GEOGRAPHICAL ANALYSIS

RESEARCH ARTICLE

Most Suitable SNIC Parameters and Classifier Algorithm for Object-Based Classification of Rice Crop Area using Remote Sensing Images in Google Earth Engine

W Khamnoi[1], S Homhuan[1]*, C Suwanprasit[1], Shahnawaz[2]

1 *Department of Geography, Faculty of Social Sciences, Chiang Mai University, 50200, Thailand*
2 *Department of Geoinformatics–Z_GIS, University of Salzburg, 5020, Salzburg, Austria*

OPEN ACCESS

Received: 03.09.2024
Accepted: 24.11.2024
Published: 22.12.2024

Citation: Khamnoi W, Homhuan S, Suwanprasit C, Shahnawaz . (2024). Most Suitable SNIC Parameters and Classifier Algorithm for Object-Based Classification of Rice Crop Area using Remote Sensing Images in Google Earth Engine. Geographical Analysis. 13(2): 71-78. https://doi.org/10.53989/bu.ga.v13i2.57

* **Corresponding author**.
sakda.homhuan@cmu.ac.th

Funding: None

Competing Interests: None

Published By Bangalore University, Bengaluru, Karnataka

ISSN
Print: 2319-5371
Electronic: XXXX-XXXX

Abstract

This study aims at identifying rice crop area in Chae Chang and Buak Kang subdistricts, of San Kamphaeng District, Chiang Mai Province, Thailand through Object Based Image Analysis approach. It investigates the best fit of three elements: (1) most suitable band combination of Sentinel-1 and Sentinel-2 satellite images; (2) optimum parameters for image segmentation; and (3) best performing classifier algorithm in Google Earth Engine [GEE]. The six bands of Sentinel-1 [VV & VH from ascending and descending orbits] and Sentinel-2 [B2, B3, B4 & B8] were used in different combinations. Simple Non-Iterative Clustering [SNIC] method was applied for image segmentation. Three algorithms, Support Vector Machine [SVM], Gradient Boosting Trees [GBT] and Random Forest [RF] were tested for classification and validation. The study analyzed the outcomes of five different band combinations, thirty sets of SNIC parameters - including Compactness [Co], Connectivity [Cn], Neighborhood Size [Ns], Segment Size [Ss] - and three classifier algorithms in GEE. The highest overall accuracy of 97% and a Kappa coefficient of 0.94 was achieved by using all the six bands of the images of the two satellites [Ascending orbit of Sentinel-1] with the SNIC parameter set including Co=0.1, Cn=8, Ns=10, and Ss=5 in RF classifier algorithm. The study reveals that higher Co levels lead to more circle like segments thus unsuitable for rectangular agricultural fields. The results validated against Regions of Interest [ROI] indicate that the optimized SNIC parameters effectively delineated the big and small rice fields covering 34.64 km^2 [73%] of the total 47.23 km^2 area of the two sub-districts. The study offers a pathway for improving classification accuracy in various contexts, particularly in rice area classification.

Keywords: Object-Based Image Analysis; Simple Non-Iterative Clustering; Sentinel-1 and Sentinel-2 images; Support Vector Machine Classifier; Gradient Boosting Trees Classifier; Random Forest Classifier

1 Introduction

Physical changes on the earth's surface due to natural processes and human interventions are continuous phenomena. The human induced changes in agricultural land use appear more frequently due to seasonal cropping patterns. Efficient technologies and advanced methodologies are required for recording and analyzing the extent and pace of such changes for monitoring and management of economic and environmental aspects. As the Remote Sensing [RS] platforms providing images of the Earth's surface at various spatial and temporal resolutions have evolved over time, so have developed the image processing and analysis approaches[1]. One of the prominent image classification methods, Geographic Object-Based Image Analysis [GEOBIA], segments the target RS image by identifying the pixels of similar characteristics and then groups them into objects corresponding to various land use land cover [LULC] classes[2]. This approach offers more accurate representation of landscape structures and patterns by reducing the noise associated with single-pixel based analysis[3]. GEOBIA is particularly effective in the processing of high-resolution images as it can capture fine details and patterns more precisely[4]. Overcoming the effects of shadows, it utilizes various object attributes like color, shape, size, and texture for efficient differentiation and classification of complex objects[5]. Moreover, it can be integrated with the technologies like Machine Learning [ML] and Artificial Intelligence [AI] to enhance analytical algorithms. As mentioned, GEOBIA involves two main steps, i.e. image segmentation and object classification. A variety of algorithms are available for segmentation and delineation of objects, and 'Simple Non-Iterative Clustering [SNIC]' is used widely. The common ML classification algorithms used include Support Vector Machines [SVM], Gradient Boosting Trees [GBT] and Random Forests [RF]. The widespread use of these classifiers in remote sensing applications is recognized for their robust performance[6,7]. Thus, Comparing the relative performance of these classifiers is desirable for selecting the most suitable model considering the crucial factors like accuracy, complexity, resource usage, and practical application flexibility. Proper implementation of ML classification algorithms improves accuracy, increases operational flexibility and allows efficient processing of large datasets. Additionally, these minimize the chances of human errors and enhance the reliability of the results[8].

