Applied Nanoscience
https://doi.org/10.1007/s13204-021-01969-3

REVIEW ARTICLE

Mining images of high spatial resolution in agricultural environments

D. V. Lalitha Parameswari[1] · CH. Mallikarjuna Rao[2] · D. Kalyani[3] · G. Ramesh[2]

Received: 8 May 2021 / Accepted: 6 July 2021
© King Abdulaziz City for Science and Technology 2021

Abstract

Satellites are widely used for remote sensing applications. High resolution images are used for different geographical applications. Using geographical objects or spatial objects for analysis became prevalent in the contemporary era. Many supervised classification techniques came into existence to have efficient classification of high-resolution imagery. There are many factors that may affect classification of geographic images. They include the presence of mixed objects, feature selection, size of training set and segmentation scale. When these factors are considered for a systematic mining of images with high resolution, it results in improved performance. Especially in agricultural environments, it is essential to have such study to ascertain which supervised learning mechanism can best deal with the factors aforementioned. An algorithm named Feature Subset Selection (FSS) is defined to enhance classification accuracy. Different classification techniques such as Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes, k-Nearest Neighbour (KNN), Adaboost.M1 and Decision Table (DT) are used for the empirical study with spatial data mining. Useful analysis of the techniques is made and thus this paper provides valuable insights on mining images of high spatial resolution in agricultural environments.

Keywords Spatial data mining · Classification · SVM · RF · KNN · DT · Adaboost.M1

Introduction

There are technological advances in space borne and air-borne remote sensing technologies. With respect to agricultural environments, it is important to have image classification and mapping to have better decision-making. Spatial data mining helps in identifying land cover, land usage and other agriculture-related aspects. There are many supervised learning methods that came into existence. Spatial data are the data related to geographical locations with purposeful coverage. For instance, spatial data are related to agriculture cultivation. Remote Sensing (RS) images related to agriculture can help in understanding crop dynamics and the usage patterns of the land available. Spatial data also may have non-spatial data associated with it. This is shown in as presented in Fig. 1, there is spatial data that are categorized into water, residential, institutional, industrial and commercial. However, the spatial data have associated non-spatial data. The non-spatial data have city blocks and corresponding land use details. In spatial data mining, it is possible to have non-spatial predicates in queries. Thus non-spatial data provide additional details for processing user queries on spatial data.

There were many studies on the image analysis. In Blanche et al. (2014) geographical object analysis is explored with paradigm shift. LIBSVM is the library with variants of SVM studied in Chang and Lin (2011) (Mallikrarjuna Rao et al. 2019; Lalita Parameswari et al. 2014) for spatial data analysis. Image segmentation with automatic approach for parameter setting is investigated in Drăguţ et al. (2014). The image segmentation scale and the impact of this factor in analysing spatial data related to agriculture is focussed in Dronova et al. (2012). Spatial image analysis include object-based and pixel-based analysis as explored in Duro et al. (2012). Different approaches found in the literature have different approaches for classification of spatial imagery. However, there is need for leveraging feature selection to enhance quality in the training process. Similarly, from the review of literature presented in Related work, it is understood that there is need for analysing various factors of image classification so as to improve the utility of

✉ G. Ramesh
ramesh680@gmail.com

1 Department of CSE, GNITS, Hyderabad, India

2 Department of Computer Science and Engineering, GRIET, Bachupally, Hyderabad 500090, India

3 Department of IT, VNR VJIET, Hyderabad, Telangana, India

Published online: 15 July 2021

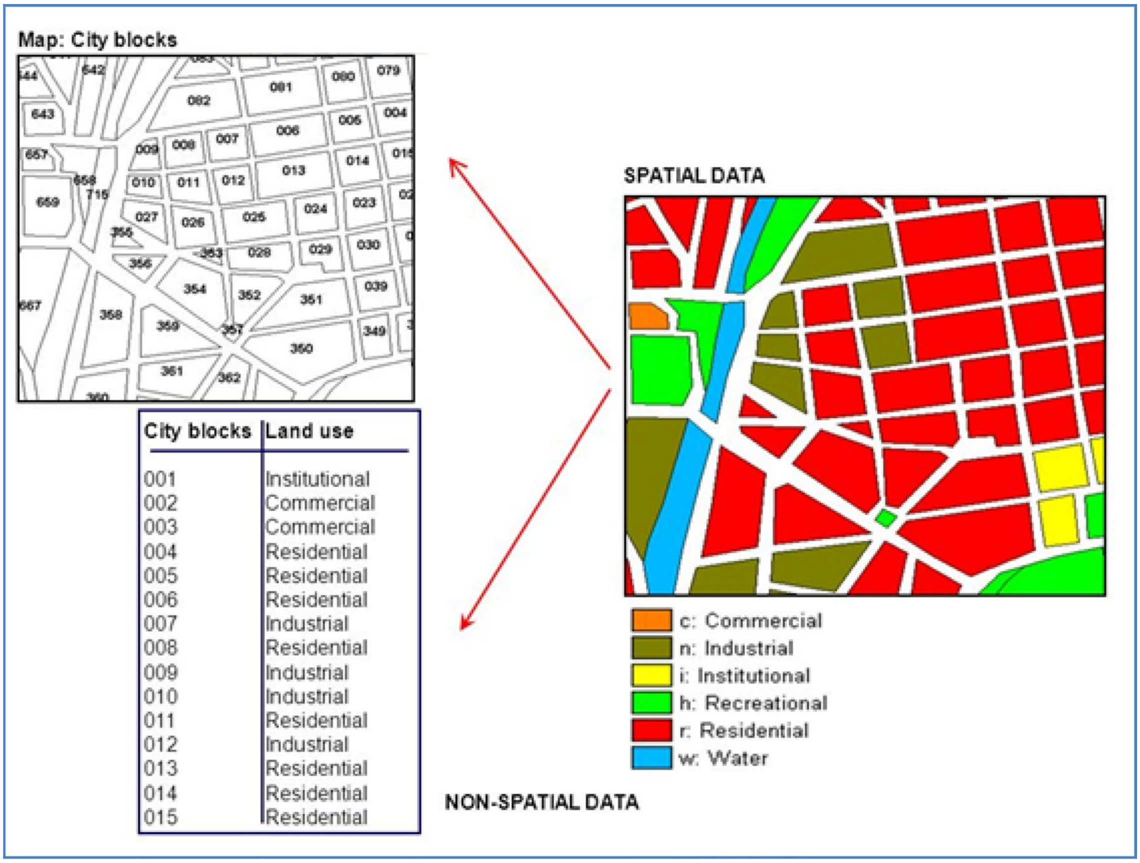

Fig. 1 Spatial data containing information about non-spatial data

classification of spatial imagery. The following are contributions of this paper.

