8th International Conference on Geomatics and Geospatial Technology (GGT 2022)
IOP Conf. Series: Earth and Environmental Science 1051 (2022) 012024
doi:10.1088/1755-1315/1051/1/012024

Oil Palm Plantation Land Cover and Age Mapping Using Sentinel-2 Satellite Imagery and Machine Learning Algorithms

A N Jarayee¹, H Z M Shafri¹, Y Ang¹, Y P Lee², S A Bakar³, H Abidin², H S Lim³ and R Abdullah⁴

¹Department of Civil Engineering and Geospatial Information Science Research Centre (GISRC), Faculty of Engineering, Universiti Putra Malaysia (UPM), 43400 Serdang, Selangor, Malaysia
²Geoinformatics Unit, FGV R&D Sdn Bhd, FGV Innovation Centre, PT23417, Lengkuk Teknologi, 71760 Bandar Enstek, Negeri Sembilan, Malaysia.
³School of Physics, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia.
⁴School of Computer Sciences, Universiti Sains Malaysia (USM), 11800 Gelugor, Penang, Malaysia.

Corresponding author: helmi@upm.edu.my

Abstract. Nowadays, there are various techniques and methods used in land cover classification using remote sensing data especially in oil palm monitoring. This study discussed the oil palm mapping using satellite imagery (Sentinel-2) and classification of land cover features using machine learning algorithms such as linear support vector classifier (LSVC), random forests (RF) and deep neural network (DNN). A total 13218 sampling points (80% of the total sampling points used as training samples and 20% applied as testing samples) were randomly selected in the study area which were then classified into six land cover features; water, bare soil, forest, immature oil palm (the age of 2-8 year), mature oil palm (age >8 year) and built-up area. These data were validated by using spectral reflectance, Google Earth Pro and ground checking. The accuracy assessment was conducted by a confusion matrix method. The results showed that classification of land features using DNN with batch size 32 and epoch 100 has the highest accuracy which is 99.35% for overall accuracy and 98.49% kappa accuracy. This study demonstrated various machine learning algorithms that may be used to detect and classify the maturity of oil palm trees, which is vital to record in tree inventories for effective plantation management.

1. Introduction
Oil palm, *Elaeis guineensis* is the main contributor to Malaysia’s economy in the agriculture industry according to the Department of Statistics Malaysia (DOSM). The oil palm plantations have conquered approximately 11.10 million acres of land in Malaysia. It is difficult for the plantation staff to monitor the development of the oil palm trees and manage the plantation area due to the large number of oil palm trees in a wide area and it will require more human power, consume time, and be costly. In the plantations industry, it is essential to ensure the development of the crops is manageable to make sure the production of yield is high which leads to an increase in profit. Other than that, it is important for
efficient management in croplands and the expansion of the plantation has to be carefully arranged in order to prevent the destruction of ecosystems. Recently, remote sensing has been widely used in monitoring the oil palm plantation, where these remote sensing will be used in the oil palm mapping and the classification of land cover that is important for better inventory management and for the manager to plan the strategy for the replanting.

Researchers [1][2][3][13] had conducted various different techniques of classification and for oil palm yield estimation that utilize different machine learning algorithms such as the random forest (RF), modified AdaBoost algorithms, support vector machine (SVM), maximum likelihood (ML), artificial neural network (ANN), and more. The land cover mapping involved the selection of the high resolution of satellite images such as Landsat-8, Sentinel-1, Sentinel-2, and others. The machine learning algorithms were applied and the confusion matrix is calculated to generate the accuracy of the model. Accuracy assessments have been conducted in most of the studies. An evaluation to determine the accuracy of the used algorithms is essential in land cover mapping as it is to compare and determine the best land cover features classifier. Oil palm data extraction and land cover classification are conducted more efficiently in RF than other machine learning algorithms [1][3].