Google Earth Engine [GEE], based on Google Cloud Platform, is an invaluable online facility for storage, retrieval, processing and applications of remote sensing datasets of various spatial and temporal scales. The platform's ability to quickly access high-resolution images allows the users to efficiently process large size datasets. Additionally, GEE offers numerous pre-built tools and algorithms for data processing, analysis and display, making it ideal for studying and monitoring contemporary environmental changes through online platforms[9]. The integration of Google Cloud ML technology with RS, particularly utilizing Sentinel-1 datasets, involves extracting key features from Synthetic Aperture Radar [SAR] data. This provides a highly efficient tool and technique, especially effective for classifying rice fields. The SAR imaging capably captures data during day and night times as well as in all weather conditions i.e. cloudy, rainy, snowy etc. Also, the SAR images provide VV [Vertical Transmit Vertical Receive] and VH [Vertical Transmit Horizontal Receive] polarizations, and the difference values of these two channels enhance the probability of identifying the physical characteristics of rice crop area more accurately[10,11].

Classifying post-harvest rice fields using satellite imagery poses significant challenges due to varying field conditions and the diverse nature of satellite data. Identifying the optimal combination of classification methods and parameters for accurate detection remains a complex task. Therefore, this study aims to explore and compare various classification methods for identifying post-harvest rice fields using Sentinel-1 satellite imagery within the GEE platform.

2 Study Area

Chae Chang and Buak Kang subdistricts, located in the San Kamphaeng District of Chiang Mai Province, Thailand, cover an area of approximately 47.225 km^2 [Figure 1].

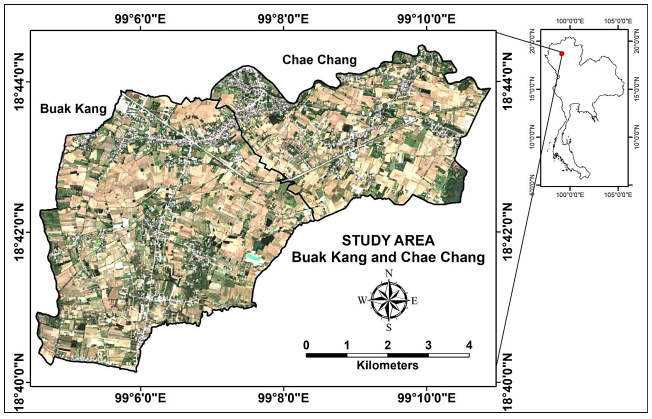

Fig. 1. Satellite View of the Study Area

The study area located between 99°04'32"E, 99°10'54"E longitudes and 18°40'07"N, 18°44'30"N latitudes includes 23 villages. Its physiography is a mixture of flat plains and low hills, and the cultural landscape is predominantly composed of agricultural fields with intermittent settlements and roads. It has tropical monsoon climate with an average annual rainfall of about 1,100 millimeters, the large part of which is received during the rainy season from May to October. The average monthly temperature is around 31°Celsius with a mean monthly minimum about 18°C in January and mean monthly maximum nearly 38°C in April. The area

also benefits from two well-developed irrigation systems, namely Mae-On and Mae Kuang, ensuring the water supply essential for winter-season rice cultivation. The advantageous topography, favorable climatic conditions and availability of irrigation facilities make these subdistricts well-suited for large-scale rice cultivation supporting the livelihoods of the residents. The population of these subdistricts is about 11,000 persons, mostly engaged in rice cultivation which is their staple food and primary source of income[12].

3 Data and Methods

3.1 Data Acquisition and Pre-Processing

This study focuses on identifying the area under rice crop which requires complete LULC classification based on satellite images of the study area and then separate the 'rice area'. Special emphasis is on the 'Harvest Stage' as this marks the period of peak growth and ready to harvest rice crop across the study area which runs from September to December depending on the time of plantation [Table 1].

Table 1. Stages of Rice Cropping

Planting Stage	Growth Stage	Harvest Stage
Seeding / Planting	Water management and fertilization	Monitoring maturity and harvesting
May – June	July – August	September – December

This study is based on the images captured by Sentinel-1 and Sentinel-2 satellites. For Sentinel-1, dual-polarization mode of VV and VH Interferometric Wide Swath [IW] Ground Range Detected [GRD] products have been used. The two Sentinel-1 images of ascending [24-09-2023][9] and descending[9,13] orbits have been filtered according to the boundary of the study area [Table 2]. Additionally, the three visible and one near-infrared band of Sentinel-2 satellite image[11,14] having up to 5% cloud cover was incorporated for comprehensive view of the Earth's surface leveraging the unique strengths of the two datasets. Both the satellites provide images of 10-meter spatial resolution facilitating detailed mapping and precise monitoring of LULC.