- We proposed a methodology with different classification techniques and parameters for evaluation of spatial data mining approaches on the imagery in agricultural environments.
- The proposed feature subset selection algorithm improves the performance of the classification techniques.
- We built a prototype application to study different classification techniques with and without feature selection.
- We analysed the results of empirical study and provided suitable insights that can help stakeholders to make well informed decisions in choosing mining choices in future.

The remainder of the paper is structured as follows. Related work provides review of literature that analyse the state of the art. Proposed solution provides the proposed methodology. Experimental results presents experimental results. Conclusions and future work gives conclusions and provides directions for future work.

Related work

This section reviews literature on classification techniques that are widely used for mining spatial imagery. Alfaro et al. (2013) explored classification techniques with bagging and boosting. It has provision for Adaboost and Adaboost.M1. Ensembles are trained to have better solution for classification. Blanche et al. (2014) proposed a new paradigm for image analysis. The images considered are geographic objects. Chang and Lin (2011) studied SVM using LIBSVM library that supports classification of imagery with different variants like one-class SVM, Support Vector Regression (SVR) and Support Vector Classification (SVC). Drăguţ et al. (2014) on the other hand proposed a framework for automated parameterisation required by multi-scale parameterisation. High resolution imagery is used for empirical study. Dronova et al. (2012) explored image segmentation with different segmentation scale, classification methods and vegetation classes. Object level texture metrics are used to make classification decisions. Wetland plant function types are considered for classification. However, they are yet to explore functional diversities in different spatial imagery.

Unlike the work in Blanche et al. (2014) where object level image analysis is made, (Duro et al. 2012) focussed on both pixel-based and object-based analysis of high-resolution images. Different classification algorithms like RF, SVM and DT are used for empirical study. In agricultural environments, it is essential to use more than one method for better results. Earth observation imagery are used for inputs. In remote sensing-based estimations, (Fassnacht et al. 2014) focussed on experimentation related to the importance of prediction methods, data types and the size of data in case of above ground forest biomass analysis. They used different analysis methods for the same and found that LiDAR (Light Detection and Ranging) is a better technique. Bamboo patches classification is employed in Ghosh and Joshi (2014) with different classification algorithms. They used agricultural environments to study land use and land cover (LULC). Recursive feature elimination method is used to have highly relevant features for classification purposes.

Object-based classification methods with non-parametric approach are studied by Luque et al. (2013). Regression tree, classification algorithms like NN, SVM and CART are used for the study of classification methods. Three study areas are used for empirical study. The results revealed that SVM classifier was better than other methods. Ma et al. (2014) threw light on extracting cultivated land information from spatial imagery. The imagery was captured from unmanned aerial vehicle. A technical framework is proposed to have cultivated land information. Ma et al. (2015) studied size, scale and features of training set as part of object-based image analysis. UAV imagery is used for empirical study. They found that UAV imagery can be used to have land cover and land use classifications.

Pu and Landry (2012) made a comparative analysis of spatial imagery for mapping urban tree species. Their research includes classification, regression trees, linear discrimination, masking and object-based image analysis. More research on object-based image analysis is made in Radoux and Bogaert (2014). Land cover classification is explored in Shao and Lunetta (2012) with limited training data as part of agricultural research. Land cover mapping is performed along with accuracy assessment. Forest habitat mapping is studied in Strasser and Lang (2015) with multiscale object-based class modelling. There was comparative study between object-based and pixel-based image analysis. Land use and land cover are studied for better understand of geography (Tehrany et al. 2014). Classified image objects are validated by Whiteside et al. (2014) using location-based and area-based approaches. Oil spill identification (Xu et al. 2014), Baysian and Network Classifier (Yang and Wang 2012) and segmentation quality evaluation (Zhang et al. 2015; Ramesh and Mallikrarjuna Rao 2018) for improving region-based precision and recall are other important researches found in the literature. A machine learning-based approach (Atul et al. 2021; Tian et al. 2021) and exploration of remote sensing possibilities (Arifjanov et al. 2021) are other important contributions found in the literature. From the review of literature, it is found that there needs to be a comprehensive approach which can cater to the needs of an application like classification of objects in spatial imagery. In this paper, we proposed methodology to achieve this. Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes, k-Nearest Neighbour (KNN), Adaboost.M1 and Decision Table (DT) are evaluated with different parameters and metrics. The results revealed that there is need for evaluating classifiers with different parameters and feature selection.

Proposed solution

This section provides the proposed methodology for finding performance of different algorithms and evaluate the same with respect to Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes, k-Nearest Neighbour (KNN), Adaboost.M1 and Decision Table (DT). Towards this end, the input remote sensing imagery is divided into two areas. Figure 2 shows the two areas of the dataset used for empirical study.

As explored in Duro et al. (2012), the study areas are considered with respect to agricultural environment to be part of a landscape. The legend associated with each image pair shows the details of buildings, woodland, crop, bare land, road and water. The two areas are used for experiments. Different classification techniques are used for evaluation. The dataset comes from the high-resolution imagery Deyang city of Sichuan province in China. The images were collected from UAV. With each classification technique, feature selection is employed. As the images are of very high resolution, the classified images are used as reference images. Different land cover types are found in the study area 1. They include buildings, woodland, crop, bare land and road. The land cover types associated with the study area 2 include buildings, woodlands, crop, bare land, road and water. Digital Orthogonal Map (DOM) data are generated. Based on the DOM data generated, different segmentation scales are used for experiments. Most suitable algorithms were used for empirical study. Correlation-based feature selection method is employed for pre-processing.

The quality of subset of features was verified. Before indulging into a classification technique, the correlation technique is applied to all features that are computed.