There is a trend using advanced machine learning algorithms to classify, and yet there is still a lack of studies using advanced machine learning, especially deep learning algorithms for land cover classification and oil palm mapping in high-resolution satellite data. Existing research or studies have used machine learning algorithms for land cover classification and crop mapping [1][2][3][13]. Furthermore, open-source plugins in GIS software using advanced machine learning algorithms especially deep learning algorithms are very limited. This study will use machine learning algorithms such as the linear support vector classifier (LSVC), random forest (RF), and deep neural network (DNN) to classify the land cover features and monitor the maturity of the oil palm trees. Therefore, the objective of this study is to examine oil palm mapping using high resolution freely available satellite imagery and carry out land cover classification and oil palm mapping using advanced machine learning algorithms.

2. Methodology

2.1 Study Area
Our study area comprises 23 blocks in a research plantation located in Pahang state in Peninsular Malaysia, covering 24-kilometer square.

![Figure 1. Map of study area location.](image)
2.2 Design of Study
The design study for this study is as shown in Figure 1. It consists of three main parts which are data acquisition, processing, and result.

2.3 Data Acquisition
The data used in this study is the Sentinel-2 imagery in the year 2020. The Sentinel-2 image is obtained from the Sentinel-2 satellite that was launched during the Copernicus Programme by the European Space Agency (ESA). This satellite consists of multispectral instruments which measure the reflection of radiance from Earth in 13 spectral bands (from near-infrared to shortwave infrared) obtained from 10, 20 and 60 m spatial resolution as shown. The Sentinel-2 image is analyzed by using Quantum GIS (QGIS), an open-source geographic information system (GIS) software. To overcome the limitation obstructed by cloud cover, we used a cloud percentage (of 10%) as a screening method for creating our cloud-free composite. Following that, the image was pre-processed using cloud masking. First, the QA60 bitmask band (a quality flag band) was utilized to detect and mask out opaque clouds and cirrus, and then sentinel2 images were scaled by 10,000. The spatial resolution of all satellite bands was resampled to 10m using the nearest neighbour method [4].

2.4 Sampling Strategy
The sampling strategy used in this study is random sampling, the most common sampling technique in remote sensing. The sampling strategy is important to improve the accuracy of the prediction of samples that have fewer samples otherwise it will cause an increase in cost and generate poor accuracy [5]. The training samples that were used to classify the S2 imagery in this research were 80% (10574 points) of the total samples and 20% (2644 points) of total samples applied as testing samples that were used to assess the accuracy of the classifiers used. The samples were verified by using Google Earth Pro and spectral reflectance signature in the EnMAP plugin.

2.5 Classification
Nowadays, machine learning is often being used to do classification of land cover classes. It is useful in remote sensing data classification because of its efficient classification of satellite imagery. Machine learning center of attention is the predicted features, the performance of the classifier to predict efficiently, and the method used to determine the success of the prediction [6]. In this study, there are
six land cover features used which are water, build up area, bare soil, forest, immature oil palm (the age of 2-8 year) and mature (age > 8 year). The classification in machine learning is dependent on the hyperparameters of the respective machine learning algorithms. These hyperparameters are derived through programming however since in this study, the classification using machine learning is processed through plugins as shown in Table 1 in QGIS.

**Table 1.** Machine learning algorithms and the parameters used.

| No. | Models | Plugins used for analysis | Parameters | Definition |
|-----|--------|---------------------------|------------|------------|
| 1   | LSVC   | EnMAP-Box                 | ‘C’ (0.001, 0.01, 0.1, 1, 10, 100, 1000). | ‘C’ = penalty parameter of the error term |
| 2   | RF     | EnMAP-Box                 | n_estimators (100), oob_score (TRUE). | n_estimators = number of trees |
| 3   | DNN    | Advanced Hyperspectral Data Analysis Software (Avhyas) | hidden_layers (2), epoch (50 and 100), learning rate (0.01) and batch sizes (32 and 64). | hidden_layers = number of hidden layer Epoch = number of times to train the entire data |