All the image datasets were processed through the Sentinel-1 toolbox within GEE and optimized for LULC classification. The Sentinel-1 toolbox includes several operations i.e. importing Sentinel-1 data, applying point filtering, and performing various analyses such as noise removal, calibration, and terrain correction.

The initial information about LULC of the study area was collected from the relevant government agencies[15]. Additionally, the research team made several visits to the study area from May to November 2023 and collected required data about each stage of rice cropping. Along with other details, they also recorded 200 segment locations representing various LULC types to be used as the Regions of Interest [ROI] in the object bases image classification process. The segments of ROI consisting of 60% rice fields, 20% built-up areas, 10% tree cover, 4% water bodies, and 6% other land types were proportionally divided between the two sub-districts. The set of 140 [70%] of the total 200 samples were used for training and remaining 60 [30%] for validation [Table 3].

Table 2. Integrated Image Datasets for LULC Classification

Date of Capture	Source	Band Combination	Spatial Resolution
21-09-2023 and 24-09-2023	Sentinel-1	VV and VH	10 meters
20-11-2023	Sentinel-2	B2, B3, B4, and B8	10 meters
	Multibands-1	VV, B2, B3, B4, and B8	10 meters
	Multibands-2	VH, B2, B3, B4, and B8	10 meters
	Multibands-3	VV, VH, B2, B3, B4, and B8	10 meters

VV = Single co-polarization, vertical transmit vertical receive. VH = Dual-band cross-polarization, vertical transmit horizontal receive. B2 = Blue [B]; B3 = Green [G]; B4 = Red [R]; B8 = Near Infra-Red [NIR]. **NOTE:** These datasets contain the products of the Sentinel-1 and Sentinel-2 Satellite, provided by the European Space Agency [ESA] and accessed through Google Earth Engine. Further details can be accessed at:https://sentiwiki.copernicus.eu/web/sentiwiki

Table 3. The Region of Interest [ROI] Segment Samples

LULC Class	Training [70%]	Validation [30%]	TOTAL
Rice area	86	37	123
Built-up	29	12	41
Tree cover	13	6	19
Water bodies	4	2	6
Other	8	3	11
TOTAL	**140**	**60**	**200**

3.2 Methods

The study aims at classifying rice and non-rice areas integrating remote sensing data with field observations. Image processing and analysis were carried out using the GEE cloud computing platform. The conceptual flow work methodology of the study is illustrated in Figure 2.

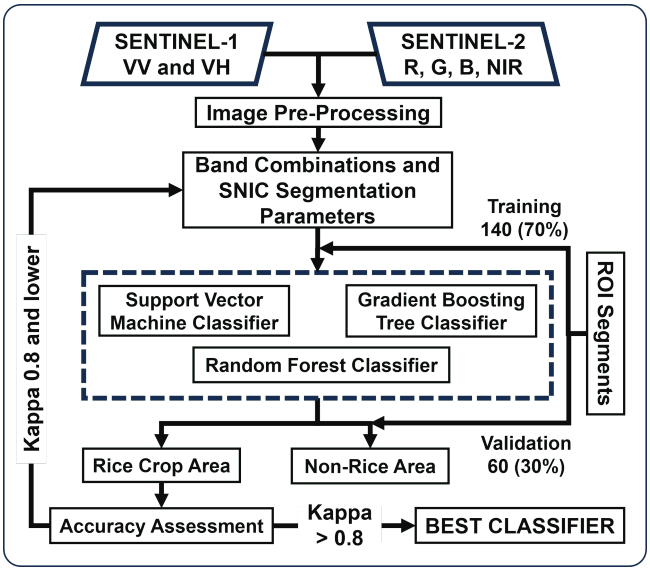

Fig. 2. Conceptual Framework and Methodology of the Study

3.2.1 Simple Non-Iterative Clustering [SNIC]

The SNIC algorithm in GEE is an image segmentation method designed to group neighboring pixels into superpixels based on their spectral and spatial proximity. This technique is particularly beneficial for reducing data dimensionality and noise, thereby enhancing the efficiency and accuracy of subsequent analyses such as image classification or change detection[16]. Unlike other segmentation algorithms, SNIC is non-iterative and focuses on creating superpixels that are uniform in size and shape. The equation of this method is following:

$$S_{[p,c]} = \alpha \times d_s[p,c] + [1-\alpha] \times d_x[p,c] \quad (1)$$

Where: $S_{[p,c]}$ measures the similarity between pixels, p is to a potential superpixel center, c is often based on a combination of d_s and d_x, α is a parameter that balances the weight between spectral and spatial contributions to the total similarity measure, $d_s[p,c]$ represents the Euclidean distance between the spectral values of pixel p and superpixel center c, $d_x[p,c]$ is the spatial distance between the location of pixel p and the centroid of superpixel c

