RESEARCH PAPER

SMART DETECTION OF ILLICIT CANNABIS PLANTATIONS USING REMOTE SENSING TECHNOLOGY AND MACHINE LEARNING

Dedi Irawadi[1,3*], Tuga Mauritsius[1], Dony Kushardono[2], Syarif Budhiman[2], Karunika Diwyacitta[2], Bayu S. Adhitama[2], Fauzan A. Ayubi[2], Olivia Maftukhaturrizqoh[2], Ika S. Supriyani[2]

[1] Information Systems Management Department, Bina Nusantara University, Jakarta, 11480, Indonesia
[2] Research Center for Geoinformatics, Research Organization for Electronics and Informatics, National Research and Innovation Agency (BRIN), Bandung, 40135, Indonesia
[3] Research Center for Satellite Technology, Research Organization for Aeronautics and Space, National Research and Innovation Agency (BRIN), Bogor, 16310, Indonesia
*Corresponding author: dedi004@brin.go.id; dedi.irawadi@binus.ac.id

ABSTRACT. Remote sensing technology and machine learning classifiers can be utilized to develop smart detection systems for illicit crops such as *Cannabis sativa L*. Machine learning algorithms for classifying medium-resolution optical satellite data can be compared to identify the best model for enhancing law enforcement's detection of illicit crops efficiently and accurately. Remote sensing-based smart detection systems have been developed in South America and Central Asia; however, these methods cannot be used effectively for Indonesia due to high cloud coverage, geographical differences, and the smaller area of *Cannabis sativa L.* plantations. This research developed an agile methodology that employs backpropagation neural networks to analyze the statistical growth phenology of cannabis derived from multitemporal medium-resolution remote sensing data. Using datasets derived from Indonesian law enforcement eradication records, the method achieved 94% accuracy and a kappa coefficient of 0.9. Further, plant growth phenology based on vegetation index values from multitemporal data was used to assess the condition of identified cannabis plantations.

INTRODUCTION

According to the World Drug Report 2022 released by UNODC, *Cannabis sativa L.*, or marijuana, is the most widely used addictive substance globally, with over 209 million users in 2022. The Indonesian National Narcotics Agency (BNN) reported that marijuana accounts for 65.5% of all illicit drug abuse cases, with the prevalence of drug use rising from 1.80% in 2019 to 1.95% in 2021[1]. The drug eradication programs, including detecting cannabis plantations, will contribute to decreasing the availability of these prohibited substances. New approaches, such as remote sensing technology and artificial intelligence, are imperative for the rapid and efficient detection of illegal cannabis plantations, particularly in extensive areas.

An herbal plant originating from Central Asia, *Cannabis sativa L.* or marijuana, has been widely used for medicinal purposes and as a source of textile materials (Andre et al. 2016). Cannabis, which has a wide range of applications, grows well with little maintenance as long as it gets enough sunlight and water. These characteristics have led to cannabis' rapid globalization and increasing attention.

Cannabis is an annual tropical plant; it can be harvested 4-6 months after planting. Its fingered leaves can grow up to 2 m, and it thrives at an elevation of 1000 meters above sea level. There are 3 species of cannabis plants (*Cannabis sativa L., Cannabis indica, and Cannabis rudelis*) with different tetrahydrocannabinol (THC) contents. This level of THC in cannabis can harm mental health and the nervous system (Karst 2018). It may make users more tolerant to its negative effects, lead to dependence, and cause withdrawal symptoms. (Lim et al. 2017; Wilkinson et al. 2016).

[1] Research Data and Information Center. (2022). Indonesia Drugs Report 2022. Jakarta: National Narcotics Board of Indonesia.

Due to the negative effects of cannabis, some countries, including Indonesia, have banned its use. Law enforcement has put a lot of effort into eliminating the distribution and the source of cannabis. The operations to eradicate cannabis plantations carried out by law enforcers have not been effective; there is still a high number of cannabis users' prevalence. Using technology, such as remote sensing technology, to detect these illegal plantations will help law enforcement operate more efficiently to eradicate cannabis plantations.

Various classification methods based on optical remote sensing and SAR satellite data are used to monitor forest dynamics (Armenteras et al. 2013; Diwediga et al. 2017; Sirro et al. 2018), extract LULC information (Chen and Tsou 2021), calculate agricultural land area for harvest prediction (Dahal et al. 2018; Duraisamy et al. 2018; Elbasi et al. 2023), and develop urban land cover mapping (Tavares et al. 2019). Leveraging information based on remote sensing technology, policymakers can improve the decision-making process.

Phenology, a dynamic feature of vegetation, can help differentiate plants from others (Kustiyo et al. 2024). Remote sensing satellite data can estimate phenology by observing vegetation growth and leaf cross-section (Kosczor et al. 2022; Yang et al. 2023). To avoid the complexity of multispectral data, plant greenness indices like Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI), can be used to analyze phenology (Huang et al. 2021; Jimenez et al. 2022; Suárez et al. 2022). These indices facilitate the classification of landscapes, including agriculture, vegetation, land use, water bodies, and forests (Choudhary et al. 2019; Zuo et al. 2019). This information can also be used in decision-making to determine plant age (Nieto et al. 2021).

The effectiveness of artificial intelligence in object identification using remote sensing data has been demonstrated in numerous studies. Classifying hyperspectral airborne data, Support Vector Machine (SVM) yields better accuracy compared to Random Forest (RF) to identify vegetation species that have similar spectral with their surroundings for heterogeneous areas, but RF yields better accuracy for more homogeneous area (Sabat-Tomala et al. 2020). Qian et al. (2015) stated that SVM and Normal Bayes (NB) are more accurate than CART and KNN for identifying land cover from high-resolution satellite images. An ensemble of CNNs with different architectures applied to various image representations functions as an effective feature extractor used to detect cannabis plantations (Ferreira et al. 2019). Sujud et al. (2021) argue that precision accuracy for classification methods depends on various factors; no one can claim that one classifier performs better than others.

Previous research has detected illegal cannabis plantations carried out in China, Afghanistan, and several South American countries, with one thing in common, namely that the average area is quite large, 3 to 9 ha (Sujud et al. 2021) or even more. Detection of cannabis plantations using Landsat 8 NDVI data (Mattiuzzi et al. 2014), IKONOS high-resolution data and deep learning (Ferreira et al. 2019), optical and SAR satellite data fusion, and machine learning (Sujud et al. 2021), as well as high-resolution data (QuickBird and SPOT 5) and machine learning (Bicakli et al. 2022) produces quite high accuracy values of more than 90%.