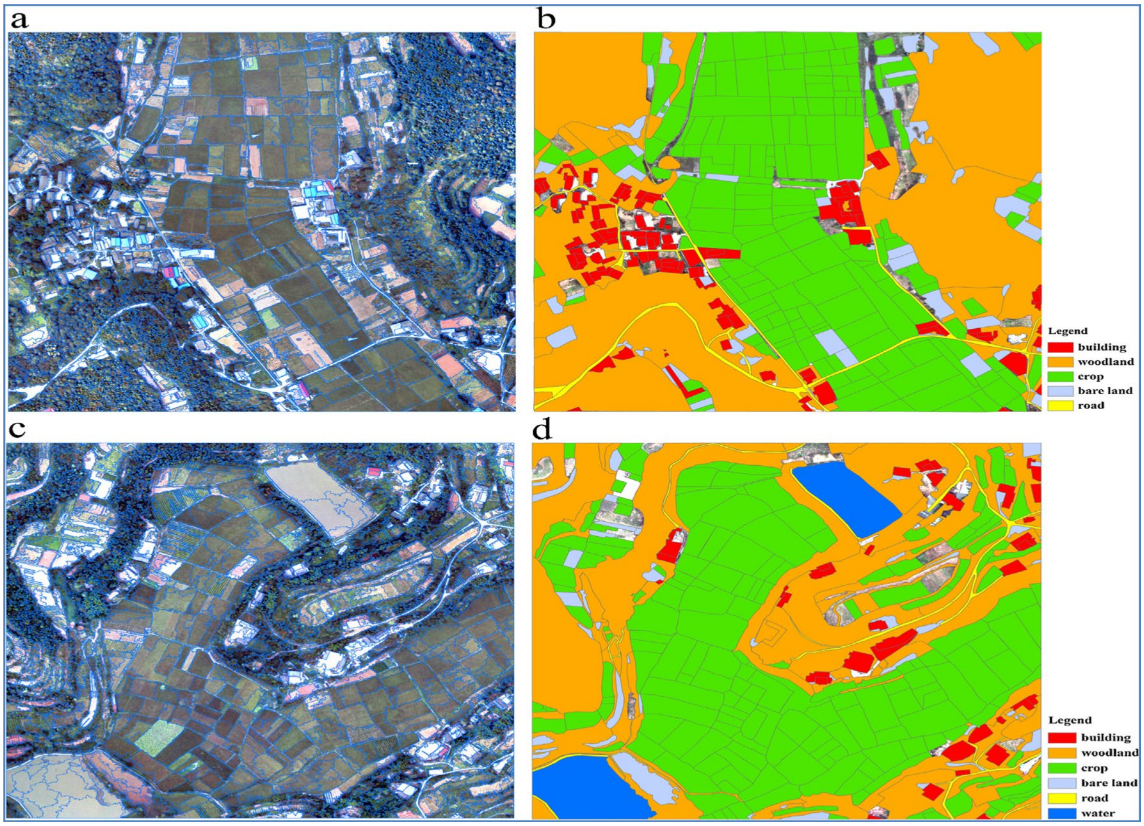

Fig. 2 Segmented image **a** and its corresponding interpretation image **b** for area 1 and segmented image **c** and its corresponding interpretation image **d** for area 2; both are taken at a scale of 100

Different measures are used for computations. They include mean red, mean green, mean blue, brightness, standard deviation red, standard deviation blue and standard deviation green. Texture measures are also used for better results. Different classifiers are used for empirical study as shown in Fig. 3.

Support vector machine: SVM is non-parametric and binary supervised learning classifier. LIBSVM library is used for implementing SVM. It supports four kinds of kernels. In this study, Radial Basis Function (RBF) is used. Penalty and kernel are two parameters to be used for the RBF. SVM is used with cross validation for the purpose of the experiments.

Random forest (RF): Random forest (RF) is another well-known classifier used. It has been used widely for remote sensing images. A random method is used by RF and it needs two parameters namely number of predication variables and classification trees desired. RF has built in functionality to be carried out for predictions.

K-nearest neighbour: KNN is widely used classifier due to its simplicity and also flexibility. It assigns an object to a class depending on the nearest neighbours in the search space. From the training set closes nearest neighbours are considered to predict a new object. Cross-validation and bootstrap values are used for searching the best K value.

Decision tree: Of late, DT is widely used to deal with remote sensing imagery. Binary recursive partitioning is employed to have a tree grown. The split process of the dataset is repeated until the terminal nodes are too small or too few to be split.

Adaboost.M1: To have ensemble of methods, boosting is a popular method. Therefore, the Adahoost.M1 is used for classification of objects in earth observation imagery. Alpha is a learning factor (a constant) used for experiments. The value 100 is set for number of trees and number of iterations.

Naïve Bayes: It is another powerful tool for classification of spatial imagery. Probabilistic representation and reasoning are employed in this algorithm. It also deals with conditions of uncertainty. The problem is then boiled down to conditional probability and it is widely used classifier for remote sensing imagery. Accuracy assessment and visual assessment are employed to know the advantages of the proposed system. The experiments are made with and without feature selection. The difference in results with a single selected