The segments are defined through the-

'ee.Algorithms.Image.Segmentation.SNIC()' function of SNIC algorithm in GEE. The algorithm identifies and groups the pixels of similar surface texture patterns to form segments. These segments are then correlated with 200 sample ROIs [Table 3] to represent LULC types for classification and validation. The SNIC function includes several parameters: Compactness [**Co**], Connectivity [**Cn**], Neighborhood Size [**Ns**], and Segment Size [**Ss**], each influencing the segmentation process differently. **Co** affects the shapes of the clusters, with higher values leading to more compact and circular-shaped clusters, aiding in object distinction. **Cn** determines how adjacent objects are linked to form connected superpixels. **Ns** is important for avoiding the tile boundary artifacts and defining the size of neighboring objects. **Ss** sets the intervals at which superpixel seeds are placed within the area. In this study, various settings of these four parameters [Table 4] were applied for identifying the most suitable combination for best segmentation leading to high accuracy of classification.

Table 4. Parameters used in SNIC for Image Segmentation

Compactness [Co]					Connectivity [Cn]		Neighborhood Segment size [Ns]	Segment size [Ss]		
0.1	0.3	0.5	0.7	1	4	8	10	5	10	20

NOTE: There are total 11 parameters out of which **Co** has 5, **Cn** has 2, **Ns** has only 1, and **Ss** has 3 parameters. Changing one value of one parameter at a time provided 30 sets of SNIC parameters for segmentation. The most suitable set of parameters was assessed through the values of Kappa Coefficient [**K**].

3.2.2 Support Vector Machine [SVM] Classifier

It is a supervised machine learning algorithm in GEE. Its ability to distinctly separate different types of data makes it particularly effective for classification tasks, especially when using the high-resolution satellite imagery[17]. In this study, the SVM algorithm determined the optimal hyperplane that maximized the margin between various classes. The scientific equation of the hyperplane is represented as following:

$$w \times x + b = 0 \quad (2)$$

The decision function:

$$f[x] = sign[w \times x + b] \quad (3)$$

Where: w is the normal vector to the hyperplane, x is the input feature vector, and b is the bias term. This equation determines the class of a given input vector x, either -1 or 1, representing two different land cover types.

3.2.3 Gradient Boosting Trees [GBT] Classifier

GBT in GEE incrementally constructs models by optimizing a loss function. This technique is especially effective in handling complex datasets and enhancing accuracy through its layered modeling approach. GBT efficiently manages imbalanced datasets and provides adjustable parameters to tackle various classifications by iteratively learning from previous models and GBT reduces errors and improves overall outcomes[18,19]. Themodel's equation is expressed as follows:

$$F[x] = \Sigma_{i=1}^{N} \gamma_i h_i[x] \quad (4)$$

Where: $F[x]$ is the final prediction model, γ_i are the weights assigned to each tree, $h_i[x]$ are the predictions from individual weak learner trees.

3.2.4 Random Forest [RF] Classifier

In GEE, the RF constructs multiple decision trees during training process and outputs the statistical mode of the classes for images classification or the statistical mean prediction for regression. Its robust ability to handle multi-dimensional and complex datasets helps efficient classification process. Moreover, the use of multiple trees reduces the risk of overfitting the model and ensures its stability[20,21]. Below is given its equation:

$$f[x] = \frac{1}{N}\Sigma_{i=1}^{N} f_i[x]$$

Where: $f_i[x]$ is the prediction from the i-th tree and N is the number of trees.

3.2.5 Accuracy Assessment and Validation

The percentage of Overall Accuracy [OA] and the values Kappa Coefficient [K] have been used for assessing and validating the accuracy of the classification. Although the values of K may range between -1 and +1 but the ones more than 0.8 represent greater agreement so only such values have been considered for identifying the best fit of SNIC parameters with the three classifiers. The equation for calculating overall accuracy is following:

$$Overall\ Accuracy\ [OA] = \frac{\Sigma_{i=1}^{N} C_{ii}}{\Sigma_{i=1}^{N} T_i} \quad (6)$$

Where: N is total number of classes, C_{ii} is number of correctly classified samples for class i, T_i is total number of samples for class i

The kappa coefficient equation is:

$$Kappa\ Coefficient\ [K] = \frac{P_o - P_e}{1 - P_e}$$

Where: P_o is observed agreement $[\frac{\Sigma_{i=1}^{N} C_{ii}}{T_i}]$ and P_e is expected agreement $[\frac{\Sigma_{i=1}^{N}[T_i \times C_i]}{T^2}]$ by T is total number of samples, T_i is total number of samples for class i, C_i is total number of samples classified into class i

4 Results and Discussion

4.1 SNIC Parameters

The analysis of 30 sets of the SNIC parameters reveals that variations in the parameter values cause variations in segmentation of the images. Changes the level of **Co** [superpixel density] result in slight alterations in the location, shape, and size of the segments. Higher levels of **Co** tend to produce more circle like segments, which are not suitable for fitting rectangular shaped rice crop fields. In case of **Cn** values, the properties of the adjacent pixels can be interpreted by using **Cn=4** [horizontal and vertical directions] and **Cn=8** [horizontal, vertical and 2 diagonal directions].