Unique features of cannabis plantations in Indonesia, such as small, un-irrigated areas, forested areas, and high cloud coverage, pose challenges for building identification models using medium-resolution optical satellite data. This research aims to develop a cannabis plantation detection system applicable in Indonesia by utilizing multitemporal medium optical data, high-resolution satellite data, and other geographic features. By compiling a valid dataset, a model to detect cannabis plantations can be built by leveraging a fast machine learning classifier. The result of this research will benefit law enforcement for efficient and effective detection of illegal cannabis plantations and provide a cannabis plant dataset in Indonesia for future research.

This research uses adjusted satellite data from the Sentinel 2A and 2B missions, collected between 2021 and 2022. The data comes from different sources, resolutions, and times. High-resolution satellite imagery and a database of cannabis eradication operations provided by law enforcement agencies are also used for validation purposes. Aceh Besar District in Nangroe Aceh Darussalam (NAD) will be the focal point of this research.

MATERIALS AND METHODS

Study Area

Although the cultivation of cannabis is prohibited under Indonesia's Narcotics Law of 2009, the practice persists, particularly in NAD[2]. This has prompted law enforcement agencies to conduct frequent eradication operations aimed at eliminating cannabis plantations[3]. In 2023, BNN together with BRIN (The National Research and Innovation Agency), BIG (The Indonesian Mapping Agency) and local authorities eradicated cannabis plantation in Aceh Besar District, NAD Province, as shown in Fig. 1a. This area was then selected as a study location due to the availability of field data. Field information, serving as secondary data was made available by law enforcement agencies, containing locations and timestamps of eradication activities in spreadsheet form.

Remote sensing data

Leveraging 5-day temporal resolution, the Sentinel-2 constellation[4] is used in this research to obtain cloud-free data (Fig. 1b). The Sentinel-2 data, acquired between October 2021 and February 2022, comprise 17 selected dates and were utilized to detect cannabis plantations. Sentinel-2 data from March to October 2022 is used to assess the condition of these plantations. The satellite data used includes three visible light spectrum bands and one near-infrared band, all with a spatial resolution of 10 meters.

The eradication site locations for this research were validated using RGB composite imagery from the WorldView-3 satellite by Maxar Technologies (Fig. 1c), available on Google Earth. Acquired in October 2024, this imagery offers a spatial resolution of 31 cm. It adds to the main dataset, which consists of Sentinel-2 images taken over nine months. These images include data in the red, green, blue, and near-infrared bands (bands 2, 3, 4, and 8).

A total of 284 scenes of data were accessible via the Sentinel portal data hub, part of the Copernicus program[5]. These datasets, which have undergone radiometric and

---

[2] Kompas, (2023). Ladang Ganja Masih Ditemukan di Aceh, Kali Ini Seluas 2,5 Hektar. [online] Available at: regional.kompas.com/read/2023/02/23/091659278/ladang-ganja-masih-ditemukan-di-aceh-kali-ini-seluas-25-hektar [Accessed 11 Dec. 2024].

[3] https://www.thejakartapost.com/indonesia/2024/05/22/aceh-authorities-to-ban-marijuana-laced-dishes-in-restaurants.html [Accessed 11 Dec. 2024]

[4] Thales Alenia Space. 2021. Sentinel-2 Products Specification Document (PSD).

[5] https://dataspace.copernicus.eu/browser/

(a) Study Location   (b) Sentinel-2   (c) WorldView-3
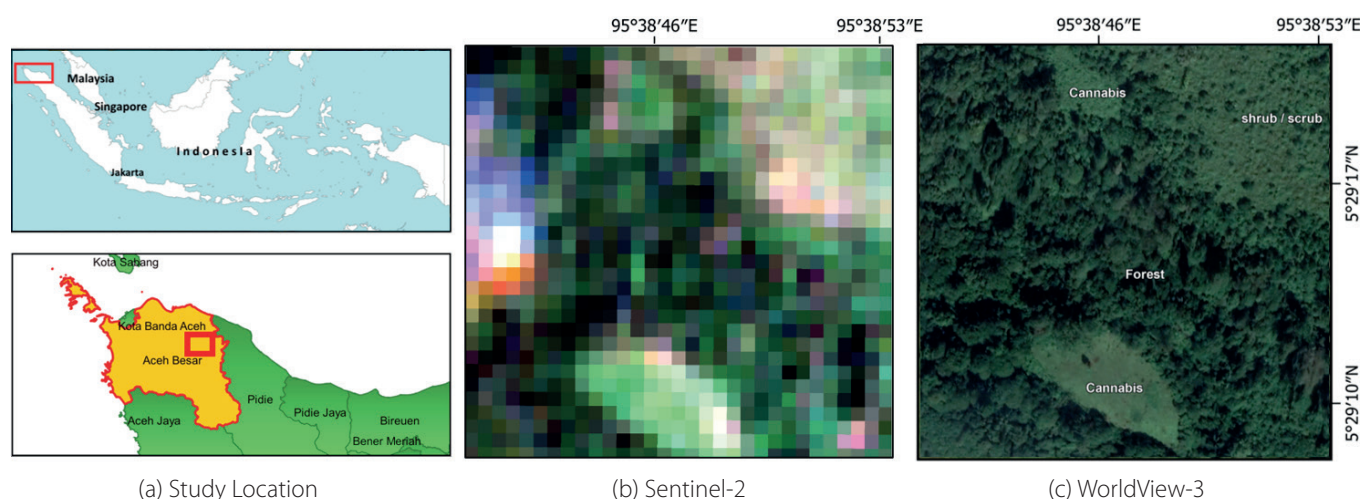
Fig. 1. (a) Study location with (b) input data and (c) reference image

geometric corrections, serve as the primary data for further processing. High-resolution satellite data was provided by BRIN (The National Research and Innovation Agency) and BIN (The National Intelligence Agency), while the base map was provided by BIG (The Indonesian Mapping Agency).

Methodology

The study aimed to develop a cannabis plantation detection system based on digital data processing utilizing remote sensing satellite data and machine learning classifiers. Additionally, the study sought to determine the condition of cannabis plants by analyzing plant growth phenology. The identification of illicit cannabis plantations within forested regions can be effectively achieved through a supervised classification approach employing machine learning classifiers. This method leverages the statistical features of plant growth phenology derived from multi-temporal optical satellite imagery, addressing challenges associated with high cloud cover and varying phenological stages of cannabis plants. Furthermore, the condition of cannabis plants was determined by analyzing their growth phenology based on vegetation index values from multi-temporal remote sensing satellite data.