2.5.1 Linear Support Vector Classifier. The purpose of Linear support vector classifier (LSVC) is to fit the data that is user-provided by producing the best fit of a hyperplane that divides the data used in this research. The prediction of the classification is determined after the production of the hyperplane, where the desired features are added to the classifiers. Before the processing, the hyperparameters such as the penalty parameter of the error term (‘C’) which consist of seven values, 0.001, 0.01, 0.1, 1, 10, 100, and 1000. The higher the value of ‘C’, the more accurate the training sample. However, if the ‘C’ value is way too high, it might cause the overfitting of the training data. This overfitting data will cause problems for future observations if new data are being added. In these plugins, it will choose the suitable ‘C’ values for the classification.

2.5.2 Random Forests. Random forest (RF) classifier is the most common algorithm used nowadays due to its efficiency in classification. RF utilizing bagging ensemble method to train the random datasets. These datasets were trained separately and the outcomes are produced via majority voting or calculating the average to generate high accuracy models. There are several studies that compare the usage of random forest algorithms with other algorithms such as maximum likelihood classifiers [7]. In this study, the classification using random forest is conducted in ENMAP plugins. The model parameter is the number of trees (n_estimators) which default equal to 100. N_estimators is the number of trees that the users desired before the maximum or average voting is being taken. The higher the n_estimator, the better the performance of the classification. Next model parameter is out of bag score (oob_score) which is used to validate the outcomes of the classification model in random forest.

2.5.3 Deep Neural Network. Deep neural networks (DNN) contain one input layer, several hidden layers and one output layer. The higher the number of hidden layers, the deeper the network. Two hidden layers was implemented in Avhyas plugin. Those layers are linked, with the outcomes of the past layers that become the source of the current layer. These layers are weighted and these weights will determine the performance of the network [8]. In this study, the hyperparameters (epoch, learning rate, and batch size) were adjusted in Avhyas plugins in QGIS.
2.6 Accuracy Assessment
Accuracy assessment is significant in remote sensing as it is used to evaluate the accurateness of the land cover classification methods. The accuracy is tested by using ground truth data of the six features of the land cover. These accuracies are the result of the analysis of the image classification by the machine learning algorithms. The confusion matrix is used to evaluate the performance of a classification model by comparing the actual classes to the machine learning model's predictions. By summing the number of correctly classified classes and dividing by the total number of classes, the overall accuracy is obtained. The correctly classified classes are computed using a confusion matrix. The resulting kappa measure accounts for chance agreement in classification and indicates how much better the classification performed. These are two accuracies that will be used in this study which are the overall accuracy and Kappa accuracy.

3. Results and Discussion

3.1 Spectral reflectances of the features
Spectral reflectances are different for each of the classes. Therefore, it is used to confirm the classification during ground-truthing alongside using Google Earth Pro. These spectral reflectances are obtained from the Sentinel-2 image. By selecting the random area, the spectral reflectance can be determined at the respective area and the value can be compared with the confirmed features and the user can manually visualize the graph of the spectral reflectance.

![Figure 3](image)

**Figure 3.** The spectral reflectance profile for respective features.

Figure 3 shows the band number against spectral reflectance for six land cover features. The band number represents the wavelength of the respective band. This spectral reflectance profile is obtained from the EnMAP by randomly selecting the respective features. As shown in the graph, each of the features have different reflectance. Vegetation absorbs more visible light and reflects near infrared radiation due to the need for photosynthesis [9]. Different species of vegetation contain different content of pigment and texture. Thus, it will result in a difference of absorption of radiation.

3.2 Classification results
The classification result was evaluated by using a confusion matrix that was generated from the classification results in the plugins. Testing samples were used to evaluate the land cover classifications features and confusion matrix were produced that included the OA and KA (see Table 2) and UA and PA (see Table 3) for each of the algorithms. This confusion matrix determined the accuracy of the data processing from satellite image Sentinel-2 compared with real world situations [10]. The overall accuracy and kappa accuracy of the three algorithms were quite close to each other as shown in Table
3. The highest overall accuracy and kappa accuracy was generated by DNN with batch size (32) and epoch (100) which is 99.35% and 98.49% respectively. Followed by another scenario with different batch size (64) but different epoch (50), which it resulted 98.63% for overall accuracy and 96.81% for kappa accuracy and RF algorithms with an overall accuracy of 98.51% and kappa accuracy 96.63% and finally the LSVC that produced 98.12% overall accuracy and 95.7% kappa accuracy.