Table 5. SNIC Parameters, OA and K Values of Three Classifiers

SNIC Parameters				SVM		GBT		RF	
Co	**Ct**	**Ns**	**Ss**	**OA%**	**K**	**OA%**	**K**	**OA%**	**K**
0.1	4	10	5	53	0.09	94	0.90	95	0.91
0.3	4	10	5	95	0.91	94	0.90	95	0.91
0.5	4	10	5	53	0.09	94	0.90	95	0.91
0.7	4	10	5	53	0.09	94	0.90	95	0.91
1	4	10	5	53	0.09	94	0.90	95	0.91
0.1	**8**	**10**	**5**	53	0.11	94	0.89	**97**	**0.94**
0.3	8	10	5	53	0.11	94	0.89	96	0.92
0.5	8	10	5	53	0.09	94	0.89	96	0.92
0.7	8	10	5	53	0.09	94	0.89	96	0.92
1	8	10	5	53	0.09	94	0.89	96	0.92
0.1	4	10	10	51	0.17	86	0.74	87	0.76
0.3	4	10	10	51	0.17	86	0.74	87	0.76
0.5	4	10	10	51	0.17	86	0.74	87	0.76
0.7	4	10	10	51	0.17	86	0.74	87	0.76
1	4	10	10	51	0.17	86	0.74	87	0.76
0.1	8	10	10	63	0.23	91	0.83	91	0.84
0.3	8	10	10	63	0.23	91	0.83	91	0.84
0.5	8	10	10	63	0.23	91	0.83	91	0.84
0.7	8	10	10	63	0.23	91	0.83	91	0.84
1	8	10	10	63	0.23	91	0.83	91	0.84
0.1	4	10	20	44	0.13	83	0.67	86	0.73
0.3	4	10	20	44	0.13	83	0.68	86	0.73
0.5	4	10	20	44	0.13	83	0.67	86	0.73
0.7	4	10	20	44	0.13	83	0.67	86	0.73
1	4	10	20	44	0.13	83	0.68	86	0.73
0.1	8	10	20	53	0.08	81	0.66	88	0.77
0.3	8	10	20	53	0.08	81	0.66	88	0.77
0.5	8	10	20	53	0.08	81	0.66	88	0.77
0.7	8	10	20	53	0.08	81	0.66	88	0.77
1	8	10	20	53	0.08	81	0.66	88	0.77

Co = Compactness; Cn = Connectivity; Ns = Neighborhood Size; Ss = Segment Size; SVM = Support Vector Machine: GBT = Gradient Boosting Trees; RF = Random Forest; OA = Overall Accuracy; K = Kappa Coefficient

Using **Cn=8** produces neater segments and reduces segment fragmentation in the areas with minimal differences. The **Ns** [necessary for avoiding artifacts at the boundaries] refers to the radius of the region used in segment computation. The larger **Ns** allows the inclusion of more similar segments within the same area, which is appropriate for classifying large and less segmented areas. Further, **Ss** is crucial for parameter adjustment and is necessary for the accurate classification of rice fields of various sizes. Setting the **Ss** smaller helps achieving more detailed and better separation of areas, though it may result in an excessive number of segments. In this study, the optimal set of SNIC parameters leading to

highest accuracy of classification across the three classifiers are presented in Table 5. The table shows that the most optimal set of SNIC parameters was found as: **Co=0.1; Cn=8; Ns=10 and Ss=5** leading to achieve **OA of 97%** and **K of 0.94** in LULC classification facilitating reliable identification of rice crop area. Finding an optimal set of parameter is crucial for accurate LULC classification [13,14,22].

4.2 Classifier Algorithms

When comparing the overall performance of the three classifier algorithms, SVM returned lowest values of **OA** and **K**, while GBT provided relatively better results and RF emerged to the best [Table 6] [23,24]. Also, it was found that using only VV and VH polarization data of Sentinel-1 resulted in poor classification in all the three classifiers. However, combining B2, B3, B4, and B8 bands of Sentinel-2 with VV and VH polarization data from Sentinel-1 significantly improved the accuracy of classification. The highest values of **OA=97%** and **K=0.94** were provided by the RF classifier from Multibands-3 combination including Ascending Orbit datasets of Sentinel-1 satellite, illustrating the benefits of multiple data fusion in remote sensing applications [25].