Cannabis cultivation necessitates direct sunlight (Morello et al. 2022) and adequate water supply (Dillis et al.

2020) for optimal growth. This dependency makes cannabis plantations visible from an aerial perspective and often results in their location near water sources. Consequently, medium-resolution satellite data serves as an ideal tool for locating cannabis plantations, as remote sensing satellites periodically pass over and collect data from areas on the Earth's surface.

Fig. 2 illustrates the flow diagram of the proposed data processing model for utilizing remote sensing to detect cannabis plantations and assess their conditions. The model utilized medium-resolution optical remote sensing data from the Sentinel-2 constellation, covering the period from 2019 to 2022.

The model development involved the following stages:

1. **Data Preparation:** This involves data correction for the satellite data, ensuring it has undergone necessary level corrections. Commonly used geometric and radiometric correction methods are applied. This research used Level-2A, radiometrically and geometrically corrected image products provided by the Sentinel Hub. The Level-2A data is an ortho-image product that is projected using a Digital Elevation Model (DEM) to correct geometric distortions. Bottom-Of-Atmosphere (BOA) radiometric correction was applied to the data, followed by basic pixel features classification (including cloud types identification)[6]. A manual selection process was employed to identify cloud-
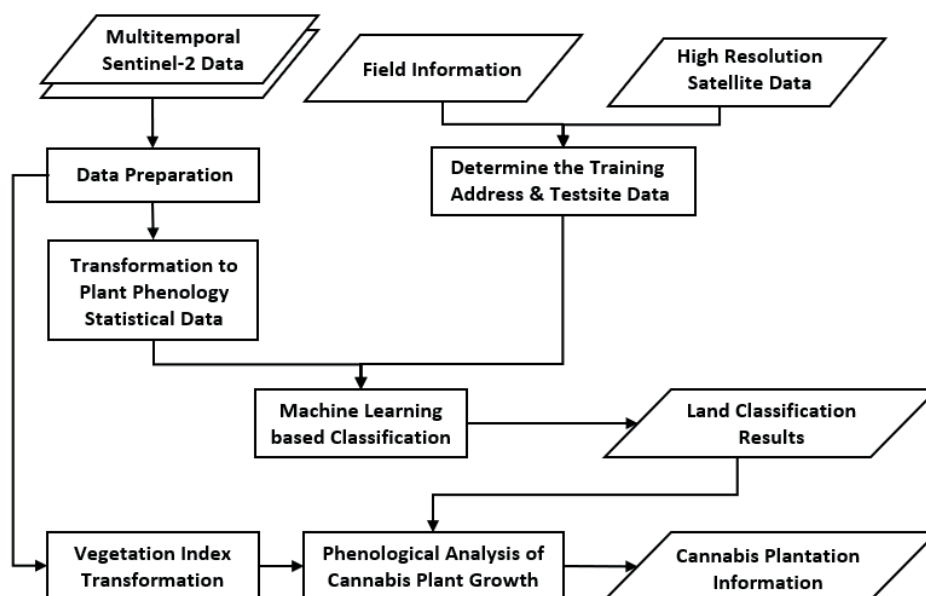

Fig. 2. Proposed model of smart cannabis plantations detection

[6] Thales Alenia Space. 2021. Sentinel-2 Products Specification Document (PSD).

free, usable data from the four-band multi-temporal Sentinel-2 satellite datasets spanning from 2021 to 2022. This process resulted in the identification of 17 usable datasets for the designated research area.

2. **Transformation of Multitemporal Satellite Data:** Data transformation aims to extract statistical parameters that consist of the minimum, maximum, average, and standard deviation of each pixel position as new datasets. Statistical parameters reflect the plant growth phenology and other land characteristics over a period of time (6 months before and 3 months after the plant destruction). NDVI was calculated from the red and NIR bands, and then NDVI along with the red, green, blue, and NIR bands is transformed into statistical parameters.

3. **Determination of Training and Test Data:** Cannabis dataset, derived from eradication activities provided by Indonesian law enforcement agencies, was validated by high-resolution remote sensing data. The locations (longitude and latitude) of all eradication activities were selected as the centers of spatial datasets. Data from the available multi-temporal Sentinel-2 constellation data were cropped into 32×32 pixels of the dataset. A manual selection was employed to choose cloud-free, usable data. Statistical features (minimum, maximum, median, mean, and standard deviation) were extracted from pixels that contained cannabis information. This dataset was subsequently used to train and test the proposed model's accuracy. Training sample data for machine learning and test site data for accuracy assessments were determined based on field data or location points where authorities have eradicated cannabis plantations. These points were verified or interpreted using high-resolution satellite imagery. Various land classes, such as cannabis plantations, forests, and bush areas, were identified based on satellite data interpretation and field data.

4. **Machine Learning based Classification for Cannabis Detection:** Machine learning-based detection of cannabis plantation was conducted using statistical features of growth phenology and other objects from multitemporal satellite data, along with the training data. The neural network architecture follows the design proposed by Kushardono et al. (1995a) illustrated in Fig. 3, consisting of three layers: an input layer, a hidden layer, and an output layer. This three-layer configuration offers faster learning rates and higher classification accuracy compared to four-layer architectures.

Input layer contains neurons corresponding to the input data volume, which in this research involves 16 neurons (calculated from 4 statistical features × 4 spectral bands). The hidden layer, based on the optimal configuration identified by Kushardono et al. (1995a), contains twice the number of neurons in the output layer. Deviations from this optimal configuration can prolong the learning process, increase the risk of convergence failure, and reduce classification accuracy. The output layer's neuron count equals the number of class categories. Optimum training parameters for backpropagation, as suggested by Kushardono et al. (1995a), include a learning rate of 0.9 and a momentum rate of 0.1. Non-optimal parameter combinations may result in a higher number of iterations or difficulties in achieving convergence to a target error rate below 0.1.

5. **Land Classification Results:** The target area was classified into four land classes using the trained neural networks, with input feature data representing land phenology statistical values derived from the optical satellite data. The classification accuracy was then evaluated using a confusion matrix through overlapping analysis with test site data. The overall accuracy was found by dividing the number of correct classifications (shown on the main diagonal of the confusion matrix) by the total number of data points in the matrix. Additionally, the kappa coefficient was used to evaluate the agreement between observed and expected classifications in a confusion matrix, considering chance agreement. This coefficient provides a robust measure of classification performance, particularly useful for imbalanced classes. Kappa is calculated by comparing observed accuracy with expected chance agreement and ranges from -1 (no agreement) to +1 (perfect agreement), with higher values indicating better classifier performance (Congalton and Green 2009).

