in cameroon using PALSAR 50-m orthorectified mosaic images," *Remote Sensing*, vol. 7, no. 2, pp. 1206–1224, 2015, doi: 10.3390/rs70201206.

[12] L. F. Yayusman and R. Nagasawa, "ALOS-Sensor data integration for the detection of smallholder's oil palm plantation in Southern Sumatra, Indonesia," *Journal of the Japanese Agricultural Systems Society*, vol. 31, no. 2, pp. 27–40, 2015, doi: 10.14962/jass.31.2_27.

[13] Y. W. Kee, A. R. M. Shariff, A. M. Sood, and L. Nordin, "Application of SAR data for oil palm tree discrimination," *IOP Conference Series: Earth and Environmental Science*, vol. 169, no. 1, 2018, doi: 10.1088/1755-1315/169/1/012065.

[14] Y. W. Kee, A. R. M. Shariff, A. M. Sood, and L. Nordin, "Application of SAR data for oil palm tree discrimination," in *IOP Conference Series: Earth and Environmental Science*, 2018, vol. 169, no. 1, doi: 10.1088/1755-1315/169/1/012065.

[15] A. C. Morel *et al.*, "Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data," *Forest Ecology and Management*, vol. 262, no. 9, pp. 1786–1798, 2011, doi: 10.1016/j.foreco.2011.07.008.

[16] C. Pohl, "Mapping palm oil expansion using SAR to study the impact on the CO 2 cycle," *IOP Conference Series: Earth and Environmental Science*, vol. 20, no. 1, pp. 0–8, 2014, doi: 10.1088/1755-1315/20/1/012012.

[17] R. Düring, F. N. Koudogbo, M. Weber, and I. Gmbh, "TerraSAR-X and TanDEM-X : Revolution in Spaceborne Radar," 2007.

[18] S. Darmawan, W. Takeuchi, A. Haryati, R. Najib A M, and M. Na'aim, "An investigation of age and yield of fresh fruit bunches of oil palm based on ALOS PALSAR 2," in *IOP Conference Series: Earth and Environmental Science*, 2016, vol. 37, no. 1, p. 012037, doi: 10.1088/1755-1315/37/1/012037.

[19] M. Shimada, "Model-Based Polarimetric SAR Calibration Method Using Forest and Surface-Scattering Targets," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 5, pp. 1712–1733, May 2011, doi: 10.1109/TGRS.2010.2090046.

[20] V. A, S. P, and K. M, "SAR Ice Image Classification Using Parallelepiped Classifier Based on Gram-Schmidt Spectral Technique," no. September 2013, pp. 385–392, 2013, doi: 10.5121/csit.2013.3540.

[21] I. Entezari, M. Motagh, and B. Mansouri, "Comparison of the performance of L-band polarimetric parameters for land cover classification," *Canadian Journal of Remote Sensing*, vol. 38, no. 5, pp. 629–643, 2012, doi: 10.5589/m12-051.

[22] J. A. Dos Santos, P. H. Gosselin, S. Philipp-Foliguet, R. Da S. Torres, and A. X. Falao, "Multiscale classification of remote sensing images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 10 PART1, pp. 3764–3775, 2012, doi: 10.1109/TGRS.2012.2186582.

[23] B. H. Trisasongko, D. R. Panuju, D. J. Paull, X. Jia, and A. L. Griffin, "Comparing six pixel-wise classifiers for tropical rural land cover mapping using four forms of fully polarimetric sar data," *International Journal of Remote Sensing*, vol. 38, no. 11, pp. 3274–3293, 2017, doi: 10.1080/01431161.2017.1292072.

[24] H. M. Rizeei, H. Z. M. Shafri, M. A. Mohamoud, B. Pradhan, and B. Kalantar, "Oil palm counting and age estimation from WorldView-3 imagery and LiDAR data using an integrated OBIA height model and regression analysis," *Journal of Sensors*, vol. 2018, 2018, doi: 10.1155/2018/2536327.

[25] S. S. Nath, G. Mishra, J. Kar, S. Chakraborty, and N. Dey, "A survey of image classification methods and techniques," *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies, ICCICCT 2014*, pp. 554–557, 2014, doi: 10.1109/ICCICCT.2014.6993023.

[26] M. B. A. Gibril, S. A. Bakar, K. Yao, M. O. Idrees, and B. Pradhan, "Fusion of RADARSAT-2 and multispectral optical remote sensing data for LULC extraction in a tropical agricultural area," *Geocarto International*, vol. 32, no. 7, pp. 735–748, 2017, doi: 10.1080/10106049.2016.1170893.

[27] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote sensing of environment*, vol. 37, no. 1, pp. 35–46, 1991.

[28] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, 2011, doi: 10.1016/j.isprsjprs.2010.11.001.

[29] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machines," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 6, pp. 1335–1343, 2004, doi: 10.1109/TGRS.2004.827257.

[30] K. C. Teng, J. Y. Koay, S. H. Tey, K. S. Lim, H. T. Ewe, and H. T. Chuah, "A dense medium microwave backscattering model for the remote sensing of oil palm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 6, pp. 3250–3259, 2015, doi: 10.1109/TGRS.2014.2372796.