4.3 LULC Classification and Rice Crop Area Identification

In GEOBIA approach, the classifier algorithms require a ROI for grouping various segments into corresponding LULC classes. The 200 segment locations representing six LULC classes collected during the fieldwork [Table 3] were divided into two parts i.e. 140 [70%] for training the classifiers and 60 [30%] for validating the accuracy of classification. As mentioned above, RF algorithm provided the best results [OA=97% and K=0.94] based on appropriate combination of bands and optimum set of SNIC parameters. Out of total 47.23 km^2 area of the two sub-districts, 34.64 km^2 [73.32%] is covered under rice crop fields [Table 7 & Figure 3]. The rice crop area falling in Buak Khang is about 22.53 km^2 and Chae Chang is nearly 12.10 km^2. The remaining 12.6 km^2 [26.68%] of the total study area turns out as non-rice area which includes various types of built-up areas, tree cover, water bodies, and other areas used for miscellaneous purposes.

Although various types of non-rice areas are small-sized and occur intermittently in the study area [Figure 3], but the methodology used for LULC classification demonstrates its ability to distinguish complex areas quite effectively. Despite the limitations in segmenting the images of complex regions, it is suitable for generating continuous and uniform stretches of rice crop areas and large-scale agricultural and environmental monitoring [26–29].

Table 6. Overall Accuracy and Kappa Coefficient Provided by the Three Classifiers

Results of Optimum Set of SNIC Parameters: Co = 0.1, Cn = 8, Ns = 10, Ss = 5							
Descending		**OA [%]**	**Kappa**	**Ascending**		**OA [%]**	**Kappa**
SVM	VV	30	0.00	**SVM**	VV	27	-0.03
	VH	36	-0.01		VH	27	-0.01
	Multibands-1	50	0.03		Multibands-1	27	-0.01
	Multibands-2	50	0.06		Multibands-2	53	0.11
	Multibands-3	45	-0.01		Multibands-3	49	0.07
GBT	VV	67	0.38	**GBT**	VV	63	0.29
	VH	73	0.47		VH	69	0.40
	Multibands-1	92	0.85		Multibands-1	92	0.87
	Multibands-2	93	0.88		Multibands-2	91	0.84
	Multibands-3	92	0.86		Multibands-3	92	0.86
RF	VV	70	0.44	**RF**	VV	66	0.30
	VH	69	0.41		VH	67	0.36
	Multibands-1	90	0.81		Multibands-1	95	0.90
	Multibands-2	94	0.90		Multibands-2	94	0.89
	Multibands-3	91	0.83		**Multibands-3**	**97**	**0.94**

Table 7. Area of Various LULC Classes

Rice Crop [km^2]	Built-Up [km^2]	Tree Cover [km^2]	Water Bodies [km^2]	Other [km^2]	TOTAL [km^2]
34.63 [73.32%]	6.06 [12.83%]	5.49 [11.62]	0.65 [1.31%]	0.40 [0.84%]	47.23 [100%]

5 Conclusions

This study attempts to identify the most suitable machine learning algorithm for delineating the area under rice crop adopting GEOBIA approach. It investigates the outcomes of several crucially relevant factors which include 5 combinations of 6 multi-sensor satellite image bands, 5 sets of 4 image segmentation parameters and 3 classifier algorithms using GEE for processing and analysis. The SNIC operation was run for segmenting the selected multi-sensor image bands and LULC classification was performed through SVM, GBT and RF algorithms.

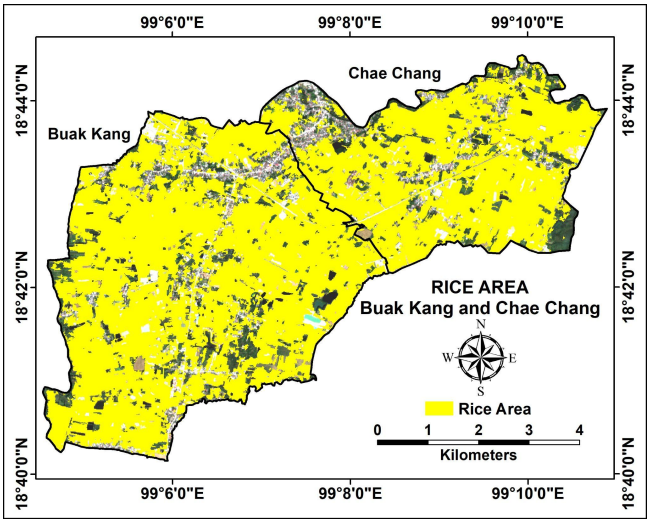

Fig. 3. The Best Rice Area Classification Result

The results indicate that minimal image segmentation provides better alignment of the segment boundaries with the boundaries of the corresponding sample areas collected in fieldwork. Although, an optimal set of 4 SNIC parameters was identified but these may or may not be suitable for other areas because shape and size of agricultural fields require adaptation of these parameters.