6. **Cannabis Plantation Information:** To identify illegal cannabis plantations in forested areas, particularly in the context of extensive satellite image coverage, classification results were filtered using a forest area map. The assessment of cannabis growth stages—ranging from vegetative growth, flowering, and nearing harvest to fallow fields—was conducted through phenological analysis of vegetation index values derived from multispectral optical satellite data.
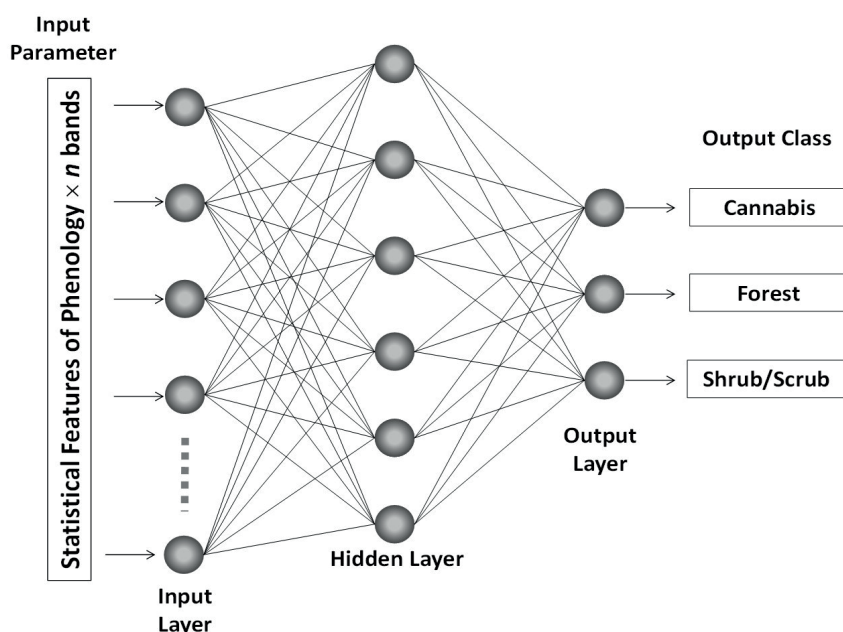

Fig. 3. Neural network architecture

RESULTS

The use of statistical features on plant growth phenology, such as minimum, maximum, average, and standard deviation from each band of multitemporal satellite data, shows good results for training classifiers. Using 3 types of training samples for 8 training areas for the cannabis plantation class, 11 training areas for the forest class, and 12 training areas for the shrub class yields the fast-decreasing tendency of the root mean square (RMS) values in the training process of a backpropagation neural network, indicating fast convergence. It shows that the RMS values decrease to less than 0.1 or 10% after 100 iterations and 0.01 or 1% after 776 iterations, as illustrated in Fig. 4.

Using the backpropagation neural network method, the classification of statistical parameters capturing growth phenology from multi-temporal Sentinel-2 data bands yields de-limitation results for three distinct classes: cannabis plantations (represented in red), forests (green), and shrubs (yellow), as illustrated in Fig. 5a.

A comparative analysis between classification results derived from 10-meter spatial resolution Sentinel-2 data and higher spatial resolution satellite imagery from October 2022, illustrated in Fig. 5, reveals substantial similarity in the spatial distribution and characteristics of land features,

encompassing 2 cannabis plantations, forest vegetation, and bush vegetation. Accuracy tests using data from 254 sample points (see Table 1) show a high accuracy rate of 95% and a kappa score of 0.92.

In a comparative analysis, the same training samples and input features utilized in the back-propagation neural network are applied to an SVM (Support Vector Machine) classifier, following the same settings mentioned in Park et al. (2018). The classification results in Fig. 4b show that the SVM classifier has an accuracy of 94% and a kappa coefficient of 0.91, as seen in Table 2. However, the accuracy results obtained from this classification process are not statistically significant. While the SVM classifier demonstrates a comparable accuracy level, Fig.5b indicates the presence of noise, particularly along the boundaries. This suggests that the SVM classifier does not perform well with training samples that contain noise, such as overlaps between the cannabis class and the bush/scrub class, as well as between the shrub class and the forest class. Conversely, the backpropagation neural network, functioning as a non-parametric classifier, can iteratively refine the training process until convergence, thereby minimizing errors and determining the class based on the highest neuron output value. Consequently, due to the encountered classification errors, particularly noise-related issues, the SVM classifier is not recommended for accurately determining the location of cannabis plantations.
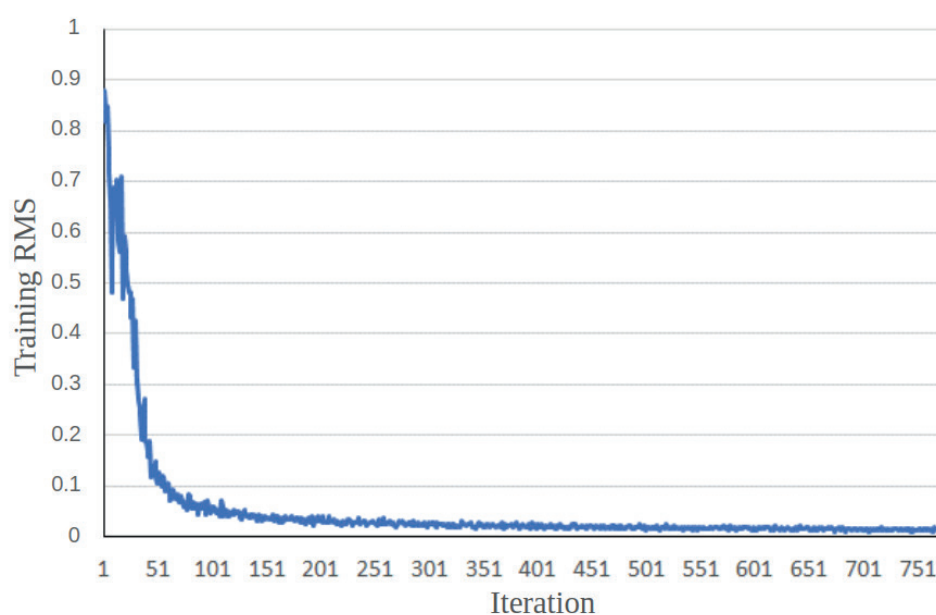
Fig. 4. RMS training reduction in backpropagation neural network learning iterations