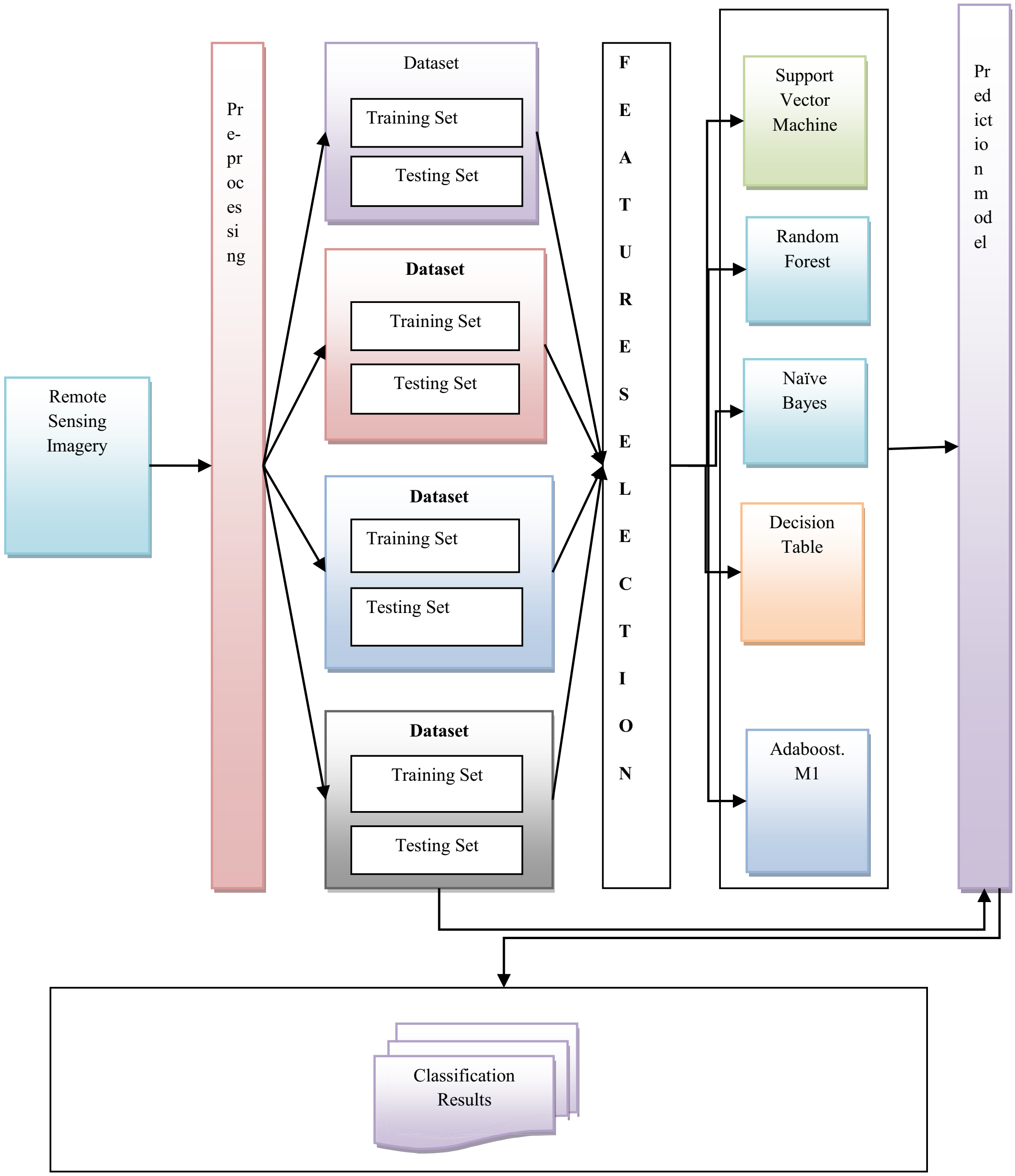

Fig. 3 Shows the methodology of the proposed approach

feature and all features are shown. The effect of training size on the accuracy of classification is provided. Homogenous and heterogeneous objects are there in the given dataset. In presence of such samples experiments are made. Agricultural environment is considered for empirical study. The proposed feature subset selection algorithm is based on two measures namely entropy and gain. Entropy measures impurity while gain reflects the amount of information a feature can help in contributing class label identification.

Algorithm: Feature Subset Selection

Inputs: Dataset ***D*** (spatial imagery)

Outputs: Relevant Features ***RF***

```
Initialize f for storing a single feature
Initialize gain to store gain value
Initialize entropy to store entropy value
Initialize RF to store candidates
Initialize RF’ to store features relevant
For f in D
   Compute entropy using Eq. (1)
   Compute gain using Eq. (2)
    IF gain and entropy are > threshold THEN
        Associate f with RF
    END IF
End For
For each f in the RF
   Obtain pair of features
   IF f has correlation with f’ THEN
        Associate f to RF’
    END IF
End For
Return RF’
```

Algorithm 1: Feature selection algorithm

The feature selection is made based on the proposed algorithm which helps in identification of features that lead to contribution of class label selection. This has potential to improve quality in training and then improve performance of the classification techniques used in the proposed approach. The proposed algorithm uses the measures like entropy and gain to determine the correlation between features. Entropy as in Eq. (1) shows the uncertainty in the dataset while the gain as in Eq. (2) is derived from the entropy. They are computed based on the data associated with features of dataset.

$$H(X) = \sum_{x \in X} p(x) \log_2 p(x) \tag{1}$$

$$\text{Gain}\ (X/Y) = H(X) - H(X/Y) \tag{2}$$

$$= H(Y) - H(Y/X)$$

Gain reveals the possible reduction in entropy and the two measures are related to each other to have a mechanism to determine correlations among features. This is very important consideration for making decisions in choosing subset of features from given dataset that are highly relevant to the given objective.

Experimental results

Experiments are made with high-resolution spatial images of two areas as shown in Fig. 2. In either case, different classifiers are built and the performance of the classifiers is evaluated. The classification algorithms include SVM, RF, Adaboost.M1, Naïve Bayes and DT.

Results on area 1

Area 1 considered in Fig. 2 is used and the results are observed. The results are presented with segmentation scale and overall accuracy as performance metrics to compare among the aforementioned classifiers. Spline line of mean is a measure used to have spatial analysis which is based on selected features or all features.

As presented in Table 1, overall accuracy of SVM classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features.

Table 1 Shows overall accuracy of SVM classifier with spline line of mean with chosen feature and all features

SVM		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.85	0.86
50	0.86	0.87
70	0.88	0.89
90	0.87	0.87
110	0.85	0.86
130	0.86	0.88
150	0.88	0.88
170	0.85	0.87
190	0.87	0.86

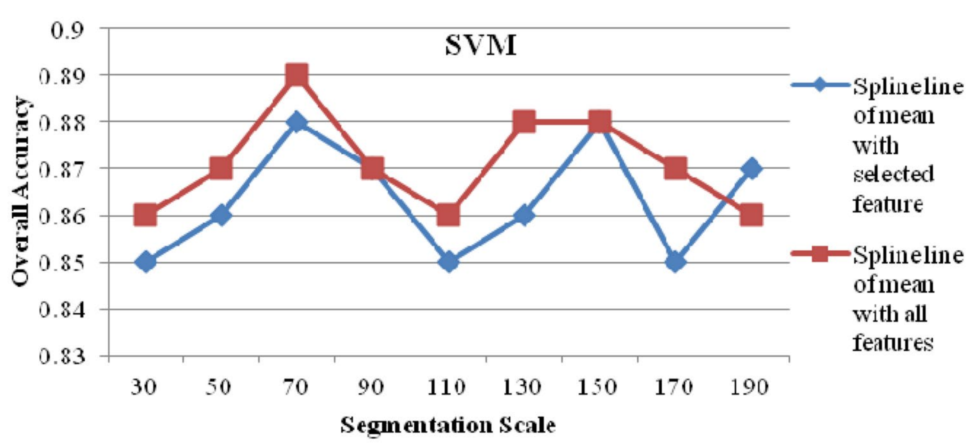

Fig. 4 Performance comparison with SVM

Table 2 Shows overall accuracy of RF classifier with spline line of mean with chosen feature and all features

RF		
Segmentation code	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.89	0.89
50	0.88	0.87
70	0.88	0.89
90	0.87	0.87
110	0.85	0.86
130	0.86	0.88
150	0.88	0.88
170	0.88	0.88
190	0.885	0.89