While comparing the 3 ML classifier algorithms, RF outperformed SVM and GBT in terms of OA [97%] and K [0.94]. SVM was less effective due to its computational intensity and reduced accuracy with distinct class boundaries, but GBT performed better particularly in case of imbalanced spatial pattern. Integration of the 3 optical bands of Sentinel-2 with the VV and VH polarization data of Sentinel-1 enhanced the classification performance significantly, highlighting the benefits of multi-sensor data fusion.

The capability of segments in capturing the objects accurately is fundamental in determining the accuracy of classification. Each classifier's capabilities vary depending on the segmentation parameters. The results show that using different classifiers with the same parameters significantly impacts classification outcomes, with RF performing best compared to SVM and GBT. Therefore, it can be concluded that classification accuracy and precision should derive from both good segmentation and good classifier.

Acknowledgments

This work was supported by the ERASMUS+ KA107: Student and Staff Mobility program. We extend our gratitude to the European Commission and the University of Salzburg for facilitating this mobility, and to the Chiang Mai University for continued support throughout this project.

References

1) Ez-Zahouani B, Kharki OE, Idé SK, Zouiten M. Determination of Segmentation Parameters for Object-Based Remote Sensing Image Analysis from Conventional to Recent Approaches: A Review. *International Journal of Geoinformatics*. 2023;19(1):23–42. Available from: https://doi.org/10.52939/ijg.v19i1.2497.
2) Kucharczyk M, Hay GJ, Ghaffarian S, Hugenholtz CH. Geographic Object-Based Image Analysis: A Primer and Future Directions. *Remote Sensing*. 2020;12(12):1–33. Available from: https://doi.org/10.3390/rs12122012.
3) Blaschke T, Hay GJ, Kelly M, Lang S, Hofmann P, Addink E. Geographic Object-Based Image Analysis - Towards a New Paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing*. 2014;87:180–191. Available from: https://doi.org/10.1016/j.isprsjprs.2013.09.014.
4) Lang S, Hay GJ, Baraldi A, Tiede D, Blaschke T. GEOBIA Achievements and Spatial Opportunities in the Era of Big Earth Observation Data. *ISPRS International Journal of Geo-Information*. 2019;8(11):1–19. Available from: https://doi.org/10.3390/ijgi8110474.
5) Qian Y, Zhou W, Yan J, Li W, Han L. Comparing Machine Learning Classifiers for Object-Based Land Cover Classification Using Very High Resolution Imagery. *Remote Sensing*. 2015;7(1):153–168. Available from: https://doi.org/10.3390/rs70100153.
6) Cheng F, Ou G, Wang M, Liu C. Remote Sensing Estimation of Forest Carbon Stock Based on. *Machine Learning Algorithms Forests*. 2024;15(4):1–25. Available from: https://doi.org/10.3390/f15040681.
7) Wu N, Crusiol L, Liu G, Wuyun D, Han G. Comparing Machine Learning Algorithms for Pixel/Object-Based Classifications of Semi-Arid Grassland in Northern China Using Multisource Medium Resolution Imageries. *Remote Sensing*. 2023;15(3):1–22. Available from: https://doi.org/10.3390/rs15030750.
8) Zhou R, Yang C, Li E, Cai X, Yang J, Xia Y. Object-Based Wetland Vegetation Classification Using Multi-Feature Selection of Unoccupied Aerial Vehicle RGB Imagery. *Remote Sensing*. 2021;13(23):1–21. Available from: https://doi.org/10.3390/rs13234910.
9) Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*. 2017;202:18–27. Available from: https://doi.org/10.1016/j.rse.2017.06.031.
10) Khamnoi W, Homhuan S, Suwanprasit C, Shahnawaz. Assessment of Post-Harvest Rice Crop Biomass and Carbon Stock Using Remote Sensing Data in Google Earth Engine. *International Journal of Geoinformatics*. 2024;20(8):88–101. Available from: https://doi.org/10.52939/ijg.v20i8.3459.
11) Luo C, Qi B, Liu H, Guo D, Lu L, Fu Q, et al. Using Time Series Sentinel-1 Images for Object-Oriented Crop Classification in Google Earth Engine. *Remote Sensing*. 2021;13(4):1–19. Available from: https://doi.org/10.3390/rs13040561.
12) National Statistical Office. Provincial Statistical Development Plan. Ministry of Digital Economy and Society, Thailand. . Available from: https://www.nso.go.th/nsoweb/official_stat/province/F?set_lang=th.