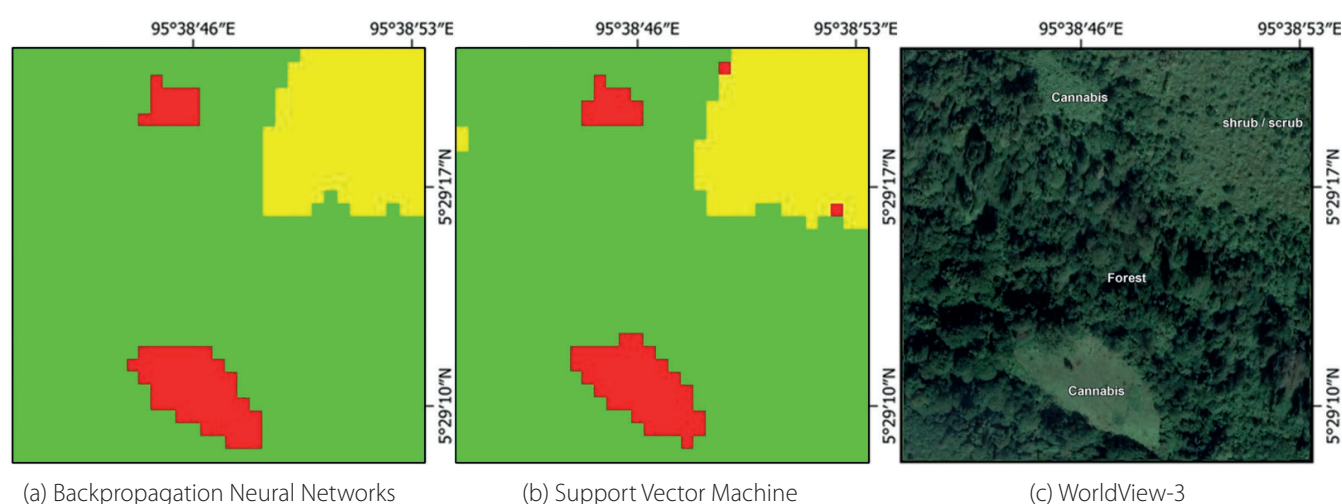(a) Backpropagation Neural Networks (b) Support Vector Machine (c) WorldView-3

Fig. 5. Classification result (red: cannabis; green: forest; yellow: shrub) based on (a) backpropagation Neural Networks and (b) Support Vector Machine, compared with (c) WorldView-3 data October 2022

**Table 1. Accuracy of Neural Network-based classification results for cannabis and non-cannabis
(Overall Accuracy = 94.88%; Kappa Coefficient = 0.92)**

Class	Ground Truth (Pixels)			
	Cannabis Plantations	Forest	Shrub / Scrub	Total
Unclassified	0	0	0	0
Cannabis Plantations	74	0	0	74
Forest	13	90	0	103
Shrub / Scrub	0	0	77	77
Total	87	90	77	254

**Table 2. Accuracy of SVM-based classification results for cannabis and non-cannabis
(Overall Accuracy = 94.09%; Kappa Coefficient = 0.91)**

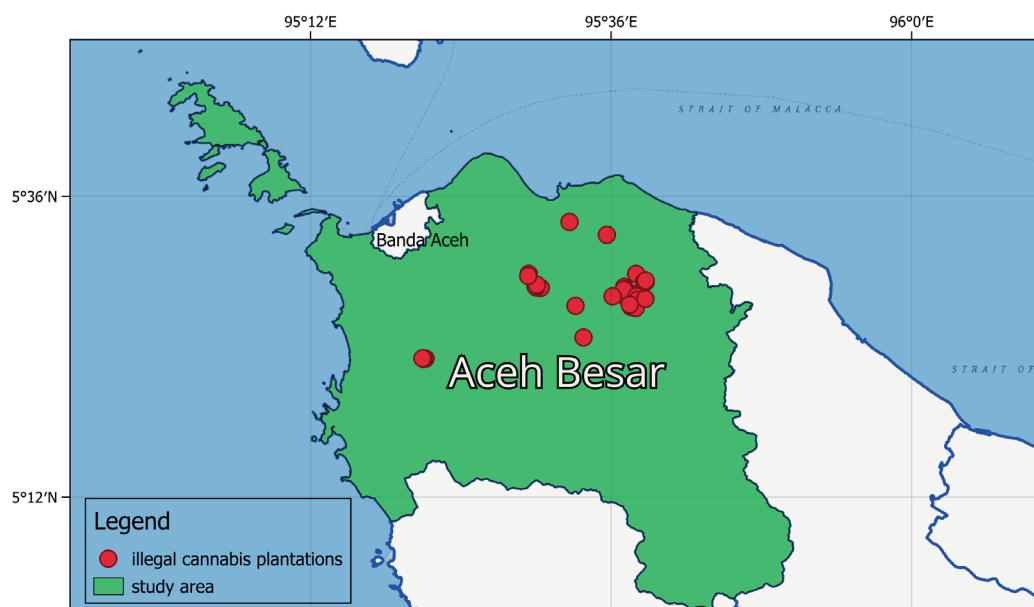| Class	Ground Truth (Pixels)			
	Cannabis Plantations	Forest	Shrub/ Scrub	Total
Unclassified	0	0	0	0
Cannabis Plantations	74	0	0	74
Forest	13	88	0	101
Shrub/ Scrub	0	2	77	79
Total	87	90	77	254

Fig. 6. Distribution of illegal cannabis plantations in Aceh Besar District

Fig. 6 illustrates several identified locations of illegal cannabis plantations in Aceh Besar District, NAD. All identified locations are situated in remote, mountainous, and forested areas. This study demonstrates that the integration of machine learning with medium-resolution remote sensing data is a reliable and effective approach for identifying cannabis plantations with high accuracy (>90%). It provides significant and effective support to law enforcement efforts in Indonesia in combating drug distribution.

Following the identification of cannabis plantation locations in the aforementioned classification results, an assessment of cannabis plantation conditions is conducted based on the vegetation index (NDVI) values derived from multi-temporal Sentinel-2 satellite data collected between March and October 2022. A phenology graph illustrating the growth stages of the cannabis plants is generated, as illustrated in Fig. 7 According to Desjardins (2018)[7], cannabis typically undergoes an initial vegetative growth phase lasting up to 8 weeks, followed by a flowering phase lasting 6-8 weeks and culminating in a harvesting phase. The NDVI values graph reveals a discernible pattern that correlates with the predicted age of cannabis plants, as depicted in Fig. 7. Understanding the NDVI pattern graph enables law enforcement agencies to estimate the condition and growth stage of cannabis within plantations, thereby enhancing the capacity to detect and eradicate illegal cannabis cultivation activities.