Fig. 5 Performance comparison with RF

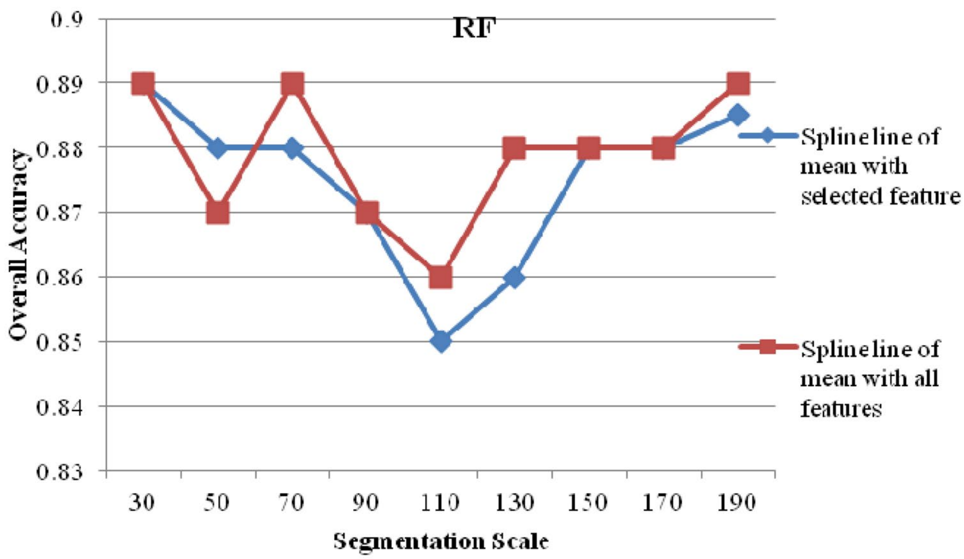

Table 3 Shows overall accuracy of adaboost.M1 classifier with spline line of mean with chosen feature and all features

Adaboost.M1		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.81	0.81
50	0.82	0.825
70	0.85	0.86
90	0.9	0.89
110	0.85	0.86
130	0.86	0.88
150	0.88	0.88
170	0.88	0.88
190	0.89	0.89

As presented in Fig. 4, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the SVM classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 2, overall accuracy of SVM classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features

As presented in Fig. 5, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the RF classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the

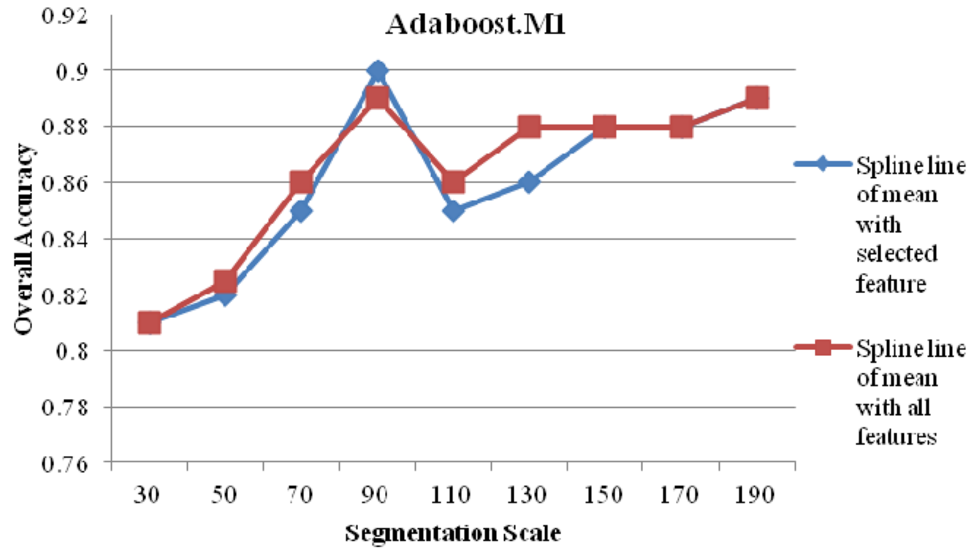

Fig. 6 Performance comparison with Adaboost.M1

Table 4 Shows overall accuracy of Naïve Bayes classifier with spline line of mean with chosen feature and all features

Naïve Bayes		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.79	0.74
50	0.8	0.76
70	0.8	0.77
90	0.82	0.79
110	0.84	0.795
130	0.83	0.798
150	0.85	0.798
170	0.88	0.8
190	0.9	0.8

overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 3, overall accuracy of adaboost.M1 classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features

As presented in Fig. 6, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the Adaboost.M1 classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 4, overall accuracy of Naïve Bayes classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features

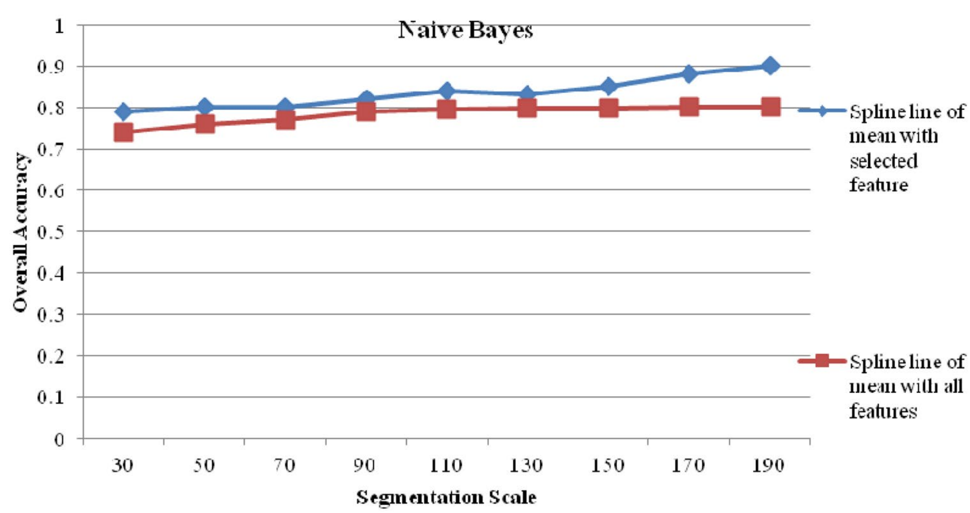

Fig. 7 Performance comparison with Naïve Bayes

Table 5 Shows overall accuracy of DT classifier with spline line of mean with chosen feature and all features

DT		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.75	0.76
50	0.78	0.78
70	0.8	0.8
90	0.83	0.84
110	0.85	0.85
130	0.87	0.86
150	0.89	0.87
170	0.9	0.88
190	0.88	0.86

As presented in Fig. 7, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the Naïve Bayes classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 5, overall accuracy of DT classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features

As presented in Fig. 8, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the DT classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

Results of area 2

The high-resolution spatial image related to agricultural environment provided in Fig. 2 (c) is used for another set of experiments. The results with different classifiers are presented in this sub section.