13) Akcay O, Avsar EO, Inalpulat M, Genc L, Cam A. Assessment of Segmentation Parameters for Object-Based Land Cover Classification Using Color-Infrared Imagery. *ISPRS International Journal of Geo-Information*. 2018;7(11):1–26. Available from: https://doi.org/10.3390/ijgi7110424.
14) El-Naggar AM. Determination of Optimum Segmentation Parameter Values for Extracting Building from Remote Sensing Images. *Alexandria Engineering Journal*. 2018;57(4):3089–3097. Available from: https://doi.org/10.1016/j.aej.2018.10.001.
15) Land Development Department. Dinonline. Ministry of Agriculture and Cooperatives, Thailand. . Available from: https://dinonline.ldd.go.th.
16) Achanta R, Süsstrunk S. Superpixels and Polygons Using Simple Non-Iterative Clustering. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 2017;p. 4651–4660. Available from: https://doi.org/10.1109/CVPR.2017.520.
17) Jamali A. Sentinel-1 Image Classification Using Machine Learning Algorithms Based on the Support Vector Machine and Random Forest. *International Journal of Geoinformatics*. 2020;16(2):15–22. Available from: https://journals.sfu.ca/ijg/index.php/journal/article/view/1809.
18) Jamali A, Mahdianpari M, Karaş İR. A Comparison of Tree-Based Algorithms for Complex Wetland Classification Using the Google Earth Engine. The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*;XLVI-4/W5-2021:313–319. Available from: https://doi.org/10.5194/isprs-archives-XLVI-4-W5-2021-313-2021,2021.
19) Sharma V, Ghosh SK. Evaluating the Potential of 8 Band Planet Scope Dataset for Crop Classification Using Random Forest and Gradient Tree Boosting by Google Earth Engine . *The International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences*;XLVIII-M-1-2023:325–330. Available from: https://doi.org/10.5194/isprs-archives-XLVIII-M-1-2023-325-2023.
20) Melati DN, Astisiasari T. An Assessment of Object-based Classification Compared to Pixel-based Classification in Google Earth Engine Using Random Forest. In: 2022 IEEE Asia-Pacific Conference on Geoscience, Electronics and Remote Sensing Technology (AGERS). IEEE. 2022;p. 73–78. Available from: https://doi.org/10.1109/AGERS56232.2022.10093267.
21) Tassi A, Gigante D, Modica G, Martino LD, Vizzari M. Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sensing*. 2021;13(12):1–20. Available from: https://doi.org/10.3390/rs13122299.
22) Selvaraj R, Amali DGB. Assessment of Object-Based Classification for Mapping Land Use and Land Cover Using Google Earth. *Global NEST Journal*. 2023;25(7):131–138. Available from: https://doi.org/10.30955/gnj.004829.
23) Vizzari M, Lesti G, Acharki S. Crop classification in Google Earth Engine: leveraging Sentinel-1, Sentinel-2, European CAP data, and object-based machine-learning approaches. *Geo-spatial Information Science*. 2024;p. 1–16. Available from: https://dx.doi.org/10.1080/10095020.2024.2341748.
24) Tassi A, Vizzari M. Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms. *Remote Sensing*. 2020;12(22):1–17. Available from: https://dx.doi.org/10.3390/rs12223776.
25) Karakuş P. Object Based Classification in Google Earth Engine Combining SNIC and Machine Learning Methods (Case Study: Lake Köyceğiz). *Turkish Journal of Remote Sensing and GIS*. 2024;5(1):125–137. Available from: https://doi.org/10.48123/rsgis.1411380.
26) Clinton N, Holt A, Scarborough J, Yan L, Gong P. Accuracy Assessment Measures for Object-based Image Segmentation Goodness. *Photogrammetric Engineering & Remote Sensing*. 2010;76:289–299. Available from: https://dx.doi.org/10.14358/pers.76.3.289.
27) Clinton N, Holt A, Yan A, Gong P. An Accuracy Assessment Measure for Object Based Image Segmentation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2008;XXXVII(Part B4):1189–1194. Available from: https://www.isprs.org/proceedings/XXXVII/congress/4_pdf/208.pdf.
28) Matarira D, Mutanga O, Naidu M, Vizzari M. Object-Based Informal Settlement Mapping in Google Earth Engine Using the Integration of Sentinel-1, Sentinel-2, and PlanetScope Satellite Data. *Land*. 2022;12(1):1–17. Available from: https://dx.doi.org/10.3390/land12010099.
29) Zhao Z, Islam F, Waseem LA, Tariq A, Nawaz M, Islam IU, et al. Comparison of Three Machine Learning Algorithms Using Google Earth Engine for Land Use Land Cover Classification. *Rangeland Ecology & Management*. 2024;92:129–137. Available from: https://dx.doi.org/10.1016/j.rama.2023.10.007.