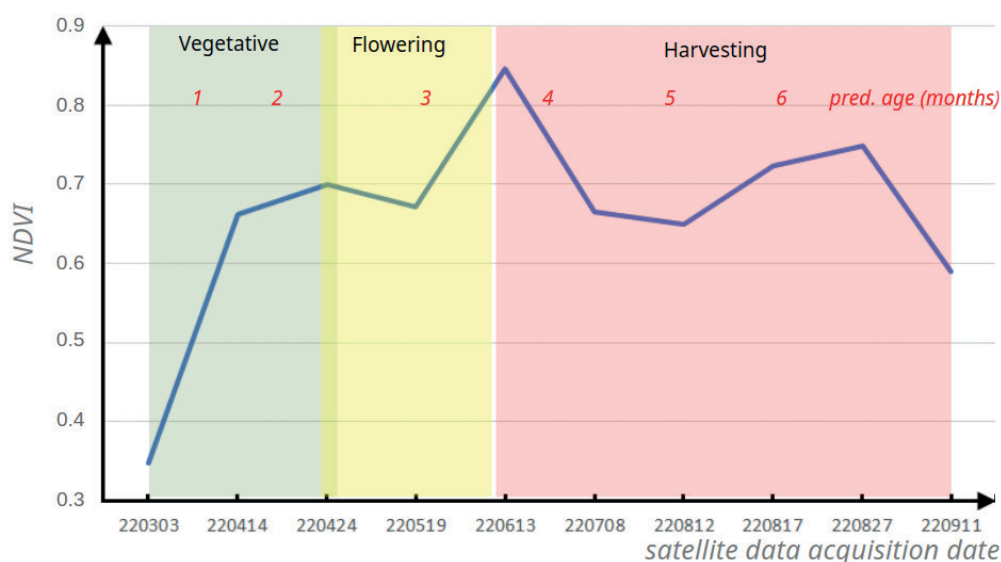[7] Desjardins, Jeff. (2018). The Anatomy of a Cannabis Plant, and Its Lifecycle. 1–9. Retrieved December 20, 2023 (https://www.visualcapitalist.com/anatomy-cannabis-plant/).

Fig. 7. Graph of NDVI values in cannabis plantation and prediction of the age of cannabis plants

DISCUSSION

The suggested model, which uses medium-resolution Sentinel-2 data, demonstrates high overall accuracy and kappa values by applying two different classifiers to distinguish between three categories: cannabis, forest, and shrub. Specifically, neural network-based classification achieves an accuracy of 94.88% and a kappa value of 0.92, while SVM-based classification yields an accuracy of 94.09% and a kappa value of 0.91. These results are comparable to prior findings. For example, Ferreira et al. (2019) found that high-resolution IKONOS images could help identify cannabis farms using a CNN classifier with an accuracy of 92.16%. Similarly, Sujud et al. (2021) found that an SVM classifier provided the best overall accuracy of more than 90%, compared to RF, GBT, and CART, when using a combination of medium-resolution Sentinel-1, Sentinel-2, and Landsat-8 data for detecting cannabis plantations. Bicakli et al. (2022) argued that an RF classifier with high-resolution Planet data yielded a better accuracy of 93% compared to Decision Tree (75.2%), Random Tree (87.1%), KStar (91.0%), and IDK (92.1%).

Previous research identified backpropagation neural networks as a preferred non-parametric classification method over conventional parametric approaches like Maximum Likelihood and Dempster-Shafer's (Kushardono et al. 1995b). Backpropagation is one of the best algorithms for training multilayer perceptron networks due to its ability to minimize errors and efficiency in pattern classification (Suliman and Zhang 2015). Neural networks excel in capturing complex non-linear relationships, handling spectral variations, mixed pixels, radiometric inconsistencies, and limited training data—challenges common in cannabis fields located on tropical hillsides with high cloud cover and small spatial extent. Guo et al. (2012) found that Support Vector Machines (SVMs) achieve slightly higher classification accuracy than neural networks but face limitations, including reliance on kernel selection, difficulty with minimal training data, and challenges in finding an optimal separating hyperplane. Future research should explore spatio-temporal features using Convolutional Neural Networks (CNNs) to better process multitemporal, high-resolution satellite remote sensing data.

Furthermore, this research found that SVM-based classification exhibits higher levels of noise compared to neural network-based classification. SVM appears to struggle with noisy training samples, such as pixels containing two classes. In contrast, the backpropagation neural network, works as a non-parametric classifier by repeatedly improving its training to reduce mistakes. This research demonstrates that free-access satellite data, specifically from the Sentinel-2 constellation, can be effectively utilized to detect the potential location of illicit crops, such as cannabis, with high accuracy over small areas. This approach reduces reliance on high-resolution commercial satellite data. The results indicate that with an appropriate model, it is possible to extract challenging and sensitive information from freely available satellite data.

The availability of cloud-free Sentinel-2 data over NAD is limited, with only 198 out of 1708 Sentinel-2 observations between 2016 and 2022 having less than 20% cloud cover. Despite this limitation, this research successfully utilized low-cloud-cover data with usable information from locations documented in a law enforcement eradication database, focusing on the cannabis growth period. NDVI analysis revealed distinct patterns compared to surrounding vegetation (Desjardins 2018). Key statistical features (minimum, maximum, median, mean, and standard deviation) were extracted to enhance machine learning classification. Areas obscured by clouds could not be analyzed, but future studies may address this limitation using SAR data.

Cannabis plants undergo distinct growth phases: fallow, seedling, vegetative, flowering, and harvest. Temporal satellite data captures these phases as changes in vegetation greenness, with digital values starting low before planting, rising during the vegetative phase, and falling at harvest. These patterns differ from other land cover types around them. Phenological changes are analyzed using statistical features for each land type, including cannabis, to account for planting time variations. Illegal cannabis, an annual crop grown in remote areas, exhibits phenological statistics that contrast sharply with surrounding land types—for instance, consistently high greenness in forests and shrubs versus frequently low greenness in open land and water bodies.

Illegal cannabis plantations in Aceh (NAD) and other regions are typically found in remote hilly areas. Growth phenology statistics for the same cannabis species show little variation across locations. Differences in soil fertility and rainfall have minimal impact on these features when plantation management is similar. Neural network classifiers effectively mitigate such variations due to their robustness to satellite data noise. This model will work for homogenous cannabis plantations. Field report show that

homogeneous cannabis plantations can be found other districts in NAD, such as: North Aceh District, South Aceh District, Gayo Lues District, etc. Mandailing Natal District in North Sumatera also exhibits this phenomenon.