![Figure 4](image1.png)

**Figure 4.** Classification results of the study area: (a) LSVC, (b) RF, (c) DNN (Batch size= 32, Epoch= 100) and (d) DNN (Batch size= 64, Epoch= 50) methods.

| No. | Algorithms | Parameters | OA (%) | KA (%) |
|-----|------------|------------|--------|--------|
| 1   | LSVC       | ‘C’= 0.001, 0.01, 0.1, 1, 10, 100, 1000 | 98.12  | 95.7   |
| 2   | RF         | n_estimators= 100, oob_score= TRUE | 98.51  | 96.63  |
| 3   | DNN        | Batch size, Learning rate, Epoch | OA (%) | KA (%) |
|     |            | 32, 0.01, 100 | 99.35  | 98.49  |
|     |            | 64, 0.01, 50  | 98.63  | 96.81  |

**Table 2.** Comparison of classification results.
Table 3. Comparison of classification results for each class.

| No. | Map Class                  | Machine Learning |           | Deep Learning |           |
|-----|----------------------------|------------------|-----------|---------------|-----------|
|     |                            | LSVC             | RF        | DNN (BS=32)   | DNN (BS=64)|
|     |                            | UA (%)           | PA (%)    | UA (%)        | PA (%)    |
|     |                            |                  |           | UA (%)        | PA (%)    |
| 1   | Build up                   | 100.00           | 98.31     | 99.14         | 97.46     |
| 2   | Forest                     | 98.96            | 99.86     | 99.69         | 99.31     |
| 3   | Bare soil                  | 84.00            | 87.50     | 59.15         | 87.50     |
| 4   | Mature oil palm (>10 yr)   | 98.15            | 93.92     | 97.60         | 96.59     |
| 5   | Immature oil palm (2-10 yr)| 82.26            | 90.27     | 87.55         | 86.46     |
| 6   | Water                      | 98.73            | 100.00    | 95.00         | 97.44     |

UA: User’s accuracy; PA: Producer’s accuracy

The results showed deep learning algorithms (DNN) produce higher accuracy compared to other machine learning algorithms (LSVC and RF). For both models of machine learning algorithms which are LSVC and RF able to precisely detect and differentiate the vegetation (mature, immature and forest) classes which shows high accuracy. According to Figure 4 (c) and Figure 4 (d), it is easily confused with the classification of bare soil and water features compared to machine learning algorithms in Figure 4 (a) and Figure 4 (b) that could distinguish bare soil and water features in the study area. The causes of these problems could be the depth of the layers in the deep neural network processing, where this layer depth gives important impacts to the accuracy of the classification [11]. For all models, the build-up classes are the considerably most accurate differentiated features, where the values of accuracy are higher than 95 %. Between the machine learning algorithms; RF and LSVC, RF are able to provide more accuracy and are more efficient in classification of features than linear support vector classifiers. This is due to its ability to solve both classification and regression through majority voting or average calculation [12]. Based on table 4 shown above, the accuracy percentage (UA and OA) for both immature and mature oil palm trees show lower accuracy than in other classes. A previous study has shown this is due to the confusion in the classifier and caused misclassification between the oil palm trees and forests since both classes belong to the same class which is vegetation that has similar spectral properties [13].