As presented in Table 6, overall accuracy of SVM classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features using area 2

As presented in Fig. 9, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the SVM classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 7, overall accuracy of RF classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features using area 2

As presented in Fig. 10, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the RF classifier. The results are observed with and without feature

Fig. 8 Performance comparison with DT

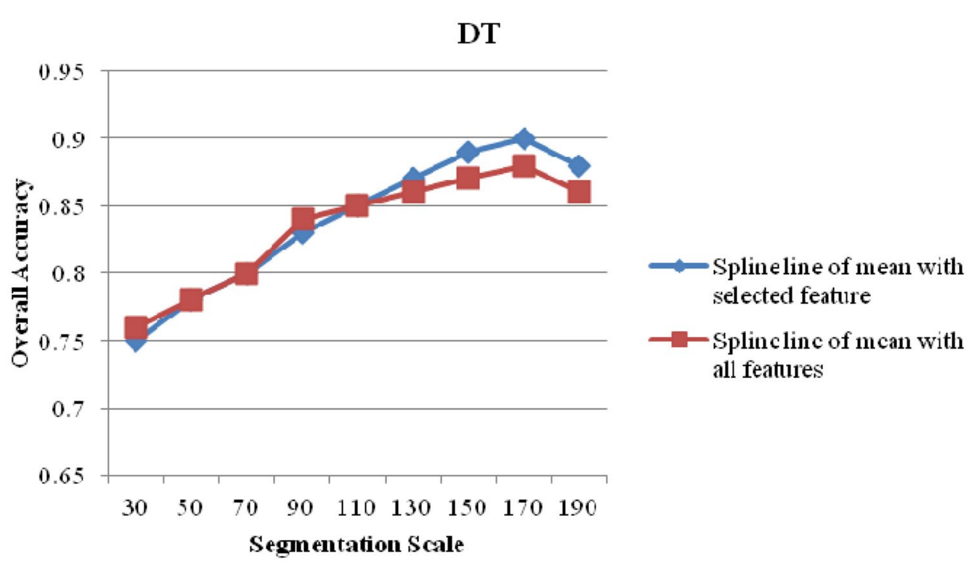

Table 6 Shows overall accuracy of SVM classifier with spline line of mean with chosen feature and all features

SVM		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.81	0.81
50	0.82	0.825
70	0.85	0.86
90	0.9	0.89
110	0.85	0.86
130	0.86	0.88
150	0.84	0.85
170	0.79	0.78
190	0.78	0.79

selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 8, overall accuracy of Adaboost.M1 classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features using area 2

As presented in Fig. 11, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the Adaboost.M1 classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 9, overall accuracy of Naïve Bayes classifier is provided against different segmentation scales.

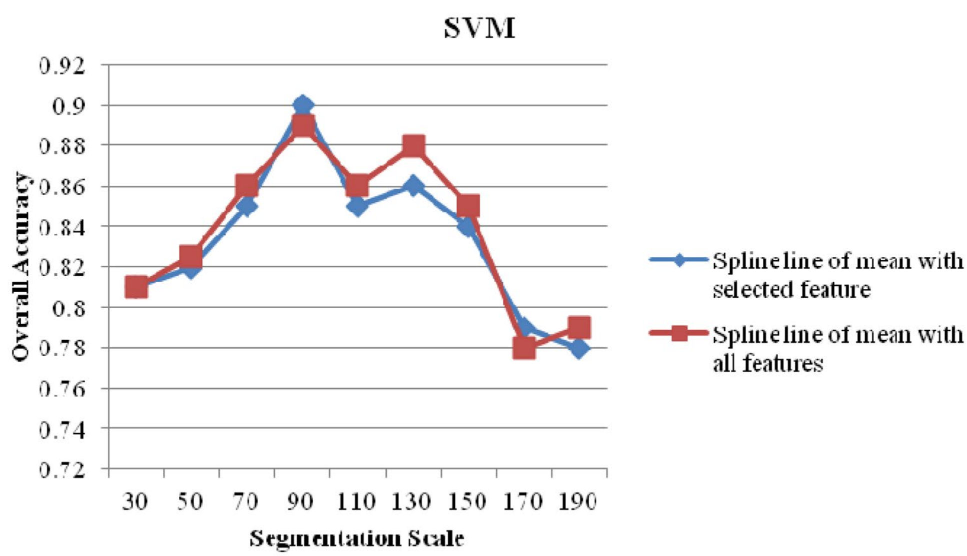

Fig. 9 Performance comparison with SVM with area 2 imagery

Table 7 Shows overall accuracy of RF classifier with spline line of mean with chosen feature and all features

RF		
Segmentation code	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.87	0.87
50	0.86	0.86
70	0.85	0.85
90	0.845	0.84
110	0.85	0.85
130	0.87	0.86
150	0.89	0.88
170	0.89	0.88
190	0.88	0.87

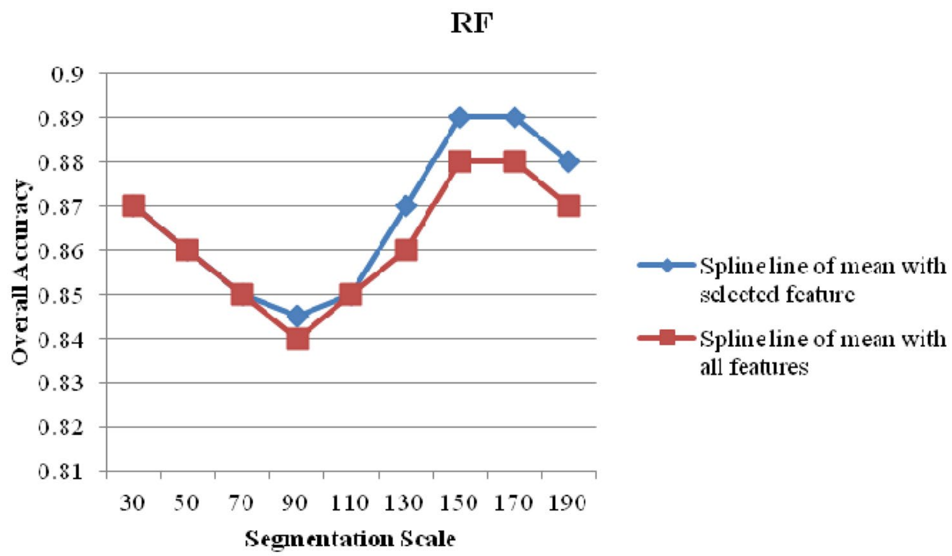

Fig. 10 Performance comparison with RF with area 2 imagery

Table 8 Shows overall accuracy of Adaboost.M1 classifier with spline line of mean with chosen feature and all features

Adaboost.M1		
Segmentation code	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.78	0.79
50	0.8	0.8
70	0.79	0.82
90	0.82	0.82
110	0.795	0.84
130	0.8	0.83
150	0.82	0.85
170	0.84	0.855
190	0.84	0.86