This research demonstrates that with good data sampling, a combination of optical medium-resolution remote sensing data and machine learning classifiers can be effectively utilized to detect potential locations of illicit crops such as *Cannabis sativa L*. The conditions of cannabis plants from potential locations determined from the proposed method can be determined using vegetation indices such as NDVI or EVI. This model offers a rapid, efficient, and low-cost approach to combating illegal cannabis cultivation. This model might be unsuitable for intercropping plantations, as it was developed using datasets from monoculture cannabis plantations.

CONCLUSION

This research presents a method for processing digital data to find an evaluate illegal cannabis plantations. It uses statistical analysis of the growth pattern of cannabis plants, which are obtained from satellite images taken over time. Additionally, this research has been successful in substantially enhancing the efficiency and speed of detecting illegal cannabis cultivation, thereby strengthening law enforcement efforts in Indonesia. Illegal cannabis plantations in forests can be spotted by classifying data based on the growth patterns of plants, using information gathered from Sentinel-2 satellite images taken at different times. The classification results have an accuracy of about 95%, and the backpropagation

neural network classifier can accurately locate cannabis farms. Assessment of cannabis plantation conditions, including estimation of plant age, is enabled through phenological analysis of cannabis growth patterns using vegetation index values derived from multi-temporal Sentinel-2 satellite data, aligned with established growth stages of cannabis plants.

Our research shows that using statistical data from multi-temporal remote sensing data in a machine learning classifier is the most effective method for detecting illegal cannabis crops in Indonesia, especially in mitigating cloud cover. Using medium-resolution satellite data to detect potential illegal cannabis plantations reduces reliance on costly high-resolution commercial data.

High cloud cover in Indonesia limits this research, despite its numerous successes. This can hinder the use of optical remote sensing satellites for detecting illegal cannabis plantations, as it can obscure the study area. However, under conditions of thin cloud cover, we can still identify the land cover changes and predict the cannabis growth stages. Other machine learning classifiers can extend the model and apply it to regions with similar geographical characteristics. And to enhance the accuracy of identifying land suitable for the cultivation of cannabis plants, the use of multi-temporal Synthetic Aperture Radar (SAR) data texture features, which are unaffected by cloud cover, can be implemented. Additionally, the application of cross-validation techniques can improve the stability and reliability of the accuracy assessment model. Future research could focus on the fusion of data and/or information to enrich feature sets and further enhance predictive accuracy. ◼

REFERENCES

Andre, C. M., Hausman, J. F., and Guerriero, G. (2016). Cannabis Sativa: The Plant of the Thousand and One Molecules. Frontiers in Plant Science 7:1–17. doi: 10.3389/fpls.2016.00019.

Armenteras, D., Cabrera, E., Rodríguez, N., and Retana, J. (2013). National and Regional Determinants of Tropical Deforestation in Colombia. Regional Environmental Change 13(6):1181–93. doi: 10.1007/s10113-013-0433-7.

Bicakli, F., Kaplan, G., and Alqasemi, A. S. (2022). Cannabis Sativa L. Spectral Discrimination and Classification Using Satellite Imagery and Machine Learning. Agriculture (Switzerland) 12(6). doi: 10.3390/agriculture12060842.

Chen, F., and Tsou, J. Y. (2021). DRSNet: Novel Architecture for Small Patch and Low-Resolution Remote Sensing Image Scene Classification. International Journal of Applied Earth Observation and Geoinformation 104:102577. doi: 10.1016/j.jag.2021.102577.

Choudhary, K., Shi, W., Boori, M. S., and Corgne, S. (2019). Agriculture Phenology Monitoring Using NDVI Time Series Based on Remote Sensing Satellites: A Case Study of Guangdong, China. Optical Memory and Neural Networks (Information Optics) 28(3):204–14. doi: 10.3103/S1060992X19030093.

Congalton, R. G., and Green, K. (2009). Assessing the Accuracy of Remotely Sensed Data: Principles and Practices. Vol. 25. Second. Florida: CRC Press.

Dahal, D., Wylie, B., and Howard, D. (2018). Rapid Crop Cover Mapping for the Conterminous United States. Scientific Reports 8(1):1–13. doi: 10.1038/s41598-018-26284-w.

Dillis, C., McIntee, C., Butsic, V., Le, L., Grady, K., and Grantham, T. (2020). Water Storage and Irrigation Practices for Cannabis Drive Seasonal Patterns of Water Extraction and Use in Northern California. Journal of Environmental Management 272:110955. doi: 10.1016/j.jenvman.2020.110955.

Diwediga, B., Agodzo, S., Wala, K., and Le, Q. B. (2017). Assessment of Multifunctional Landscapes Dynamics in the Mountainous Basin of the Mo River (Togo, West Africa). Journal of Geographical Sciences 27(5):579–605. doi: 10.1007/s11442-017-1394-4.

Duraisamy, V., Bendapudi, R., and Jadhav, A. (2018). Identifying Hotspots in Land Use Land Cover Change and the Drivers in a Semi-Arid Region of India. Environmental Monitoring and Assessment 190(9). doi: 10.1007/s10661-018-6919-5.

Elbasi, E., Zaki, C., Topcu, A. E., Abdelbaki, W., Zreikat, A. I., Cina, E., Shdefat, A., and Saker, L. (2023). Crop Prediction Model Using Machine Learning Algorithms. Applied Sciences (Switzerland) 13(16). doi: 10.3390/app13169288.

Ferreira, A., Felipussi, S. C., Pires, R., Avila, S., Santos, G., Lambert, J., Huang, J., and Rocha, A. (2019). Eyes in the Skies: A Data-Driven Fusion Approach to Identifying Drug Crops from Remote Sensing Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 12(12):4773–86. doi: 10.1109/JSTARS.2019.2917024.

Guo, Y., De Jong, K., Liu, F., Wang, X., and Li, C. (2012). A Comparison of Artificial Neural Networks and Support Vector Machines on Land Cover Classification. In Z. Li, X. Li, Y. Liu, and Z. Cai (Eds.), Computational Intelligence and Intelligent Systems 316:531-539. ISICA. doi: 10.1007/978-3-642-34289-9_59.

Huang, S., Tang, L., Hupy, J. P., Wang, Y., and Shao, G. (2021). A Commentary Review on the Use of Normalized Difference Vegetation Index (NDVI) in the Era of Popular Remote Sensing. Journal of Forestry Research 32(1):1–6. doi: 10.1007/s11676-020-01155-1.