4. Conclusion
This study represents the usage of advanced machine learning in production of land cover classification and oil palm age mapping using Sentinel-2 data. The results show the oil palm mapping using Sentinel-2 imagery is satisfying and acceptable. Based on the obtained result, the highest accuracy for both overall and kappa accuracy was generated by DNN that comprises batch size 32 and epoch 100 which is 99.35% and 98.49% respectively. This shows that deep learning algorithms are able to generate more accurate and satisfying classification over other machine learning algorithms such as RF and LSVC. In future studies, precise age separation of oil palm will be performed to facilitate inventory management in the plantation.
References

[1] Shaharum N S N, Shafri H Z M, Ghani W A W A K, Samsatli S, Al-Habshi M M A and Yusuf B 2020 Oil palm mapping over Peninsular Malaysia using Google Earth Engine and machine learning algorithms Remote Sens. Appl. Soc. Environ. 17 100287

[2] Prasad P, Loveson V J, Chandra P and Kotha M 2022 Evaluation and comparison of the earth observing sensors in land cover/land use studies using machine learning algorithms Ecol. Inform. 68 101522

[3] Talukdar S, Singha P, Mahato S, Shahfahad, Pal S, Liou Y A and Rahman A 2020 Land-use land-cover classification by machine learning classifiers for satellite observations-A review Remote Sens. 12

[4] Roy D P, Li J, Zhang H K and Yan L 2016 Best practices for the reprojection and resampling of Sentinel-2 Multi Spectral Instrument Level 1C data Remote Sens. Lett. 7 1023–32

[5] Mu X, Hu M, Song W, Ruan G, Ge Y, Wang J, Huang S and Yan G 2015 Evaluation of sampling methods for validation of remotely sensed fractional vegetation cover Remote Sens. 7 16164–82

[6] Loussaief S and Abdelkrim A 2017 Machine learning framework for image classification 2016 7th Int. Conf. Sci. Electron. Technol. Inf. Telecommun. SETIT 2016 3 58–61

[7] Jin Z, Shang J, Zhu Q, Ling C, Xie W and Qiang B 2020 RFRSF: Employee Turnover Prediction Based on Random Forests and Survival Analysis Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics) 12343 LNCS 503–15

[8] Vasuki A and Govindaraju S 2017 Deep neural network for image classification. Adv. Parallel Comput. 31 27 - 49

[9] Ozuyavuz M, Bilgili B C and Salici A 2015 Determination of vegetation changes with NDVI method J. Environ. Prot. Ecol. 16 264–73

[10] Ramadhan Kete S C, Suprihatin, Tarigan S D and Effendi H 2019 Land use classification based on object and pixel using Landsat 8 OLI in Kendari City, Southeast Sulawesi Province, Indonesia IOP Conf. Ser. Earth Environ. Sci. 284

[11] Zhang P, Ke Y, Zhang Z, Wang M, Li P and Zhang S 2018 Urban land use and land cover classification using novel deep learning models based on high spatial resolution satellite imagery Sensors (Switzerland) 18

[12] Sheykhmousa M, Mahdianpari M, Ghanbari H, Mohammadimanesh F, Ghamisi P and Homayouni S 2020 Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 13 6308–25

[13] Ang Y, Shafri H Z M, Bakar S A, Abidin H, Lee Y P, Hashim S J, Che’Ya N N, Hassan M R, Lim H S and Abdullah R 2022 A Novel Ensemble Machine Learning and Time Series Approach for Oil Palm Yield Prediction Using Landsat Time Series Imagery Based on NDVI Geocarto Int. 0 1–24

Acknowledgements

The researchers would like to acknowledge the Ministry of Higher Education (MOHE), Malaysia for the resources and financial support through Long Term Research Grant Scheme (LRGS)-Malaysia Research University Network (MRUN). This project is conducted under the research programme: ‘A Big Data Analytics Platform for Optimizing Oil Palm Yield Via Breeding by Design (Grant No: 203.PKOMP.6770007)’ with specific project: ‘Geoinformatics Data for Palm Oil Yield Prediction Using Machine Learning (Vote No: 6300268-10801)’. The authors would also like to extend a special acknowledgment to the team from FGV R&D Sdn. Bhd. for their expertise in supporting this research.