The results show the overall accuracy with a single feature and all features using area 2

As presented in Fig. 12, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the Naïve Bayes classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

As presented in Table 10, overall accuracy of DT classifier is provided against different segmentation scales. The results show the overall accuracy with a single feature and all features using area 2

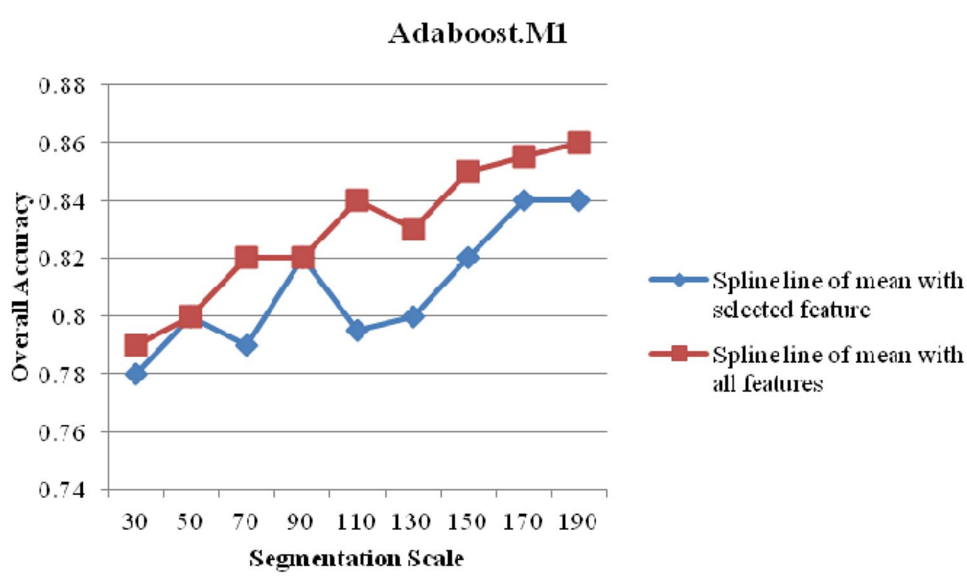

Fig. 11 Performance comparison with RF with area 2 imagery

Table 9 Shows overall accuracy of Naïve Bayes classifier with spline line of mean with chosen feature and all features

Naïve Bayes		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.81	0.79
50	0.8	0.76
70	0.8	0.77
90	0.82	0.79
110	0.84	0.795
130	0.83	0.798
150	0.85	0.798
170	0.88	0.8
190	0.9	0.8

Table 10 Shows overall accuracy of DT classifier with spline line of mean with chosen feature and all features

DT		
Segmentation Scale	Overall accuracy	
	Spline line of mean with selected feature	Spline line of mean with all feature
30	0.87	0.87
50	0.86	0.865
70	0.85	0.86
90	0.85	0.85
110	0.85	0.85
130	0.87	0.86
150	0.89	0.88
170	0.89	0.88
190	0.88	0.87

As presented in Fig. 13, the segmentation scale used from 30 to 190 incremented by 20. It is presented in horizontal axis. The vertical axis showed overall accuracy of the DT classifier. The results are observed with and without feature selection. The segmentation scale has its influence on the overall accuracy. Another important observation is that when feature selection is used, it results in improved performance.

Conclusions and future work

In this paper, we considered different factors of high-resolution images collected through remote sensing in agricultural environments for spatial data mining. In other words, the spatial imagery is subjected to supervised learning with different state of the art mining algorithms. The factors include the