Jimenez, R. B., Lane, K. J., Hutyra, L. R., and Fabian, M. P. (2022). Spatial Resolution of Normalized Difference Vegetation Index and Greenness Exposure Misclassification in an Urban Cohort. Journal of Exposure Science and Environmental Epidemiology 32(2):213–22. doi: 10.1038/s41370-022-00409-w.

Karst, A. 2018. Weighing the Benefits and Risks of Medical Marijuana Use: A Brief Review. Pharmacy 6(4):128. doi: 10.3390/pharmacy6040128.

Kosczor, E., Forkel, M., Hernández, J., Kinalczyk, D., Pirotti, F., and Kutchartt, E. (2022). Assessing Land Surface Phenology in Araucaria-Nothofagus Forests in Chile with Landsat 8/Sentinel-2 Time Series. International Journal of Applied Earth Observation and Geoinformation 112. doi: 10.1016/j.jag.2022.102862.

Kushardono, D., Fukue, K., Shimoda, H., and Sakata, T. (1995a). Optimized Neural Network for Spatial Land Cover Classification with the Aid of Co Occurrence Matrix. Journal of the Japan Society of Photogrammetry and Remote Sensing 34(4):22–35. doi: 10.4287/jsprs.34.4_22.

Kushardono, D., Fukue, K., Shimoda, H., and Sakata, T. (1995b). Comparison of multi-temporal image classification methods. 1995 International Geoscience and Remote Sensing Symposium, IGARSS '95. Quantitative Remote Sensing for Science and Applications, 1282–1284. doi: 10.1109/IGARSS.1995.521726

Kustiyo, K., Rokhmatuloh, R., Saputro, A. H., and Kushardono, D. (2024). Rice Fields Classification through Spectral-Temporal Data Fusion during the Rainy and Dry Seasons Using Sentinel-2 Optical Images in Subang Regency, West Java, Indonesia. Paddy and Water Environment 22(3):375–385. doi: 10.1007/s10333-024-00972-y.

Lim, K., See, Y. M., and Lee, J. (2017). A Systematic Review of the Effectiveness of Medical Cannabis for Psychiatric, Movement and Neurodegenerative Disorders. Clinical Psychopharmacology and Neuroscience 15(4):301–312. doi: 10.9758/cpn.2017.15.4.301.

Mattiuzzi, M., Bussink, C., and Bauer, T. (2014). Analysing Phenological Characteristics Extracted from Landsat NDVI Time Series to Identify Suitable Image Acquisition Dates for Cannabis Mapping in Afghanistan. Photogrammetrie, Fernerkundung, Geoinformation 2014(5):383–392. doi: 10.1127/1432-8364/2014/0231.

Morello, V., Brousseau, v. D., Wu, N., Wu, B. S., MacPherson, S., and Lefsrud, M. (2022). Light Quality Impacts Vertical Growth Rate, Phytochemical Yield and Cannabinoid Production Efficiency in Cannabis Sativa. Plants 11(21). doi: 10.3390/plants11212982.

Nieto, L., Schwalbert, R., Prasad, P. V. V., Olson, B. J. S. C., and Ciampitti, I. A. (2021). An Integrated Approach of Field, Weather, and Satellite Data for Monitoring Maize Phenology. Scientific Reports 11(1):1–10. doi: 10.1038/s41598-021-95253-7.

Park, S., Im, J., Park, S., Yoo, C., Han, H., and Rhee, J. (2018). Classification and Mapping of Paddy Rice by Combining Landsat and SAR Time Series Data. Remote Sensing 10(3):1–22. doi: 10.3390/rs10030447.

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(1):153–168. doi: 10.3390/rs70100153.

Sabat-Tomala, A., Raczko, E., and Zagajewski, B. (2020). Comparison of Support Vector Machine and Random Forest Algorithms for Invasive and Expansive Species Classification Using Airborne Hyperspectral Data. Remote Sensing 12(3). doi: 10.3390/rs12030516.

Sirro, L., Häme, T., Rauste, Y., Kilpi, J., Hämäläinen, J., Gunia, K., de Jong, B., and Pellat, F., P. (2018). Potential of Different Optical and SAR Data in Forest and Land Cover Classification to Support REDD+ MRV. Remote Sensing 10(6). doi: 10.3390/rs10060942.

Suárez, H. L. A., Angarita, G. P. G., Castañeda, L. N. R., and Castro, P. P. C. (2022). Illicit Crops, Planning of Substitution with Sustainable Crops Based on Remote Sensing: Application in the Sierra Nevada of Santa Marta, Colombia. Climate Emergency – Managing, Building , and Delivering the Sustainable Development Goals 483–494. doi: 10.1007/978-3-030-79450-7_36.

Sujud, L., Jaafar, H., Hassan, M. A. H., and Zurayk, R. (2021). Cannabis Detection from Optical and RADAR Data Fusion: A Comparative Analysis of the SMILE Machine Learning Algorithms in Google Earth Engine. Remote Sensing Applications: Society and Environment 24. doi: 10.1016/j.rsase.2021.100639.

Suliman, A., and Zhang, Y. (2015). A Review on Back-Propagation Neural Networks in the Application of Remote Sensing Image Classification. Journal of Earth Science and Engineering 5:52–65. doi: 10.17265/2159-581X/2015.

Tavares, P. A., Beltrão, N. E. S., Guimarães, U. S., and Teodoro, A. C. 2019. Integration of Sentinel-1 and Sentinel-2 for Classification and LULC Mapping in the Urban Area of Belém, Eastern Brazilian Amazon. Sensors (Switzerland) 19(5). doi: 10.3390/s19051140.

Wilkinson, S. T., Yarnell, S., Radhakrishnan, R., Ball, S. A., and D'souza, D. C. (2016). Marijuana Legalization: Impact on Physicians and Public Health. Annual Review of Medicine 67:453–466. doi: 10.1146/annurev-med-050214-013454.

Yang, J., Dong, J., Liu, L., Zhao, M., Zhang, X., Li, X., Dai, J., Wang, H., Wu, C., You, N., Fang, S., Pang, Y., He, Y., Zhao, G., Xiao, X., and Ge, Q. (2023). A Robust and Unified Land Surface Phenology Algorithm for Diverse Biomes and Growth Cycles in China by Using Harmonized Landsat and Sentinel-2 Imagery. ISPRS Journal of Photogrammetry and Remote Sensing 202:610–636. doi: 10.1016/j.isprsjprs.2023.07.017.

Zuo, L., Liu, R., Liu, Y., and Shang, R. (2019). Effect of Mathematical Expression of Vegetation Indices on the Estimation of Phenology Trends from Satellite Data. Chinese Geographical Science 29(5):756–767. doi: 10.1007/s11769-019-1070-y.