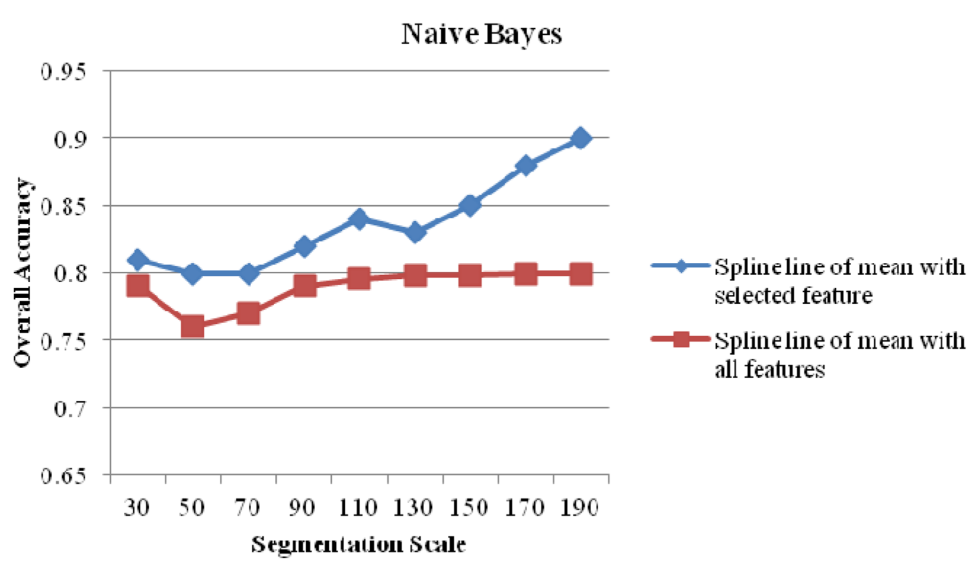

Fig. 12 Performance comparison with Naïve Bayes with area 2 imagery

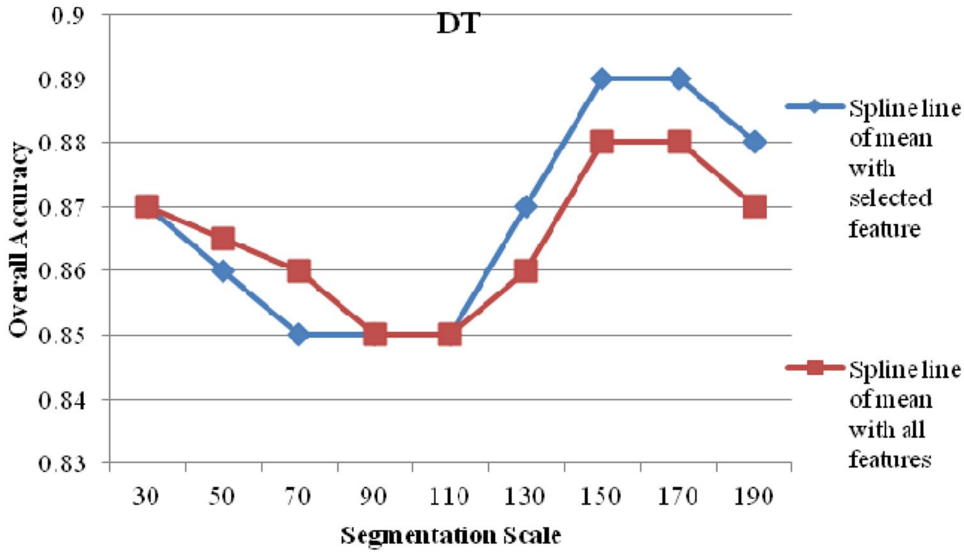

Fig. 13 Performance comparison with DT with area 2 imagery

presence of mixed objects, feature selection, size of training set and segmentation scale and the classification techniques studied are Support Vector Machine (SVM), Random Forest (RF), Naïve Bayes, k-Nearest Neighbour (KNN), Adaboost. M1 and Decision Table (DT). With respect to SVM and RF, overall accuracy changed linearly when segmentation scale is increased. Highest overall accuracy is exhibited by RF when training ratio is increased. The RF and DT were found to be very stable with and without feature selection technique. Feature selection has improved performance of Naïve Bayes but there was no improvement with Adaboost.M1. When number of mixed objects are increased in the imagery, accuracy of performance is different for each method. The imagery used for the empirical study are related agricultural environments. Feature selection using the proposed algorithm has its impact on the overall accuracy. Having used the techniques for spatial data analysis with high-resolution images, it is an important direction for future work to investigate on the low and medium resolution images. Finding trends in land cover usage is another direction for future work.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Alfaro E, Gamez M, Garcıa N (2013) adabag: an R package for classification with boosting and bagging. J Stat Softw 54(2):1–35

Arifjanov A., Akmalov S., Shodiev S., Haitov A (2021) Discussion of different remote sensing satellite possibilities for scientifical Earth observations. E3S Web of Conferences, 264, p1–11.

Atul DJ, Kamalraj R, Ramesh G, Sakthidasan Sankaran K, Sharma S, Khasim S (2021) A machine learning-based IoT for providing an intrusion detection system for security. Microprocess Microsyst 82:103741

Blanche T, Hay GJ, Kelly M, Lang S, Hofmann P, Addink E, Queiroz Feitosa R, van der Meer F, van der Werff H, van Coillie F, Tiede D (2014) Geographic object-based image analysis—towards a new paradigm. ISPRS J Photogram Remote Sens 87:180–191

Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol. Software available at. http://www.csie.ntu. edu.tw/~cjlin/libsvm/

Drăguţ L, Csillik O, Eisank C, Tiede D (2014) Automated parameterisation for multi-scale image segmentation on multiple layers. ISPRS J Photogram Remote Sens 88:119–127

Dronova I, Gong P, Clinton NE, Wang L, Fu W, Qi S, Liu Y (2012) Landscape analysis of wetland plant functional types: the effects of image segmentation scale, vegetation classes and classification methods. Remote Sens Environ 127:357–436

Duro DC, Franklin SE, Dubé MG (2012) A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5HRG imagery. Remote Sens Environ 118:259

Fassnacht FE, Hartig F, Latifi H, Berger C, Hernández J, Corvalán P, Koch B (2014) Importance of sample size: data type and prediction method for remote sensing-based estimations of aboveground forest biomass. Remote Sens Environ 154:10

Ghosh A, Joshi PK (2014) A comparison of selected classification algorithms for mapping bamboo patches in lower Gangetic plains using very high resolution WorldView 2 imagery. Int J Appl Earth Obs Geoinf 26:298–311

Lalita Parameswari DV, Seetha M, Ragha Deepika K (2014) Fuzzy Bayesian classification for spatial data streams with P-tress. Int J Adv Res Comput Commun Eng 3(6):7336–7340

Luque IF, Aguilar FJ, Álvarez MF, Aguilar MÁ (2013) Non-parametric object-based approaches to carry out ISA classification from archival aerial ortho images. IEEE J Sel Top Appl Earth Obs Remote Sens 6(4):2058–2071

Ma L, Cheng L, Han WQ, Zhong LS, Li MC (2014) Cultivated land information extraction from high-resolution unmanned aerial vehicle imagery data. J Appl Remote Sens 8(1):836731–8367325

Ma L, Cheng L, Li MC, Liu YX, Ma XX (2015) Training set size scale, and features in geographic object-based image analysis of very high resolution unmanned aerial vehicle imagery. ISPRS J Photogram Remote Sens 102:14–27

Mallikrarjuna Rao Ch, Ramesh G, Lalitha Parameswari DV, Madhavi Karanam, Sudheer Babu K (2019) Feature selection based supervised learning method for network intrusion detection. Int J Recent Technol Eng (IJRTE) 8(1):2796–2802

Pu R, Landry S (2012) A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. Remote Sens Environ 124:516–533

Radoux J, Bogaert P (2014) Accounting for the area of polygon sampling units for the prediction of primary accuracy assessment indices. Remote Sens Environ 142:9–19

Ramesh G, Mallikrarjuna Rao CH (2018) Code-smells identification by using PSO approach. Int J Recent Technol Eng (IJRTE) 7(4):323–326

Shao Y, Lunetta RS (2012) Comparison of support vector machine neural network, and CART algorithms for the land-cover classification using limited training data points. ISPRS J Photogram Remote Sens 70:78–87

Strasser T, Lang S (2015) Object-based class modelling for multi-scale riparian forest habitat mapping. Int J Appl Earth Obs Geoinf 37:29–37

Tehrany MS, Pradhan B, Jebuv MN (2014) A comparative assessment between object and pixel-based classification approaches for land use/land cover mapping using SPOT 5 imagery. Geocarto Int 29:351–369

Tian Y, Yang C, Huang W et al (2021) Machine learning-based crop recognition from aerial remote sensing imagery. Front Earth Sci 15:54–69. https://doi.org/10.1007/s11707-020-0861-x

Whiteside TG, Maier SW, Boggs GS (2014) Area-based and location-based validation of classified image objects. Int J Appl Earth Obs Geoinf 28:117–130

Xu LL, Li J, Brenning A (2014) A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. Remote Sens Environ 141:14–23

Yang J, Wang Y (2012) Classification of 10m-resolution SPOT data using a combined Bayesian network classifier-shape adaptive neighbourhood method. ISPRS J Photogram Remote Sens 72:36–45

Zhang XL, Feng XZ, Xiao PF, He GJ, Zhu LJ (2015) Segmentation quality evaluation using region-based precision and recall measures for remote sensing images. ISPRS J Photogram Remote Sens 102:73–84

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.